On Modeling the Double and Multiplicative Binomial Models as Log-Linear Models

Bayo H Lawal*
Department of Statistics & Mathematical Sciences, Kwara State University, Malete, Kwara State, Nigeria

Abstract
In this paper we have fitted the double binomial and multiplicative binomial distributions as log-linear models using sufficient statistics. This approach is not new as several authors have employed this approach, most especially in the analysis of the Human sex ratio in [1]. However, obtaining the estimated parameters of the distributions may be problematic, especially for the double binomial where the parameter estimate of \( \pi \) may not be readily available from the Log-Linear (LL) parameter estimates. Other issues associated with the LL approach is its implementation in the generalized linear model with covariates. The LL uses far more parameters than the procedure that employs conditional log-likelihoods functions where the marginal likelihood functions are minimized over the parameter space. This is the procedure employed in SAS PROC NLMIXED. The two procedures are essentially equivalent for frequency data. For models with covariates, the LL uses far more parameters and the marginal likelihood functions approach are employed here on three data set having covariates.

Keywords
Double Binomial; Multiplicative Binomial; Log-Linear; Marginal Likelihood Functions

Introduction
In the formulations of the multiplicative binomial distribution, in Altham and its corresponding double binomial distribution in Efron, both distributions were characterized with intractable normalizing constants \( c(n, \psi) \) and \( c(n, \pi) \) respectively [2,3]. Consequently, these models were implemented in by utilizing a generalized linear model with a Poisson distribution and log link to the frequency data. This approach has earlier being similarly employed in [1,4]. This approach which employs joint sufficient statistics in both distributions was earlier proposed in Lindley & Mersch [5].

Both distributions are fitted using a Poisson regression model having sufficient statistics from both distributions as explanatory variables with the frequencies being the mean dependent variables. For the Double binomial model (DBM), the sufficient statistics are \( y \log(y) \) and \( (n − y) \log(n − y) \). Similar joint sufficient statistics for the multiplicative binomial model (MBM) are \( y \) and \( y(n − y) \) with the offset being

\[
Z = \log \left( \frac{n}{y} \right) \quad \text{for both models. For instance, for the DBM the model would be:}
\]

\[
\log \left( \frac{n}{y} \right) = y + \theta
\]

Where \( \theta = \theta_1 + \theta_2 \)

\[
\begin{align*}
\theta_1 &= \begin{cases} 
0 & \text{if } y = 0 \\
y \log(y) & \text{otherwise}
\end{cases} \\
\theta_2 &= \begin{cases} 
0 & \text{if } y = n \\
n − y \log(n − y) & \text{otherwise}
\end{cases}
\end{align*}
\]

Similarly, for the multiplicative binomial, the model is estimated by the log-linear model (or Poisson Model):

\[
\log \left( \frac{n}{z} \right) = y + \delta
\]

Where, \( \delta = \delta_1 + \delta_2 \) and

\[
\begin{align*}
\delta_1 &= \begin{cases} 
0 & \text{if } y = 0 \\
y & \text{otherwise}
\end{cases} \\
\delta_2 &= \begin{cases} 
0 & \text{if } y = n \\
y(n − y) & \text{otherwise}
\end{cases}
\end{align*}
\]

However, in recent times, both models have been fully formulated with the intractable...
normalizing constants fully formulated. These distributions are described in this paper with the normalizing constants fully formulated. In this study, we compare fitting these two probability models to two example frequency data, two data set examples arising from teratology studies, and randomized complete block design example having binary outcome by the method of sufficient statistics described above and by the method of numerically maximizing the marginal likelihood function arising from engaging the problem as a mixed generalized linear model. The SAS PROC NLMIXED which performs the Maximum Likelihood estimation numerically by using the Adaptive Gaussian Quadrature and Newton-Raphson optimization algorithm. We shall designated the sufficient Statistics-Poisson regression approach as LL, while the marginal likelihood function maximization via PROCNLIMIXED is designated Mgl in this study. The sufficient statistics procedure uses a Poisson regression with an offset and is implemented in SAS PROC GENMOD.

Models Under Consideration

We describe in the following section, the two probability distribution models employed in this paper.

The Multinomial Binomial Model-MBM

[26,7] Lovinson proposed an alternative form of the two-parameter exponential family generalization of the binomial distribution first introduced by [2] which itself was based on the original Cox representation as:

\[ f(y) = \frac{n!}{y!(n-y)!} \pi^{y}(1-\pi)^{n-y} \omega^{(n-y)}, y=0,1,\ldots,n \]  

(3)

where \(0 < \psi < 1\) and \(\omega > 0\). When \(\omega = 1\) the distribution reduces to the binomial with \(\pi = \psi\). If \(\psi = 1, n \to \infty\), and \(\psi = 0\), then \(\psi = 1\) and the MBD reduces to Poisson(\(\mu\)).

The normalizing constant is

\[ c(n,\psi) = \sum_{y=0}^{n} \frac{n!}{y!(n-y)!} \pi^{y}(1-\pi)^{n-y} \omega^{(n-y)}, \]

the denominator expression in (3) in this case. [8] presented an elegant characteristics of the multiplicative binomial distribution including its four central moments. His treatment includes generation of random data from the distribution as well as the likelihood profiles and several examples-some of which are similarly employed in this presentation. Following [8] the probability \(P(Y)\) of success for the Bernoulli trial, that is, \(P(Y=1)\) can be computed from the following expression in (4) as:

\[ p_i = \psi_i \frac{k_{+i}(\psi, \omega)}{K_n(\psi, \omega)} \text{, for } i = 1 \]

(4)

Where:

\[ k_{+i}(\psi, \omega) = \sum_{j=i}^{n} \frac{n-a}{y} \pi^j(1-\pi)^{n-j} \omega^{j-n+\psi i} \text{, for } a = 1,2,\ldots,n \]

(5)

with \(p\) defined as in (4). \(\Psi\) therefore can be defined as the probability of success weighted by the intra-units association measure \(\omega\) which measures the dependence among the binary responses of the \(n\) units. Thus if \(\omega = 1\) then \(p = \Psi\) and we have independence among the units. However, if \(\omega = 1\), then \(p = \Psi\) and the units are not independent.

The mean and variance of the LMPD are given respectively as:

\[ E(Y) = np_i \]

(6a)

\[ \text{Var}(Y) = np_i (n+1-p_i) \]

(6b)

The Double Binomial (DBM) Model

In Feier et al. [7], the double binomial distribution was presented, having the pdf form:

\[ f(y;\pi,\phi) = \sum_{j=0}^{y} \binom{n}{j} \pi^j(1-\pi)^{n-j} \phi^j(1-\phi)^{n-j} \]

(6b)

Again, the normalizing constant in this case is the denominator expression given by

\[ c(n,\psi) = \sum_{y=0}^{n} \binom{n}{y} \pi^y(1-\pi)^{n-y} \phi^{n-y}(1-\phi)^y \]

Applications

We apply the models discussed above to two frequency data and to teratology data sets having four and two treatment groups. We first present the analyses for the two frequency data sets in Tables 1 through 5. The estimation of the parameters under each model for the Mgl approach uses SAS PROC NLMIXED, using the following log-likelihoods for the MBM (LL1) and DBM (LL2) respectively. The procedure was discussed earlier in the paper.

\[ LL1 = \log \left[ \sum_{i=1}^{n} \left( \frac{y_i}{\pi_i} \phi \right) \right] \]

\[ LL2 = \log \left[ \sum_{i=1}^{n} \left( \frac{y_i}{\pi_i} \phi \right) \right] \]

Example Data Set I-Geissler Data

The Distribution of males in 6115 families with 12 children in Saxoy, previously analyzed in Sokal & Rohlf [10] is presented in Table 1. The data is originally from Geissler [11] and had similarly been analyzed in [12]. Here \(Y \sim \text{binomial} (12, \pi)\). The frequencies are presented as counts having a total sum of 6115. The observed mean for the data is \(\bar{Y} = 6.2306\) and the corresponding variance is \(s^2 = 3.4498\). Under the binomial model, the estimated mean is 6.2304 and estimated variance is 2.9956. Hence the estimated dispersion parameter \(DP = s^2 / \bar{Y} = 2.07\) indicating over-dispersion in the data. The estimated probability of occurrence under the binomial model is \(\hat{\pi} = 0.5165\). The binomial does not fit the data \((X^2 = 110.5051\) on 11 d.f., \(p\)-value=0.0000) because the variance of the data is grossly under estimated by the model. The results of the application of the double binomial and the multiplicative models to this data are presented in Table 2. Further, the mixed model approach is based on one more degree of freedom as it estimates one parameter less than the LL model approach. The Mixed model approach gives the parameter estimates of the distribution. We can obtain the equivalent parameters estimates from the Log-linear (LL) approach for the multiplicative models as follows:

\[ \hat{\alpha} = \alpha \times p\hat{\phi} = \alpha \times p(-0.02615) = 0.9742, \hat{\pi} = 1 / (1 + \alpha \times p(-0.02615)) = 0.5165 \]

(8)

For the DB, \(\hat{\phi}\) can equivalently be obtained as \(\hat{\phi} = 1 + 0.140205 = 0.8598\), but the estimated probability \(\hat{\pi}\) seems intractable in this case and no equivalent solution is available in this case. We may note here that the estimate \(\hat{\pi} = 0.5165\) under the multiplicative model is not an estimate of the success probability \(\pi\). For this data, we must use the expressions in (4) and (5) to obtain this estimate. Here, \(\hat{K}(q-1) = 0.42723\) and \(\hat{K}(q-1) = 0.42499\). Consequently, \(\hat{\pi} = \hat{\psi} \frac{k(q-1)}{k(q-1)} = 0.5165 \left( \frac{0.42723}{0.42499} \right) = 0.5192\).

The mean and Variance can therefore be computed respectively from (6a) and (6b). Alternatively, the means and variances can be empirically obtained from the fitted models using the elementary principles of

\[ E(Y) = \sum_{i=0}^{n} Y_i \hat{p}_i \]
Table 1: Distribution of Males in 6115 families with 12 children

| Count | 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 |
|--------|----|----|----|----|----|----|----|----|----|----|----|----|----|
|        | 3  | 24 | 104| 286| 670| 1033| 1343| 1112| 829| 478| 181| 45 | 7  |

Table 2: Parameter estimates under the five Models

| Y     | Count | Double Binomial | Multiplicative Binomial | Marginal Likelihood |
|-------|-------|-----------------|-------------------------|--------------------|
|       |       | p̂_i            | η̂_i p̂_i                | V̂_i               |         | η̂_i p̂_i               | η̂_i p̂_i               | V̂_2               |         |
| 0     | 3     | 0.0005          | 0.0000                   | 0.0000             | 0.0000  | 0.0004                   | 0.0004                   | 0.0000             | 0.0000  |
| 1     | 24    | 0.0038          | 0.0043                   | 0.0038             | 0.0038  | 0.0037                   | 0.0041                   | 0.0037             | 0.0037  |
| 2     | 104   | 0.0171          | 0.0214                   | 0.0179             | 0.0179  | 0.0171                   | 0.0212                   | 0.0179             | 0.0179  |
| 3     | 286   | 0.0503          | 0.0717                   | 0.1869             | 0.5250  | 0.4993                   | 0.7192                   | 0.1905             | 0.4936  |
| 4     | 670   | 0.1068          | 0.1785                   | 0.6160             | 2.2333  | 1.8538                   | 2.1793                   | 0.6194             | 2.2454  |
| 5     | 1033  | 0.1698          | 0.3483                   | 1.4652             | 6.4792  | 4.3325                   | 0.1694                   | 0.3487             | 1.4665  |
| 6     | 1343  | 0.2067          | 0.5550                   | 2.7056             | 13.9220 | 6.6015                   | 0.2057                   | 0.5544             | 2.7008  |
| 7     | 1112  | 0.1938          | 0.7488                   | 4.0622             | 23.4178 | 6.9164                   | 0.1933                   | 0.7478             | 4.0542  |
| 8     | 829   | 0.1390          | 0.8878                   | 5.1743             | 32.3146 | 5.5144                   | 0.1396                   | 0.8874             | 5.1712  |
| 9     | 478   | 0.0748          | 0.9626                   | 5.8472             | 38.3710 | 4.1810                   | 0.0755                   | 0.9629             | 5.8511  |
| 10    | 181   | 0.0289          | 0.9915                   | 6.1364             | 41.2628 | 3.6074                   | 0.0291                   | 0.9920             | 6.1418  |
| 11    | 45    | 0.0074          | 0.9989                   | 6.2178             | 42.1579 | 3.4972                   | 0.0071                   | 0.9992             | 6.2204  |
| 12    | 7     | 0.0011          | 1.0000                   | 6.2306             | 42.3116 | 3.4915                   | 0.0008                   | 1.0000             | 6.2306  |

Table 3: Empirical Means and Variances for both the DBM and MBM

Var of Y being E[Y]-[E(Y)]^2. These distributions are displayed in Table 3.

In the above Table, Some of the columns are self explanatory. The columns labeled V̂_i and V̂_2 are cumulative values of

\[ \sum y^2 \hat{p}_i \left( \sum y \hat{p}_i \right)^2 \]

for both models respectively. Thus, the mean is the value at y = 12.

The variance for the DBM for instance, is computed as 42.3116−(6.2306)^2 = 3.4915.

In Table 4 are presented the expected values under both models for the two approaches (LL & Mgl), both approaches give exact results as expected. The Table also displays the mean of the distributions under both approaches as well as the empirical variances, designated here as var. We recall that for the observed data in Table 1, \( \hat{y} = 6.2306 \) and \( s^2 = 3.4898 \). We see from Table 4, that while the two models estimate the mean of the data well, the estimated variance under the binomial model of 12(0.5192)(1−0.5192) = 2.9956 understimates the observed variance of the data, and this explains the poor fit to the data by the binomial model. On the other hand, for the two models, the variance of the observed data is reasonably well estimated, because of the extra parameter in the model (dispersion parameter) of \( \phi \) and \( \omega \) for the DBM and MBM respectively.

Table 4 also displays the corresponding Pearson's X^2 and the corresponding degrees of freedom (d.f.). Clearly, for this data set, both the double binomial and the multiplicative models fit the data well with the Double binomial being slightly providing a better fit. Although the expected values generated are the same for both fitting approaches, we see that the marginal likelihood (Mgl) approach gives a more parsimonious model because it is based on one more degree of freedom.

Data Example II

This example is taken from Nelder & Mead [13] and relates to the number of candidates having an "alpha", i.e. at least 15 scores out of a total 20 points from each of nine questions employed in assessing the final class of candidates in an examination. There were a total of 209 candidates for the exam and Table 5 gives the distribution of these scores for the 209 candidates.
### Table 4: Expected values under the two models and Approaches with corresponding Pearson’s $X^2$ Statistic values

| Y   | Count | DBM MgL | LL MgL | DBM LL | LL MgL |
|-----|-------|---------|--------|--------|--------|
| 0   | 3     | 2.956   | 2.956  | 2.3486 | 2.3487 |
| 1   | 24    | 23.3861 | 23.386 | 22.5809| 22.581 |
| 2   | 104   | 104.3138| 104.3137| 104.8482| 104.8484|
| 3   | 286   | 307.7531| 307.7531| 310.8921| 310.8923|
| 4   | 670   | 652.8854| 652.8858| 655.6551| 655.6551|
| 5   | 1033  | 1038.546| 1038.546| 1036.077| 1036.077|
| 6   | 1343  | 1264.242| 1264.242| 1257.907| 1257.907|
| 7   | 1112  | 1185.039| 1185.04  | 1182.293| 1182.293|
| 8   | 829   | 850.0639| 850.0639| 853.7724| 853.7724|
| 9   | 478   | 457.2186| 457.2188| 461.9646| 461.9659|
| 10  | 181   | 176.8358| 176.8359| 177.7841| 177.785 |
| 11  | 45    | 45.2369 | 45.2369 | 43.6925 | 43.6928 |
| 12  | 7     | 6.5246 | 6.5246 | 5.1058 | 5.1858 |

### Results

The results of applying both models DBM and MBM to the data using both approaches (LL and MM) are presented in Table 5. Again, both approaches lead to the same results in terms of expected values. However, the MgL models have one more degree of freedom under both models than the LL approach. Again, to get equivalent parameter estimates from the LL model, we have, for the multiplicative model, \( \hat{\phi} = \exp(\hat{\phi}) = \exp(-0.2168) = 0.8051 \). As discussed earlier, the corresponding estimate for \( \pi \) is not readily available. Thus in this study, we will employ the alternative MgL procedure that utilizes PROC NLMIXED in SAS. One advantage of this is that it will allow for more degrees of freedom than the log-linear model. We see for the frequency data in Tables 1 for instance, that the LL approach is based on 1 d.f. more than the MgL model based.

### Regression Model Formulations

When there are covariates in our data, the sufficient statistic approach here into referred to as Log-linear (LL) does not lend itself to easier formulation and implementation. Lindsey and Altham [13] employed this approach to fitting amongst others, the two models considered in this study to the distribution of males in families in Saxony during 1885-1976 (the human sex ratio data). This approach employs far too many parameters, 13 to be precise when the same group of models can be implemented with only four parameters with the same results. Further, the implementations under this approach are not readily available. Thus in this study, we will employ the alternative MgL procedure that utilizes PROC NLMIXED in SAS. One advantage of this is that it will allow for more degrees of freedom than the log-linear model. We see for the frequency data in Tables 1 for instance, that the LL approach is based on 1 d.f. more than the MgL model based.

### Example I: Teratology-Ossification on the Phalanges

Teratology is the study of abnormalities of physiological development. The offspring of animals that were exposed to a toxin during pregnancy are studied for malformation. The number of malformed offspring in a litter of size \( n \) is not typically distributed binomial because the responses of the offspring from the same litter are not independent; hence their sum does not constitute a binomial r.v. Thus, data in teratological studies exhibit over-dispersion because of the correlation among responses from off springs in the same litter. If in the same litter, the probability of fetal death is modeled with the logit link viz:

\[
\log \left( \frac{p_{ij}}{1-p_{ij}} \right) = \beta_0 + \beta_2 z_{2i} + \beta_3 z_{3i} + \beta_4 z_{4i} \tag{9}
\]

We have assumed here that \( \beta_0 \) is similar across litters, and that,
Results

From the results in Table 6, the two cases (II & III) with variable dispersion parameters fit better than the model in case I, where the dispersion is uniform across the four groups. Of the models in Cases II and III, the models in case III fits much better than those in case II. Case II models assume that the four groups have a common dispersion parameter. That is, $\phi_0$ is the same for all groups, and the dispersion parameters $\phi_i$ for the four groups are functions of the covariates in the control and other treatment groups.

Here, $p_{ij} = \frac{1}{1 + e^{-\phi} p(-\beta_0)}$, $\phi = e^{\beta_0}$ and $\phi = \exp(c_{ij})$ with $a_{ij} \neq c_{ij}$. These ensure that the dispersion parameters are positive.

The model here has $p_i = p_{i,j}$ and the dispersion parameters are functions of the covariates. That is, $\phi = \exp(a_{ij} + a_{ij} z_i + a_{ij} z_j)$ and $\phi = \exp(c_{ij} + c_{ij} z_i + c_{ij} z_j)$.

The model here has $p_i \neq p_j$ with the ps modeled as in (9) and the dispersion parameters are modeled as functions of the covariates as in the preceding case.

Results

From the results in Table 6, the two cases (II & III) with variable dispersion parameters fit better than the model in case I, where the dispersion is uniform across the four groups. Of the models in Cases II and III, the models in case III fits much better than those in case II. Case II models assume that the four groups have a common estimated probability $\pi$, which are estimated respectively as 0.2125 and 0.2158 in both the DBM and MBM. However, the models in III which assume heterogeneous success probabilities across the four groups and variable dispersion parameters (that are functions of the covariates) fit better than those in case II. The DBM here is based on $X^2 = 115.1333$ on 72 d.f. The estimated $\pi$s under the MBM are functions of $n$, hence these values are different for different $n$ in the final model (Case III). We may note here that the $\Psi$s should not be mistaken for the success probabilities.

Data Example II-Trout Egg Data

The data in Table 7 from Manly [16] relate to the number of surviving eggs from boxes of eggs that were buried at five different locations in a stream and at four different times a box from a location was sampled. The data is presented as $y/n$ where $y$ is the number surviving and $n$ is the number of eggs in the box.

The model of interest here is:

$$\log i t(p_{ij}) = \beta_0 + \sum_{k=1}^{4} \beta_k x_i + \sum_{l=1}^{3} \beta_{l} x_i$$

where $x_i$ are four dummy variables for location effects, and $x_i$ are three dummy variables representing the Time effects. The structure here is that of a randomized block design having locations as blocks and Survival times as treatments. Thus, the structure of the Pearson’s $X^2$ would be for Location (L) and Survival time (T):

| Source | d.f. |
|--------|------|
| L|T | 4 |
| T|L | 3 |
| Residual | 12 |

The degree of freedom of 12 refers only to the binomial model. For all other distributions, the d.f. must account for the additional dispersion parameter estimates. Under the Binomial model $X^2 = 63.9639$ on 12 d.f. giving an estimated dispersion parameter of 5.3303 > 1, indicating that the data is highly overdispersed.

Because of the overdispersion in the data, we now apply our models, DBM and the MBM to the data, giving the results in Table 8.
Table 6: Parameter estimates for the Models in all the Cases

| Location in stream | Survival Period (weeks) | Case I | Case II | Case III |
|--------------------|-------------------------|--------|---------|----------|
|                    | 4                       | 7      | 8       | 11       |
| 1                  | 89/94                   | 94/98  | 77/86   | 141/155  |
| 2                  | 106/108                 | 91/106 | 87/96   | 104/122  |
| 3                  | 119/123                 | 100/130| 88/119  | 91/125   |
| 4                  | 104/104                 | 80/97  | 67/99   | 111/132  |
| 5                  | 49/93                   | 11/113 | 18/80   | 0/138    |

Table 7: Number of Surviving eggs against number of eggs in a box

Models in (A) fit both the double binomial and the multiplicative binomial with constant dispersion parameter. For this group of models, the multiplicative binomial performs much better with constant dispersion parameter of 0.9884, very close to 1, indicating there is partial independence in the data ignoring the effects of locations. Models in B, have variable dispersion parameters that are functions of the covariates (Time), that is, \( \phi = \exp(a_0 + a_1 \times x) \) and \( \omega = \exp(c_0 + c_1 \times x) \). Under this formulation, the double binomial computation does not converge, but that of the multiplicative binomial converged. This model gives a Pearson \( X^2 \) of 5.0120 on d.f.

The results of this final model are presented in Table 9. Note that for the multiplicative, the estimated probabilities of success \( \pi \) which are not the same as the \( \phi \) in the model formulation in (3) are computed using expressions in (4) and (5). Note that \( \psi = \pi \). The column labeled \( \sum X^2 \) gives the cumulative contributions of observation towards \( X^2 \). The value 5.0120 is the sum of all 20 contributions towards \( X^2 \). Under the final multiplicative model, the estimated average probabilities of surviving in the first 4, 7, 8 and 11 weeks are respectively \( \{0.8854, 0.8799, 0.8656\} \).
Table 8: Results of Analysis of Data in Table 7.

| # | n  | y     |  \( \psi \) |  \( \pi_1 \) |  \( \pi_2 \) |  \( m_i \) |  \( \omega_i \) | \( \sum x^2 \) |
|---|----|-------|-------------|-------------|-------------|------------|-------------|--------------|
| 1 | 94 | 89    | 0.9895      | 0.9863      | 0.9727      | 92.7093    | 1.003       | 1.2635       | 0.1484       |
| 2 | 98 | 94    | 0.9042      | 0.9059      | 0.8207      | 88.7817    | 0.9997      | 8.3863       | 0.4551       |
| 3 | 86 | 77    | 0.9226      | 0.8714      | 0.7592      | 74.9446    | 1.009       | 8.2376       | 0.5115       |
| 4 | 155| 141   | 0.7915      | 0.9327      | 0.87        | 144.5644   | 0.9903      | 12.016       | 0.5994       |
| 5 | 108| 106   | 0.9868      | 0.9821      | 0.9645      | 106.0685   | 1.003       | 1.8759       | 0.5994       |
| 6 | 106| 91    | 0.8823      | 0.8844      | 0.7822      | 93.7495    | 0.9997      | 10.894       | 0.6801       |
| 7 | 96 | 87    | 0.9044      | 0.8413      | 0.7076      | 80.7662    | 1.009       | 10.4533      | 1.1612       |
| 8 | 122| 104   | 0.7509      | 0.8805      | 0.7755      | 107.4158   | 0.9903      | 17.1125      | 1.0626       |
| 9 | 123| 119   | 0.9736      | 0.9633      | 0.928       | 118.4912   | 1.003       | 4.2347       | 1.272        |
| 10| 130| 100   | 0.787       | 0.7902      | 0.6244      | 102.7271   | 0.9997      | 21.7874      | 1.3444       |
| 11| 119| 88    | 0.8235      | 0.7383      | 0.5447      | 87.8609    | 1.009       | 16.3371      | 1.3446       |
| 12| 125| 91    | 0.5977      | 0.7114      | 0.5077      | 88.9207    | 0.9903      | 50.801       | 1.3933       |
| 13| 104| 104   | 0.9782      | 0.9711      | 0.943       | 100.9938   | 1.003       | 2.8698       | 1.4827       |
| 14| 97 | 80    | 0.8183      | 0.8206      | 0.6734      | 79.5995    | 0.9997      | 14.3821      | 1.4848       |
| 15| 99 | 67    | 0.8504      | 0.7776      | 0.6042      | 76.9785    | 1.009       | 13.1452      | 2.7782       |
| 16| 132| 111   | 0.6442      | 0.7911      | 0.6268      | 104.4282   | 0.9903      | 37.8028      | 3.1918       |
| 17| 93 | 49    | 0.5274      | 0.5241      | 0.2743      | 48.7372    | 1.003       | 20.3948      | 3.1932       |
| 18| 113| 11    | 0.1006      | 0.0986      | 0.0097      | 11.1422    | 0.9997      | 10.9944      | 3.1932       |
| 19| 88 | 18    | 0.1238      | 0.1869      | 0.0346      | 16.4498    | 1.009       | 10.8223      | 3.4141       |
| 20| 138| 0     | 0.0431      | 0.0121      | 0.0001      | 1.6708    | 0.9903      | 1.7055       | 5.012        |

Table 9: Parameter estimates under the multiplicative model with variable dispersion parameter.

The results of these models are presented in Table 10.

**Appendix A. SAS Program for the Example**

```sas
options nodate nonumber ls=85 ps=66;
data ex1;
do L=1 to 5;
do T=1 to 4;
input y n @@;
output;
end;
end;
datalines;
89 94 98 77 86 141 155
106 108 91 106 87 96 104 122
119 123 100 130 88 119 91 125
104 104 80 97 67 99 111 132
49 93 11 113 18 88 0 138
;run;
proc print;
run;
/*generate indicator variables for Location*/;
data w1;
array x(5) z1-z5;
do j=1 to 5;
  if j=L then x(j)=1;
  else x(j)=0;
end;
drop j;
run;
data w2;
array d(4) x1-x4;
do k=1 to 4;
  if k=T then d(k)=1;
  else d(k)=0;
end;
drop k;
run;
data new;
merge w1 w2;
/*generate indicator variables for Time*/;
run;
set ex1;
array x(5) z1-z5;
do j=1 to 5;
  if j=L then x(j)=1;
  else x(j)=0;
end;
drop j;
run;
data w2;
set ex1;
array d(4) x1-x4;
do k=1 to 4;
  if k=T then d(k)=1;
  else d(k)=0;
end;
drop k;
run;
data new;
merge w1 w2;
```

run;
proc sort data=new;
by T;
run;
proc nlmixed data=new tech=newrap maxit=2000;
parms b0=-0.1 b1=1.1 b2=0.4 b3=.1 b4=0.2 s1=s3=0.0 a0=0 a1=0 a2=0 a3=0;
lp=b0+b1*z1+b2*z2+b3*z3+b4*z4+s1*x1+s2*x2+s3*x3;
lr=a0+a1*x1+a2*x2+a3*x3;
omega=exp(lr);
p=1/(1+exp(-lp));
sum=0.0;
do j=0 to n;
z1=lgamma(n+1)-lgamma(j+1)-lgamma(n-j+1);
 u1=zz1+ j*log(psi) + (n-j)*log(1-p) + j*(n-j)*log(omega);
suma=suma+exp(u1);
end;
keep suma;
zz2=lgamma(n+1)-lgamma(y+1)-lgamma(n-y+1);
L2=z2+y*log(psi) + (n-y)*log(1-p) + y*(n-y)*log(omega)-log(sum);
model y~general(L2);
predict p out=aa;
predict omega out=bb;
run;
Ods rtf close;
format psi p1 p2 exp omega var xx LRT Wald 10.4;
proc print data=qq4;
var=n*p1+(n*(n-1)*p2)-(n*n*p1*p2);
/* Generate Wald, LRT and Pearson's GOFs */
wald=(y-exp)**2/var;
if y=0 then lrt=0;
else lrt+2*y*log(y/exp);
XX+((y-exp)**2)/exp;
run;
proc print data=new;
by T;
proc sort data=new;
by T;
run;
merge q1 q2;
data qq4;
run;
where T=1;
/* Generate Wald, LRT and Pearson's GOFs */
var=n*p1+(n*(n-1)*p2)-(n*n*p1*p2);
/* Generate Wald, LRT and Pearson's GOFs */
wald=(y-exp)**2/var;
if y=0 then lrt=0;
else lrt+2*y*log(y/exp);
XX+((y-exp)**2)/exp;
run;
proc print data=qq4;
var=n*y psi p1 p2 exp omega var xx LRT Wald;
format psi p1 p2 exp omega var xx LRT Wald 10.4;
run;

Results
Results from Table 10 show that for cases I to III, the model for case II is the most parsimonious. The difference between the Likelihood-test statistic, G^2 between models II and III being 0.2307 on 1 d.f (p-value=0.6310), which is not significant. We have used the G^2 rather than the Wald or Pearson's X^2 because only the G^2 statistic has the partitioning property, (see, [18]). However, while this model seems the best, it does not tell us much about the probability of success (\( \pi_i \), i = 0, 1) for each group. The model assumes a common dispersion parameter and constant across each treatment, the estimated probabilities and other variables under the multiplicative model for control and treatment groups. We present in Table 11 the estimated probabilities of fetal deaths for the control and experimental groups on 16 d.f. Here, under the double binomial model, the estimated probabilities of fetal deaths for the control and experimental groups are significant in model III. Thus, a reduced model of case III seems the best, it does not tell us much about the probability of success (\( \pi_i \)), which models the probability of success separately for the treatments but assumes a common dispersion parameter. The models are based on 16 d.f. Here, under the double binomial model, the estimated probabilities of fetal deaths for the control and experimental groups are respectively 0.0552 and 0.2332 and these estimated probabilities are constant across the treatment levels. The corresponding goodness-of-fit values are G^2 = 9.1437 and X^2 = 22.9671 with common dispersion parameter estimate being \( \hat{\phi} = 0.4900 \). We notice a considerable discrepancy between the values of G^2 and X^2 for this data here. This is because, of the twenty observations in the data, nine of them have zeros for the values of Y. Consequently, these observations do not contribute to the overall G^2 and this accounts for the lower values of G^2 compared to their corresponding X^2.

For the MBM, while the estimated \( \psi^* \) are specific to each treatment and constant across each treatment, the estimated probabilities \( \pi_i \) of successes vary by the number of litters n as outlined in expression (4). Thus for n = 8, \( \pi_i \) equals 0.0769 and 0.2469 respectively for the control and treatment groups. We present in Table 11 the estimated probabilities and other variables under the multiplicative model for this case.

We may note here that for this data, we have also computed the Wald’s test Statistic and it seems to give the lowest value of 23.2406. The GOF values are cumulated so that the last values give the sums over all observations.

Conclusion
Results presented in the preceding sections showed that while it is relatively easier to fit both the double binomial and the multiplicative binomial with joint sufficient statistics employing Poisson regression for frequency data, this approach cannot easily be implemented with data having co-variates. Further, the sufficient statistics approach is based on more degrees of freedom than the MGL method, which makes the MGL method more parsimonious in all cases. We would encourage the use of the MGL methods in applications of these models.
### Table 10: Parameter estimates under the three cases for the two probability models.

| Parameters | Case I | Case II | Case III |
|------------|--------|---------|----------|
| MLE Est.   |        |         |          |
| $\hat{\pi}$ | 0.1269 | 0.1086  | 0.1269   |
| $\hat{\psi}$ | 0.3033 | 0.5467  | 0.1269   |
| $\phi$ | 0.3648 | 0.8314  | 0.1772   |
| $\omega$ | 0.8035 | 0.8035  | 0.8035   |

| -2LL   | 60.3121 | 63.5982 | 63.5982 |
| AIC    | 64.3121 | 67.5982 | 67.5982 |
| $X^2$  | 28.6119 | 41.6813 | 22.8846 |
| $G^2$  | 11.7274 | 39.9362 | 9.6845  |
| d.f.   | 17      | 17      | 16      |

### Table 11: Parameter estimates under the multiplicative model with constant dispersion parameter.

| # | TRT | n | y | $\hat{\psi}$ | $\hat{\pi}_1$ | $S^2$ | $\sum G^2$ | $\sum X^2$ | $\sum Wald$ |
|---|-----|---|---|-------------|-------------|------|------------|------------|-----------|
| 1 | 0   | 5 | 0 | 0.1565 | 0.1086 | 0.5467 | 0 | 0.543 | 0.5392 |
| 2 | 0   | 6 | 2 | 0.1565 | 0.0974 | 0.6076 | 4.9215 | 3.9724 | 3.8374 |
| 3 | 0   | 7 | 0 | 0.1565 | 0.0868 | 0.6485 | 4.9215 | 4.5799 | 4.4064 |
| 4 | 0   | 7 | 0 | 0.1565 | 0.0868 | 0.6485 | 4.9215 | 5.1873 | 4.9754 |
| 5 | 0   | 8 | 0 | 0.1565 | 0.0769 | 0.6702 | 4.9215 | 5.8023 | 5.5396 |
| 6 | 0   | 8 | 0 | 0.1565 | 0.0769 | 0.6702 | 4.9215 | 6.4172 | 6.1039 |
| 7 | 0   | 8 | 0 | 0.1565 | 0.0769 | 0.6702 | 4.9215 | 7.0321 | 6.6681 |
| 8 | 0   | 9 | 1 | 0.1565 | 0.0677 | 0.6743 | 5.9118 | 7.2823 | 6.8942 |
| 9 | 0   | 9 | 2 | 0.1565 | 0.0677 | 0.6743 | 10.6648 | 10.4546 | 9.7616 |
| 10| 0   | 10| 1 | 0.1565 | 0.0594 | 0.6637 | 11.7067 | 10.7322 | 10.0101 |
| 11| 1   | 5 | 0 | 0.3409 | 0.2949 | 1.3354 | 11.7067 | 12.2067 | 11.6382 |
| 12| 1   | 5 | 2 | 0.3409 | 0.2949 | 1.3354 | 12.296 | 12.394 | 11.945 |
| 13| 1   | 7 | 1 | 0.3409 | 0.2643 | 1.9864 | 11.6956 | 12.7845 | 12.2088 |
| 14| 1   | 8 | 0 | 0.3409 | 0.2469 | 2.3069 | 11.6956 | 14.7597 | 13.8999 |
| 15| 1   | 8 | 2 | 0.3409 | 0.2469 | 2.3069 | 11.7456 | 14.76 | 13.9002 |
| 16| 1   | 8 | 3 | 0.3409 | 0.2469 | 2.3069 | 14.2534 | 15.2918 | 14.3554 |
| 17| 1   | 9 | 0 | 0.3409 | 0.2282 | 2.6004 | 14.2534 | 17.3455 | 15.9774 |
| 18| 1   | 9 | 4 | 0.3409 | 0.2282 | 2.6004 | 19.5864 | 19.1899 | 17.4341 |
| 19| 1   | 10| 1 | 0.3409 | 0.2084 | 2.8434 | 18.1179 | 19.7537 | 18.7843 |
| 20| 1   | 10| 6 | 0.3409 | 0.2084 | 2.8434 | 30.8076 | 27.1123 | 23.2406 |

Adv Comput Sci Volume: 1.1

References

1. Lindsey JK, Altham PK. Analysis of the human sex ratio by using over dispersion models. Appl Statist. 1998; 47: 149-157.
2. PM Altham. Two generalizations of the binomial distribution. Appl Statist. 1978; 27: 162-167.
3. Efron, B. Exponential families and their use in generalized linear regression. J Amer Statist Assoc. 1986; 81: 709-721.
4. Lindsey JK. Modelling Frequency and Count Data. Oxford University Press; 1995.
5. Lindsey JK, Mersch G. Fitting and comparing probability distributions with log linear models. Comput Stat Data Anal. 1992; 13: 373-384.
6. Lovinson G. An alternative representation of Altham’s multiplicative-binomial distribution. Stat Proba lett. 1998; 36:
7. DR Cox. The analysis of multivariate binary data. Appl Stat 1972; 21: 113-120.
8. EAH Elamir. Multiplicative-binomial distribution: Some results and characterization, inference and random data generation. J Stat Theory Appl. 2013; 12: 92-105.
9. Feier V, Friedl H, Hirn U. Modeling over-and under dispersed frequencies of successful ink transmissions onto paper. J Appl Stat. 2013; 40: 626-643.
10. Sokal RR, Rohlf FJ. Biometry: the Principles and Practice of Statistics in Biological Research. San Francisco: Freeman; 1969.
11. Geissler A. Contribution to the question of the sexual relationship of the born. Sachs Super numerary Bur. 1889; 35: 1-24.
12. Borges P, Rodrigues J, Balakrishnan N. A Com-Poisson type generalization of the Binominal distribution and its properties and applications. J Stat Prob Letters. 2014; 87: 158-166.
13. Nelder JA, Mead R. A simplex method for function minimization. Comput J. 1965; 7: 308-312.
14. Morel JG, Neerchal NK. Clustered binary logistic regression in teratology data using a finite mixture distribution. Stat Med. 1997; 16: 2843-2853.
15. Morel JG, Neerchal NK. Overdispersion Models in SAS. SAS publication 2012.
16. Manly B. Regression models for proportions with extraneous variance. Biometrie-Praximetrie. 1978; 18: 1-18.
17. Lawal HB. Categorical Data Analysis with SAS and SPSS Applications. Taylor & Francis. 2003; New York, USA.