Nigerian Journal of Mathematics and Applications
Volume **25**, (2016), 88 − 110

http://www.njmaman.com

Estimating Infant and Maternal Mortality in Nigeria using Overdispersion Modelling Methods

C. I. Orabuche, M. O. Adamu* and O. E. Oladokun

Abstract

In this study, models that provide more satisfactory statistical fit for data and factors that affect infant and maternal mortality rates in Nigeria are presented. Models that detect and correct overdispersion in data are used to analyze the data provided by the National Bureau of Statistics, Nigeria. Some of the models considered are Poisson, Poisson Normal, Negative Binomial, Quasi-likelihood and Bayesian models. The results presented are promising.

1. Introduction

A common task in applied statistics is choosing a parametric model to fit a given set of empirical observations. This necessitates an assessment of the fit of the chosen model. It is usually possible to choose the model parameters in such a way that the theoretical population mean of the model is approximately equal to the sample mean. However, especially for simple models with few parameters, theoretical predictions may not match empirical observations for higher moments. When the observed variance is higher than the variance of a theoretical model, overdispersion has occurred. Conversely, underdispersion means that there was less variation in the data than predicted. Overdispersion is a very

Received: 05/09/2016, Accepted: 12/12/2016, Revised: 20/12/2016.
2015 *Mathematics Subject Classification.* 49Nxx & 00Axx. * Corresponding author.
Key words and phrases. Nonlinear Integro-differential Equations, Variation iteration method, Lagrange multiplier, Decomposition method
Department of Mathematics, University of Lagos, Lagos.;
Email: collinsorabuche@googlemail.com

common feature in applied data analysis because in practice, populations are frequently heterogeneous contrary to the assumptions implicit within widely used simple parametric models.

Poisson regression is one of the most popular techniques for the analysis of count data. The Poisson regression model may be the foremost method, but it rarely explains the data due to several important constraints. One important constraint is the mean of the distribution must be equal to the variance. In this case, the standard errors, usually estimated by the Maximum Likelihood method, will be biased and the test statistics derived from the models will be incorrect. Therefore, this problem leads to an overdispersion. Failure to take overdispersion into account leads to serious underestimation of standard errors and misleading inference for the regression parameters. Consequently, a number of models and associated estimation methods have been proposed for handling overdispersed data. Such models include those based on the negative binomial distributions as well as regression models based on mixtures of Poisson and the generalized Poisson distributions with regression models.

In this work, we propose extensions of the different methods useful for modelling overdispersed data used by Adrian, et al. (2012). Based on data that consider the study of the departmental factors affecting mortality of children under 5 years and maternal mortality, a series of different overdispersion modelling approaches are considered. The results for these modelling approaches considering overdispersion can be compared with those from standard overdispersion models that do not include this extra variation. We will briefly introduce some of the most commonly used and well-known methods to model overdispersion for the cases in which the response variable presents overdispersion associated with the binomial or Poisson distributions. Among these methods, we mention the General Linear Models (GLMs) with random mean, the quasi-likelihood functions and the double exponential families. Motivated by the data on the analysis of the factors affecting infant and maternal mortality in Nigeria, we propose some Bayesian model extensions that generalize the commonly used overdispersion approaches in that they provide a better fit to the data set under study, allow the researchers to take into account additional factors that cannot be included or modelled by earlier models and illustrate their usefulness in the light of the application. We present some of the tools commonly used to detect extra variation, some of the methods used to estimate the parameters in the aforementioned models, and we also present the motivating data set for the analysis of the factors affecting infant and maternal mortality in Nigeria.

This article show how to detect, correct overdispersion in count data and criteria for choosing the best fit model for such data using data collected from the National Bureau of Statistics (NBS) Publication (2009); to estimate infant and maternal mortality in Nigeria during the period of 2001- 2008.

1.1. **Operational Definitions.**

(1) **Maternal Mortality:** Maternal mortality is the death of a woman who is pregnant or who has been pregnant within the preceding six weeks, irrespective of the duration or site of the pregnancy from any cause related to or aggravated by pregnancy or its management, but not from accidental or incidental causes.
(2) **Infant:** Children under the age of five. The term infant is typically applied to young children between the ages of 1 month and 12 months; however, definitions may vary between birth and 5 years of age.

2. Overdispersion Models

Overdispersion may also rise as a result of unexplained heterogeneity. To account for this heterogeneity a random effects model can be fitted to the data. Generalized Linear Mixed Models (GLMM) was proposed as a general framework by Breslow (1990). They include an unobserved vector of random effects in a GLM, assumed to arise from a normal distribution, and use an approximation of the marginal quasi-likelihood based on Laplace's method, leading to equations based on penalized quasi-likelihood. Lin and Zhang (1999) extended the idea by using smoothing cubic splines to propose generalized additive mixed models. To avoid the complex numerical integration required to estimate the model, they proposed a double penalized marginal quasi-likelihood also based on a Laplace approximation.

Schall (1991) proposed a general algorithm for the estimation of random effects and dispersion parameters applicable in GLMs, regardless of the structure of the linear predictor, and without the need to specify the distribution of the random effect. In his application section, he used random effects to explain extra-binomial variation; however, he did not examine this case in detail.

2.1. **Overdispersion: Causes of Overdispersion.**

2.1.1. *Causes of Overdispersion.* There are many different possible causes of overdispersion and in any modelling situation a number of these could be involved. Some possibilities are:

(1) Variability of experimental material: this can be thought of as individual variability of the experimental units and may give an additional component of variability which is not accounted for by the basic model;
(2) Correlation between individual responses: for example in cancer studies involving litters of rats we may expect to see some correlation between rats in the same litter;
(3) Cluster sampling;
(4) Aggregate level data: the aggregation process can lead to compound distributions;

(5) Omitted unobserved variables: in some sense the other categories are all special cases of this, but generally in a rather complex way. Lee and Nelder (2002)

In some circumstances, the cause of the overdispersion may be apparent from the nature of the data collection process, although it should be noted that different explanations of the overdispersion process can lead to the same model; so in general it is difficult to infer the precise cause, or underlying process, leading to the overdispersion.on: Causes and Consequences

2.1.2. *Consequences of Overdispersion.* When we identify the possible presence of overdispersion, what are the consequences of failing to take it into account? Firstly, the standard errors obtained from the model will be incorrect and may be seriously underestimated and consequently we may incorrectly assess the significance of individual regression parameters. Also, changes in deviance associated with model terms will also be too large and this will lead to the selection of overly complex models. Finally, our interpretation of the model will be incorrect and any predictions will be too precise.

2.2. **Methods of Detecting Overdispersion.** Lindsey (1999) presented an alternative intuitive way of detecting overdispersion that compares the Akaike information criterion (AIC) for the standard model with that of the model that incorporates overdispersion in the data and then detects if there is a significant improvement in the goodness of fit of the model. If so, there is some evidence that there is extra variation in the data and, thus, the researcher should consider fitting the model that incorporates overdispersion. Another statistic that has been commonly used to detect for overdispersion in the data is the *deviance*, which has also been previously used to test for the goodness of fit of a model. It is defined as twice the difference between the log-likelihood for the saturated model and that of the estimated model. *If the ratio of the deviance for the model being tested over its degrees of freedom is greater than one, there could be some indication for the presence of overdispersion in the data.* However, some authors recommend checking for overdispersion only if this ratio is greater than two (J.K. Lindsey, 1999). The deviance for data following a binomial distribution is given by $D\left(y,\ \widehat{\mu}\right) = 2\sum_{i=1}^{n}\left\{y_i ln\left(\frac{y_i}{m_i\widehat{p}_i}\right)\right\} +\ (m_i -\ y_i)\,ln\left[\frac{(m_i -\ y_i)}{(m_i -\ m_i\widehat{p}_i)}\right]$, (**??**)
and for data following a Poisson distribution, it is given by

$$D\left(y,\ \widehat{\mu}\right) = 2\sum_{i=1}^{n}\left[y_i ln\left(\frac{y_i}{\widehat{\lambda}_i}\right) -\ \left(y_i -\ \widehat{\lambda}_i\right)\right], \quad (1)$$

An additional graphical tool to determine if the fitted model is an adequate one or if there is overdispersion in the data is the envelope plot. This tool is a part of the normal percentile plot (i.e. Q–Q plot), for which deviance residuals, obtained from the fitted model, are plotted against the theoretical residuals obtained from the normal distribution. If the plot differs from that of a straight line, there would be clear indications that the deviance residuals do not follow a normal distribution, implying that the fitted model is not adequate for the data. In addition, this plot can also be used to detect the existence of possible outliers in the data; for more details on this, see A.C. Atkinson (1985). The envelope plot simulates empirical confidence bands to determine if the residuals significantly differ from the straight line. It is based on the simulation of several samples for the response variable, which are, in turn, generated from the estimations obtained in the fitted model and from the assumed distribution for the response variable being analysed. Hinde and Demétrio (1998) illustrated the procedure to generate a similar plot (i.e. the half-normal plot) that can be used as well to assess the appropriate model's fit and the existence of outliers. Atkinson (1985) suggested that half-normal plots are more effective than full-normal or envelope plots in detecting outliers and influential observations for moderate-sized samples but that, for larger sample sizes, full-normal or envelope plots can be expected to be more informative.

2.3. **Some Commonly Used Overdispersion Models.** The most commonly used models that can adjust for extra variation are obtained after generating a family of "overdispersed" models, in which a random component is initially added to the linear predictor so that a larger variance for the response variable is allowed for. Such models are known as random mean models. Now, we describe some well-known models of this type used to adjust for overdispersion in the Poisson and binomial cases Dean (1992).

2.3.1. *Models Associated with the Poisson Distribution.* The first standard overdispersion model for Poisson data is the negative binomial model. This model, originally proposed by Margolin, et al. (1981), is highly recommended to handle overdispersion in Poisson type data, especially when the researcher believes that data come from a population having different subpopulations. The model assumes that the random variable Y_i conditioned on λ_i^* follows a Poisson distribution with parameter λ_i^* and that the λ_i^*'s follow a gamma distribution with $\mathrm{E}(\lambda_i^*) = \lambda_i$ and $\mathrm{var}(\lambda_i^*) = \frac{\lambda_i^2}{\tau}$. In this way, if τ is a fixed value, the unconditional distribution of Y_i is negative binomial Efron (1986)

$$f(y_i) = \binom{y_i + \tau - 1}{y_i} \left(\frac{\lambda_i}{\lambda_i + \tau}\right)^{y_i} \left(\frac{\tau}{\lambda_i + \tau}\right)^{\tau} \tag{2}$$

so that $\mathrm{E}(Y_i) = \lambda_i$ and $var(Y_i) = \lambda_i + \frac{\lambda_i^2}{\tau}$. These settings allow for the possibility of having different models of this type assuming that the parameters in the assumed distribution for λ_i^* vary. For example, if λ_i^* follows a gamma $\Gamma(\lambda_i, \tau\lambda_i)$ distribution, the unconditional variance of Y_i would be $\lambda_i \left(1 - \frac{1}{\tau}\right)$.The second standard overdispersion model is the Poisson normal model, a model that

assumes that the natural logarithm of the distributional mean can be modelled by introducing a random component having a normal distribution. In this model, the λ_i's are modelled so that $ln\,(\lambda_i) = X_i^T\beta + \ v_i$ (**??**)
where the ν_i's are independently identically distributed random variables having a normal distribution with mean zero and finite variance τ. In this model and if τ is small, the Y_i random variables are mixed Poisson variables with approximate mean and variance given by λ_i and $\lambda_i(1 + \tau\lambda_i)$, respectively. (C.D. Dean, 1992)

2.3.2. *Models Associated with the Binomial Distribution.* The first standard overdispersion model for binomial data is the beta-binomial model. This model starts by assuming that the random variables Y_i, $i = 1, \ . \ . \ . \ , n$, conditioned on π_i^*, are independent and have a binomial distribution and that the π_i^*'s follow a beta distribution with $\mathrm{E}(\pi_i^*) = \pi_i$ and $var\,(\pi_i^*) = \tau\pi_i(1 - \pi_i)$. In this way, the unconditional distribution of Y_i is beta binomial with mean $m_i\pi_i$ and variance $m_i\pi i(1 - \pi_i)[1 + \tau(m_i - 1)]$. In addition, its probability density function is given by Lindsey (1999).

$$f\,(y_i) = \begin{pmatrix} m_i \\ y_i \end{pmatrix} \frac{B\,(k_i + \ y_i, \ \ v_i + \ m_i - \ y_i)}{B\,(k_i, v_i)} \tag{3}$$

with $k_i = \frac{(1- \ \tau)\pi_i}{\tau}$, $v_i = \frac{(1- \ \tau)(1- \ \pi_i)}{\tau}$ and $B(\,)$ being the beta function. The idea is to fit a regression model that assumes that the response variable follows the beta-binomial distribution. This assumption allows us to have a larger variance than the one considered in the GLM with binomial response variable. Moreover, the use of this model is recommended when the researcher believes that data come from a population having different subpopulations and, in addition, when there is correlation between the Bernoulli events within each binomial observation.
The second standard overdispersion model for binomial data is the binomial normal model. Some researchers such as Williams (1975), incorporated the extra-binomial variation with a model of the form

$$ln\left(\frac{\pi_i}{1- \ \pi_i}\right) = \ X_i^T + \ v_i, \quad i = 1, 2, \ldots, n \tag{4}$$

where the ν_i's are a random sample of variables having mean zero and variance τ . Under this model and when τ is small, the conditional means and variances of the random variables Y_i's are approximately given by $m_i\pi_i$ and $m_i\pi_i(1 - \pi_i)[1 + \tau(m_i - 1)\pi_i(1 - \pi_i)]$, respectively Dean (1992). In addition, it is usually assumed that the ν_i's follow a Gaussian distribution, so that the binomial normal model is obtained. As can be seen clearly, this approach is commonly used when the researcher believes that the modelling of the logit of the binomial proportion can be assumed to include a random error.

2.3.3. *Model Extensions Associated with the Poisson Distribution.* As in section 2.3.1, we first describe the proposed generalizations for the negative binomial model. In this model, we assume that the unconditional distribution of Y_i is given by

$$f\left(y_i\right) = \begin{pmatrix} y_i + \tau - 1 \\ y_i \end{pmatrix} \left(\frac{\lambda_i}{\lambda_i + \tau}\right)^{y_i} \left(\frac{\tau_i}{\lambda_i + \tau_i}\right)^{\tau_i} \tag{5}$$

and that the dispersion parameter varies as a positive function of a set of explanatory variables **Z**, so that a joint modelling of the mean and overdispersion (or variance) is now considered. An example of this situation could be the one where we have that the corresponding link functions are given by $ln\left(\lambda_i\right) = X_i^T\ and \ln\left(\tau_i\right) = Z_i^T \gamma$.

Along the same lines, the second overdispersion model to be generalized is the Poisson normal model. In this model, the extra variation for the Poisson response is incorporated by using the random variable ν_i already included in the linear predictor, as in Equation (**??**), and assuming that $\mathrm{Var}(\nu_i) = \tau_i$ varies as a function of some set of explanatory variables. In this way, this variance function for τ_i can be an appropriate real valued positive function such as the one described above for the negative binomial model. In order to verify if the overdispersion varies as a function of a single explanatory variable, it suffices to check if the parameter associated with this variable is statistically different from zero. That is, if, for example, we consider the model $\ln(\tau_i) = \gamma_0 + \gamma_1 Z_i$ and the null hypothesis $H_0 : \gamma_1 = 0$ is rejected, then we have that the functional dependence of the overdispersion parameter on the explanatory variable Z is given by this specific function. We will illustrate the application, motivation and usefulness of these generalizations to the data set under study.

2.3.4. *Model Extensions Associated with the Binomial Distribution.* As in section 2.3.2, the first overdispersion model to be generalized is the beta-binomial model. In this model, the unconditional distribution of Y_i is assumed to be beta binomial with mean $m_i\pi_i$ and variance now given, in this extended version of the model, by $m_i\pi_i(1 - \pi_i)[1 + \tau_i(m_i - 1)]$. That is, we now assume that its probability density function is given by Equation (3), with $k_i = \frac{(1- \tau_i)\pi_i}{\tau_i}$, and $v_i = \frac{(1- \tau_i)(1- \pi_i)}{\tau_i}$. In this way, the dispersion parameter is allowed to vary as a function of a set of some of the explanatory variables. The underlying idea of this new proposal is the joint modelling of the mean and the overdispersion parameter under the assumption that the response variable follows a beta-binomial distribution, an assumption that allows for the possibility of having a larger variance than the one considered in the standard GLM or of having a non-constant overdispersion parameter, thus, also generalizing the constant parameter case described in section 2.3.2 for the beta-binomial model.

Along the same lines, the second overdispersion model to be generalized is the binomial normal model. In this model, the extra variation for the binomial response is incorporated by using the random variable ν_i already included in the linear predictor, as in Equation (4), and assuming that $\text{Var}(\nu_i) = \tau_i$ varies as a function of some set of explanatory variables. In this way, this variance function for τ_i can be an appropriate real valued positive function such as $\ln(\tau_i) = Z_i^T \gamma$, where $\mathbf{Z}_i$ is a vector of explanatory variables and $\boldsymbol{\gamma}$ is the vector of unknown parameters in this variance model.

2.4. **Generalized Linear Models.** Let Y be a random variable that belongs to the exponential family of distributions. That is, its probability density function Mathew (2004) is given by

$$f(y;\theta,\emptyset) = \exp\left\{\frac{[y(\theta) - b(\theta)]}{a(\emptyset)} + c(y,\ \emptyset)\right\} \tag{6}$$

Here, we assume that $a(\emptyset) = \emptyset$, a constant positive value known as dispersion parameter. It can be shown that, for this family, $E\left(\dot{Y}\right) = \mu = b^{\imath}(\theta)$ and $var\left(\dot{Y}\right) = \emptyset V(\mu)$, where $V(\mu) = b^{\imath\imath}(\theta)$ is a known positive variance function and θ is a parameter known as location or canonical parameter. If $Y_1, \ldots, Y_n$ are independent random variables with distribution belonging to the exponential family of distributions, the model relating the observations $y_i, i = 1, \ldots, n$, with the explanatory variables is of the form

$$g(\mu) = \eta_i\ = X_i^T \beta,\ i = 1, \ldots, n \tag{7}$$

where η_i is the linear predictor for the ith observation, $\mathbf{X}_i$ is the $q \times 1$ vector of explanatory variables for the ith observation, $\boldsymbol{\beta}$ is a $q \times 1$ vector of unknown regression parameters and $g(\mu)$ is a monotone and differentiable link function. Usually, the canonical link function $g(\mu) = \theta$ is chosen, so that both the existence of a maximum-likelihood estimation of the parameter vector $\boldsymbol{\beta}$ and an easier interpretation of the model are guaranteed.

2.5. **General Overdispersion Models.** The aforementioned extra variation for the binomial and Poisson cases can alternatively be modelled with extensions of the GLMs Nelder and Wedderburn (1972), so that the dispersion parameter is included in the models. In order to do this, there are three possible approaches:

(1) The GLMs with random mean Cox, (1983);
(2) The quasi-likelihood methods, which allow for a more flexible relation between the mean and the variance of the variable to be modelled Wedderburn (1974),
(3) The double exponential family, which allows for the fitting of a regression model to the dispersion parameter Efron (1986).

GLMs with random mean start by adding a random component to the linear predictor and this is the main reason for also labelling them as composite or mixed models. They are mainly useful to work with overdispersion in binomial and Poisson-type data. The models assume that the probability density function for the random variable Y_i depends on the parameters θ_i^* and that the θ_i^*'s are continuous and independent random variables with finite mean and variance given by $E(\theta_i^*) = \theta_i$ and $var(\theta_i^*) = \tau b_i(\theta_i) > 0$, where θ_i, defined in Equation (6), is the parameter defining the distribution of the random variable Y_i for the case where there is no extra variation in the model. This family encompasses a great number of different models, such as the ones described in earlier.

Quasi-likelihood methods, originally introduced by Wedderburn (1974), only require that a relation $V(\mu)$ between the mean and the variance of the response variable be specified. In this way, it is assumed that there is a relation, such as the one in Equation (7), between the mean and the explanatory variables and, in addition, that $var(Y) = \sigma^2 V(\mu)$, where σ^2 is the dispersion parameter. Quasi-likelihood models can be very useful when modelling overdispersion because they allow for a more flexible variance function $var(\mu)$ than the one assumed in standard GLMs. As is usually done with the likelihood function, we define the quasi-likelihood function for the random variable Y for these models as follows:

$$Q\left(\mu, y\right) = \frac{1}{\sigma^2} \int_y^{\mu} \frac{(y-t)}{V\left(t\right)} dt,$$

so that for a set of independent random variables $Y_1, \ldots, Y_n$, it will be given by $Q(\boldsymbol{\mu}; \mathbf{Y}) = Q(\mu_i;\, y_i)$. Even though, in principle, the function $V(\mu)$ can be any positive function, a simple way of trying to model the overdispersion in binomial and Poisson-type data is to assume that $Var(Y) = \sigma^2 V(\mu)$, where $V(\mu)$ is the variance function for the standard GLM and $\sigma^2 > 1$, so that it allows for the possibility of increasing the variance of the GLM by increasing the σ^2 value. Defining the quasi-likelihood model in this way results in the fact that its estimations are the same as those obtained in the standard GLM, except for $\widehat{\sigma}^2$. Therefore, the estimations obtained from this overdispersion model for the binomial and Poisson cases can be obtained from the estimations in the standard GLM, and the researcher only needs to correct the estimations for the asymptotic variance–covariance matrix for $\widehat{\beta}$, the deviance, the residuals, etc. by making use of the estimation obtained for σ^2 Paula (2004).

The double exponential family of distributions allows the researcher to obtain the double binomial and double Poisson models, depending on the nature of the variable under study. The double exponential family was originally proposed by Efron (1986) and later discussed in Nelder and Pregibon (1987). It allows us to include a second parameter that, independently from the mean, controls for the variance of the response variable. That is, the dispersion parameter can be modelled from a subset of some explanatory variables, at the same time that the usual regression model for the mean is fitted. If we have $Y_1, \ . \ . \ . \ , Y_n$ independent response variables, Gordon (1989) wrote these models in the form given by

$$E(Y_i) = \mu_i = g^{-1}(\eta_i); \qquad \eta_i = X_i^T \beta \tag{8}$$

with the variance of the response variable Y_i given by $var(Y_i) = \phi_i V(\mu_i);\ \phi_i = g_\phi^{-1}(\gamma_i);\ \gamma_i = \lambda + Z_i^T \alpha$ (**??**), where the functions g and V and the vectors $\mathbf{X}_i$ and $\boldsymbol{\beta}$ are defined as earlier. The function g_φ is assumed to be positive and doubly differentiable, λ is a scalar parameter, $\mathbf{Z}_i$ is a $q_\varphi \times 1$ vector of some explanatory variables ($q_\varphi \leq$) and $\boldsymbol{\alpha}$ is a $q_\varphi \times 1$ vector of unknown parameters. We should note that if $\boldsymbol{\alpha}$ is a vector of zeroes, φ_i would be a constant dispersion parameter and, thus, Equations (8) and (**??**) would specify a GLM.

2.6. **Infant And Maternal Mortality Rate.** Examination of the trend in infant mortality rate (U5MR) in Nigeria shows that the projection of 55 per 1000 for 2015 may be unattainable because of the unabating increase recorded in the past five years. The secondary data show that about 78 percent of the children die before their first birthday while 37 percent die before the end of their first month. The most common killers of the under five children are found to be Bronco pneumonia, Sepsis, Anaemia and Malaria. Infant mortality are found to be related to mothers education and income. However, age of mothers and their occupation are not significant contributors to U5MR. Traditional beliefs are found to be important in the treatment of diseases of children as mothers who see their sick children as spiritually afflicted will rather seek for spiritual assistance than take them to hospital for treatment.

Existing research works in prints and electronic among others were used in the search and collation of facts and figures contained in the study. Nigeria is the most populous Black Nation and has the maternal mortality rate of 280 to 1150 per 100,000 live births Onwumere (2010).

Maternal and child mortality is closely linked to poverty with malnutrition as an underlying contributor in over half of these deaths. Factors associated with these problems include, poor socio-economic development, weak health care system and low socio-cultural barriers to care utilization Ibeh (2008).

A research on population-based study of effect of multiple births on infant mortality in Nigeria showed the multiple births are strongly negatively associated with infant survival in Nigeria independent of other risk factors. Mother's education played a protective role against infant death. This evidence suggests that improving maternal education may be key to improving child survival in Nigeria. A well-educated mother has a better chance of satisfying important factors that can improve infant survival: the quality of infant feeding, general care, household sanitation, and adequate use of preventive and curative health services (Olalekun

2008). A study by Mondal (1997) also observed that socio-economic status and literacy levels influenced utilization of antenatal service. The importance of quality antenatal care cannot be questioned (McDonagh, 1996). How tragic it is that pregnancy and childbirth, a joyous event for women in developed countries becomes a source of lifelong disability or worse for women in the developing world. For a country like Pakistan where technology has advanced to produce nuclear bomb, it is ironical to have the maternal mortality rate which Sweden had in the 1880s, political will backed by adequate resources is urgently needed to implement programs to reduce maternal mortality and morbidity through provision of good antenatal care (Inam and Khan, 2002).

3. Poisson Distribution Model

The first standard overdispersion model for Poisson data is the negative binomial model. This model, originally proposed by Margolin, et al (1981), is highly recommended to handle overdispersion in Poisson type data, especially when the researcher believes that data come from a population having different subpopulations. The model assumes that the random variable Y_i conditioned on λ_i^* follows a Poisson distribution with parameter λ_i^* and that the λ_i^*'s follow a gamma distribution with $E(\lambda_i^*) = \lambda_i$ and $var(\lambda_i^*) = \frac{\lambda_i^2}{\tau}$. In this way, if τ is a fixed value, the unconditional distribution of Y_i is negative binomial Efron, (1986)

$$f(y_i) = \binom{y_i + \tau - 1}{y_i} \left(\frac{\lambda_i}{\lambda_i + \tau}\right)^{y_i} \left(\frac{\tau}{\lambda_i + \tau}\right)^{\tau} \tag{9}$$

so that $E(Y_i) = \lambda_i$ and $var(Y_i) = \lambda_i + \frac{\lambda_i^2}{\tau}$. These settings allow for the possibility of having different models of this type assuming that the parameters in the assumed distribution for λ_i^* vary. For example, if λ_i^* follows a gamma $\Gamma(\lambda_i, \tau\lambda_i)$ distribution, the unconditional variance of Y_i would be $\lambda_i\left(1 - \frac{1}{\tau}\right)$. The second standard overdispersion model is the Poisson normal model, a model that assumes that the natural logarithm of the distributional mean can be modelled by introducing a random component having a normal distribution. In this model, the λ_i's are modelled so that $ln(\lambda_i) = X_i^T\beta + v_i$ (**??**)
where the ν_i's are independently identically distributed random variables having a normal distribution with mean zero and finite variance τ. In this model and if τ is small, the Y_i random variables are mixed Poisson variables with approximate mean and variance given by λ_i and $\lambda_i(1 + \tau\lambda_i)$, respectively. Dean, (1992).

3.1. Model Extensions Associated with the Poisson Distribution.

As in section 2.3.3, we first describe the proposed generalizations for the negative binomial model. In this model, we assume that the unconditional distribution of

Y_i is given by

$$f(y_i) = \begin{pmatrix} y_i + \tau - 1 \\ y_i \end{pmatrix} \left(\frac{\lambda_i}{\lambda_i + \tau}\right)^{y_i} \left(\frac{\tau_i}{\lambda_i + \tau_i}\right)^{\tau_i} \tag{10}$$

and that the dispersion parameter varies as a positive function of a set of explanatory variables **Z**, so that a joint modelling of the mean and overdispersion (or variance) is now considered proposed suggested by Adrian et al. (2012).
Along the same lines, the second overdispersion model to be generalized is the Poisson normal model. In this model, the extra variation for the Poisson response is incorporated by using the random variable ν_i already included in the linear predictor, as in Equation (**??**), and assuming that $\mathrm{Var}(\nu_i) = \tau_i$ varies as a function of some set of explanatory variables. In this way, this variance function for τ_i can be an appropriate real valued positive function such as the one described above for the negative binomial model. In order to verify if the overdispersion varies as a function of a single explanatory variable, it suffices to check if the parameter associated with this variable is statistically different from zero. We illustrate the application, motivation and usefulness of these generalizations to the data set under study.

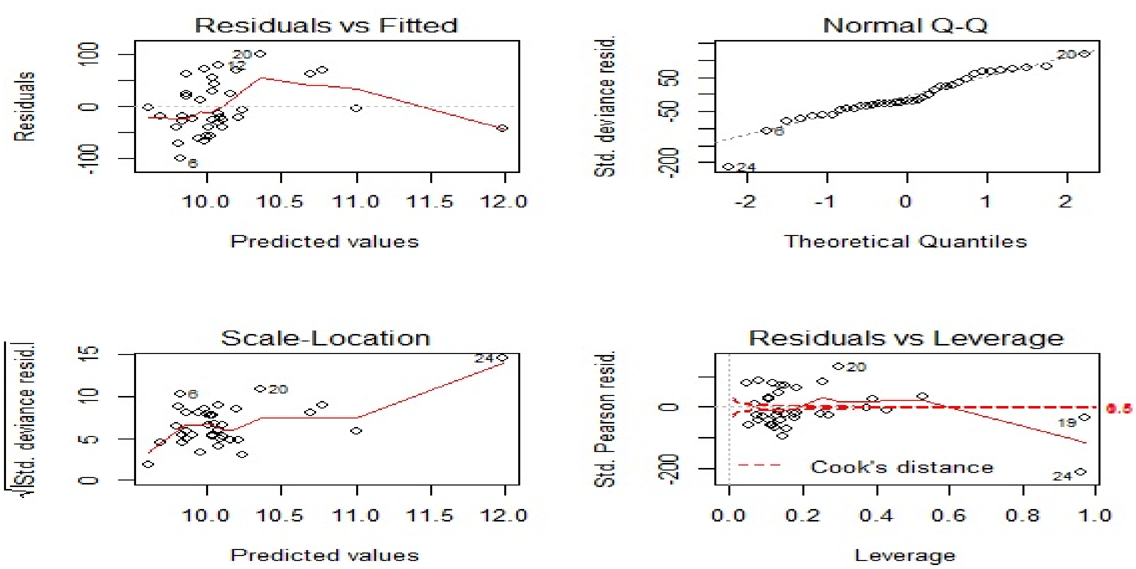
4. Data Analysis: Introduction

This section test for overdispersion in the data collected from the NBS publication (2010) used to estimate the infant and maternal mortality rate in Nigeria. It does this by fitting a Poisson model on the data. The variables used include; average number of children under 5 years who died in the 8-year period 2000–2007 (AND), number of registered live births in each of these 36 states and F.C.T. (NLB), percentage of pregnant women who registered with clinics, which show the level of awareness of the mothers in that state (PWC), percentage of pregnant women who received anti tetanus injection (PWT), the average number of infants immunized in each states (including F.C.T.) with the three doze of antigen (i.e. variables OPV3, DPT3 & MV), percentage of infants not immunized in each of the 36 states and F.C.T. (NOT), average number of maternal death between 2001-2008 (NDA), the number of pregnant mother diagnosed of malaria in 2007 (MRW), number of women that received antenatal care from a health professional in 2008 (ANTE), number of women whose last live birth was protected against neonatal tetanus in 2008 (TET), number of women delivered by a health professional which includes; Doctors, Nurse, Midwife or Auxiliary Midwife in 2008 (DHP) and number of women delivered in a health facility (DHF). The data used covers a period of eight years (2001-2008) from the 36 states and the Federal Capital Territory (Abuja) in Nigeria.

4.1. **Presentation And Interpretation of Results.** In order to determine the factors that significantly affect infant mortality, we assume that the average number of children under 5 years who died in the 8-year period 2000–2007 (i.e. variable AND) in the i^{th} state follows a Poisson distribution with mean λ_i, so that for the data set under study, it may be reasonable to propose the initial model given by

$$ln(\lambda_i) = \beta_0 + \beta_1 NLB_i + \beta_2 OPV3_i + \beta_3 DPT_i + \beta_4 MV_i + \beta_5 NOT_i + \beta_6 PWT_i + \beta_7 PWC_i$$

.



Figuare 1: A Plot of Poisson Model for Estimating Infant Mortaliy in Nigeria.
Table 1: Estimates for the fitted Poisson model for the average number of children under 5 years who died between 2000 and 2007.

Coefficients	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	9.611e+00	9.704e-03	990.35	<2e-16 ***
NLB	1.390e-06	5.972e-09	232.72	<2e-16 ***
OPV3	4.314e-07	4.213e-08	10.24	<2e-16 ***
DPT3	-7.286e-06	1.004e-07	-72.58	<2e-16 ***
MV	5.961e-06	6.286e-08	94.83	<2e-16 ***
NOT	-5.438e-03	1.591e-04	-34.19	<2e-16 ***
PWC	2.048e-03	6.181e-05	33.13	<2e-16 ***
PWT	2.725e-03	7.989e-05	34.11	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for poisson family taken to be 1)
Null deviance: 493643 on 36 degrees of freedom

Residual deviance: 87611 on 29 degrees of freedom
AIC: 88066
Number of Fisher Scoring iterations: 4

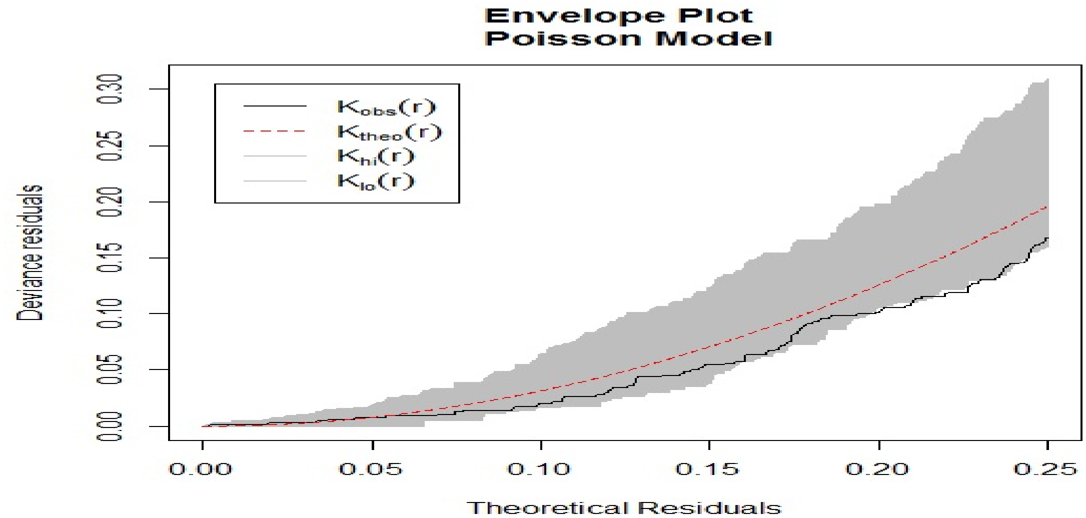


Figure 2: Envelope Plot of theoretical 95% confidence bands for Poisson Model

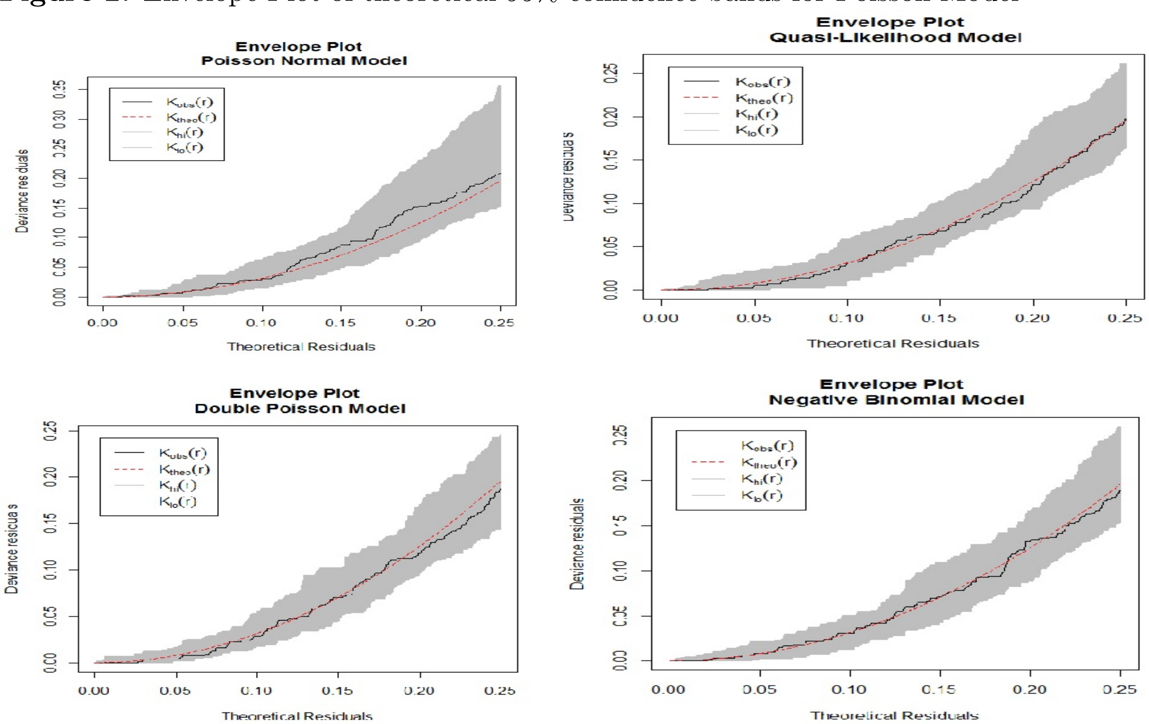


Figure 3: Envelope Plot of theoretical 95% confidence bands for Quasi-Likelihood, Poisson Normal, Double Poisson and Negative Binomial Models

Clearly, from Figure 2 we can see the existence of overdispersion in the data, thus the Poisson model is inadequate for estimating infant mortality using this data set. The fitting of the Poisson model using R statistical package, resulted in a deviance value of 87611, with 29 degrees of freedom, which seems to indicate the possible existence of overdispersion in the data. That is on using the ratio of the deviance for the model above, over the degrees of freedom; which is (3021.07) greater than 2 as recommend by Lindsey (1999); this also confirms the result from the envelope plot in Figure 2. One possible reason for this may be the belief that mortality rates for the different areas are not independent from each other, mainly due to factors such as their geographical location and the common characteristics that they may have. In order to solve this problem and as motivated in section 2.2 and 2.3, we suggest the alternative fitting of the negative binomial, quasi-likelihood, and double Poisson overdispersion models.

In Figure 3, we can clearly see from the envelope plots that the Quasi-likelihood, Negative Binomial and Double Poisson models has successfully taken the overdispersion in the data set into consideration and hence there models would give a better estimate. Also, we can see from Figure 3 that the Poisson Normal model has shown evidence of overdispersion.

Table 2 provides the value for the AIC for each of these fitted models, except for the quasi-likelihood model, for which it is not possible to find these values. The reason for this is that, technically speaking, the quasi-likelihood models are not generally fitted by maximizing a true likelihood (i.e. by knowing and using the data's assumed distribution function, which is not possible in this case) as suggested by Adrian (2012). They are fitted by solving a quasi-score equation. Therefore, the quantity that is used in place of the likelihood is not one, and so the AIC measures are, technically, not defined (and not reliable).

Table 2: Estimates for the fitted models for the average number of children under 5 years who died between 2000 and 2007.

		Intercept	**NLB**	**OPV3**	**DPT3**	**MV**	**NOT**	**PWC**	**PWT**
POISSON AIC= 87611	ESTIMATE SE *p*-value	9.611e+00 9.704e-03 <2e-16	1.390e-06 5.972e-09 <2e-16	4.314e-07 4.213e-08 <2e-16	-7.286e-06 1.004e-07 <2e-16	5.961e-06 6.286e-08 <2e-16	-5.438e-03 1.591e-04 <2e-16	2.048e-03 6.181e-05 <2e-16	2.725e-03 7.989e-05 <2e-16
Neg Binomial AIC=780.85 BIC=762.85 $\widehat{\tau}$= 10.4102	ESTIMATE SE Pr(>|z|)	9.367e+00 4.327e-01 <2e-16	1.527e-06 3.262e-07 2.84e-06	3.350e-07 2.289e-06 0.884	-1.905e-06 5.150e-06 0.711	3.403e-06 2.961e-06 0.250	-3.200e-03 7.161e-03 0.655	2.499e-03 3.491e-03 0.474	2.754e-03 2.737e-03 0.314
Quasi-like AIC=N/A BIC=N/A $\widehat{\sigma}$= 0.1191519	ESTIMATE SE Pr(>|t|)	9.367e+00 4.817e-01 <2e-16	1.527e-06 3.633e-07 0.000229	3.349e-07 2.549e-06 0.896386	-1.904e-06 5.734e-06 0.742300	3.403e-06 3.297e-06 0.310590	-3.199e-03 7.974e-03 0.691222	2.499e-03 3.887e-03 0.525320	2.755e-03 3.047e-03 0.373436
PoissonNormal AIC=13348.3 BIC=13328.3	ESTIMATE SE *t*-value	9.261404e+00 9.966011e-03 929.298999	1.606960e-06 5.946270e-09 270.246644	-3.6641e-07 4.1920e-08 -8.740648	-1.5746e-05 1.03113e-07 -152.706366	9.9236e-06 6.3577e-08 156.086593	-3.016043e-03 1.639866e-04 -18.392009	6.468532e-03 8.134732e-05 -18.392009	6.477955e-03 6.778675e-05 95.563729
DoublePoisson AIC=781.4687 $\widehat{\alpha}_1 = 2$	ESTIMATE SE Pr(>|z|)	9.681490e+00 4.921543e-01 3.772703e-86	1.486588e-06 2.828843e-07 1.479358e-07	7.300830e-07 2.158512e-06 7.351866e-01	-9.2536e-06 4.821256e-06 5.494166e-02	6.32746e-06 3.18089e-06 4.66782e-02	-5.86257e-03 7.821706e-03 4.535402e-01	2.884057e-03 4.046723e-03 4.760379e-01	1.840055e-03 3.090469e-03 5.515783e-01

Table 3: Estimates for the fitted Double Poisson and Negative Binomial model without the existence of extra variation for the average number of children under 5 years who died between 2000 and 2007.

		Intercept	**NLB**	**OPV3**	**DPT3**	**MV**	**NOT**	**PWC**	**Vi**
Neg Binomial AIC=779.39 BIC=763.388 $\widehat{\tau}$= 10.2647	ESTIMATE SE Pr(>|z|)	9.541e+00 3.447e-01 <2e-16	1.539e-06 3.285e-07 2.79e-06	2.467e-07 2.304e-06 0.915	-1.442e-06 5.139e-06 0.779	3.123e-06 2.955e-06 0.290	-3.901e-03 7.028e-03 0.579	3.421e-03 2.628e-03 0.193	
DoublePoisson AIC=779.9762 $\widehat{\alpha}_1 = 2$	ESTIMATE SE Pr(>|z|)	9.912769e+00 3.696190e-01 1.946156e-158	1.473366e-06 2.850306e-07 2.351595e-07	6.5484e-07 2.1675e-06 7.6256e-01	-8.6501e-06 4.78970e-06 7.09209e-02	6.08575e-06 3.18357e-06 5.59267e-02	-7.762662e-03 7.480270e-03 2.993857e-01	2.495830e-03 2.954452e-03 3.982399e-01	
Estimates for the final fitted Negative Binomial model for the average number of children under 5 years who died between 20									
Neg Binomial AIC=753.21 BIC=735.207 $\widehat{\tau}$= 21.6262	ESTIMATE SE *t*-value	1.064e+02 1.421e+01 6.89e-14	2.411e-05 3.327e-06 4.25e-13	-2.193e-06 1.632e-06 0.1789	8.396e-06 3.846e-06 0.0291	-1.682e-06 2.166e-06 0.4375	-1.159e-04 4.877e-03 0.9810	2.941e-03 1.818e-03 0.1058	- 1.306e+00 1.914e-01 8.89e-12

Tables 1-2 also provide the parameter estimates, their standard errors (SEs) and the *p*-values associated with their tests of significance in the model. Based on the AIC values, the fit of all the proposed overdispersion models is quite similar, and all of them show a much better fit than the Poisson and Poisson Normal models that does not seem to be able to take into account the extra variation (overdispersion) in the data.

We can now explore the possible usefulness of the corresponding models proposed in section 3.1 for this type of response variable. That is, we consider a Negative

Binomial model similar to the one described in section 3.1. In order to determine how some of the variables under study affect the infant mortality rate in each of the areas under study, and based on the models previously fitted for this data set in this chapter, it seems reasonable to assume that the average number of children under-5 years who died in the i^{th} state and in the 8-year period 2000–2007 follows the model given by

$$\begin{aligned} ln(\lambda_i) &= \beta_0 + \beta_1 NLB_i + \beta_2 OPV3_i + \beta_3 DPT3_i + \beta_4 MV_i \\ &\quad + \beta_5 NOT_i + \beta_6 PWC_i + \beta_7 v_i \end{aligned}$$

where $\nu_i \sim N(0, \tau_i)$ and $ln(\tau_i) = \gamma_0 + \gamma_1 NLB_i$. Estimates, as well as their corresponding Standard Errors (SEs), are obtained after applying the aforementioned Bayesian methodology proposed in section 3.1, and hence, the fitted model is

$$\begin{aligned} ln(\lambda_i) &= (1.064e+02) + (2.411e-05)NLB_i + (-2.193e-06)OPV3_i + (8.396e-06)DPT3_i \\ &\quad + (-1.682e-06)MV_i + (1.159e-04)NOT_i + (2.941e-03)PWC_i + (-1.306e+00)v_i \end{aligned}$$

In addition, we have for this model that the AIC values are given by 753.21 and a much higher standard error (**??**) than that of the negative binomial model in table 3. If we compare these results with those for the model reported in Table 3, we can see that the estimates for the regression parameters and their SEs are quite similar, but that, based on the reasonable reduction in the AIC value, this would represent a better fitting model proposal for this specific data set, which motivates and justifies the new methodological proposal.

Table 4: Estimates for the fitted Poisson model for the average number of maternal death between 2001 and 2008.

Coefficients	**Estimate**	**Std. Error**	**z-Value**	**Pr(>|z|)**
Intercept	9.184e+00	3.129e-03	2934.72	<2e-16 ***
MRW	1.300e-05	1.539e-07	84.51	<2e-16 ***
ANTE	4.557e-04	3.376e-05	13.50	<2e-16 ***
TET	-2.852e-03	6.029e-05	-47.31	<2e-16 ***
DHP	5.211e-03	8.264e-05	63.06	<2e-16 ***
DHF	-3.983e-03	7.764e-05	-51.30	<2e-16 ***

Null deviance: 119281 on 36 degrees of freedom.
Residual deviance: 109589 on 31 degrees of freedom.
The Dispersion parameter for poisson family taken is 1.
$AIC = 110008$ and $BIC = 110018.1$

$$\begin{aligned} ln(\lambda_i) &= 9.184e+00 + 1.300e-05MRW_i + 4.557e-04TYANTE_i - 2.852e-03TET_i \\ &\quad + 5.211e-03DHP_i - 3.983e-03DHF_i \end{aligned}$$

where $i = 1, \ldots, 37$ (States in Nigeria), where λ is the poisson variable.

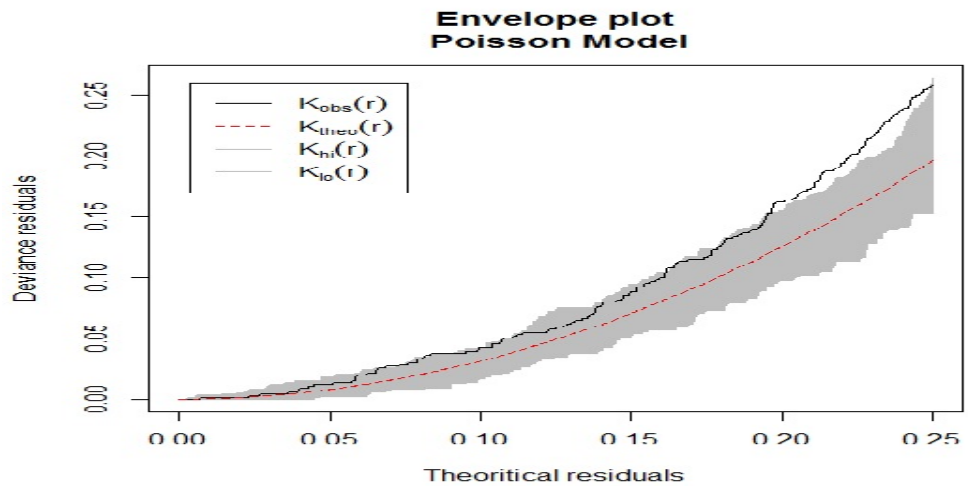


Figure 4: Envelope plots for the data set with clear signs of overdispersion with theoretical 95% confidence bands.

The envelope plot above differs from the straight line this is a clear indication that the deviance residuals do not follow a normal distribution, implying that the Poisson fitted model is not adequate for the data. It also has it data plotted outside the 95% confidence band.

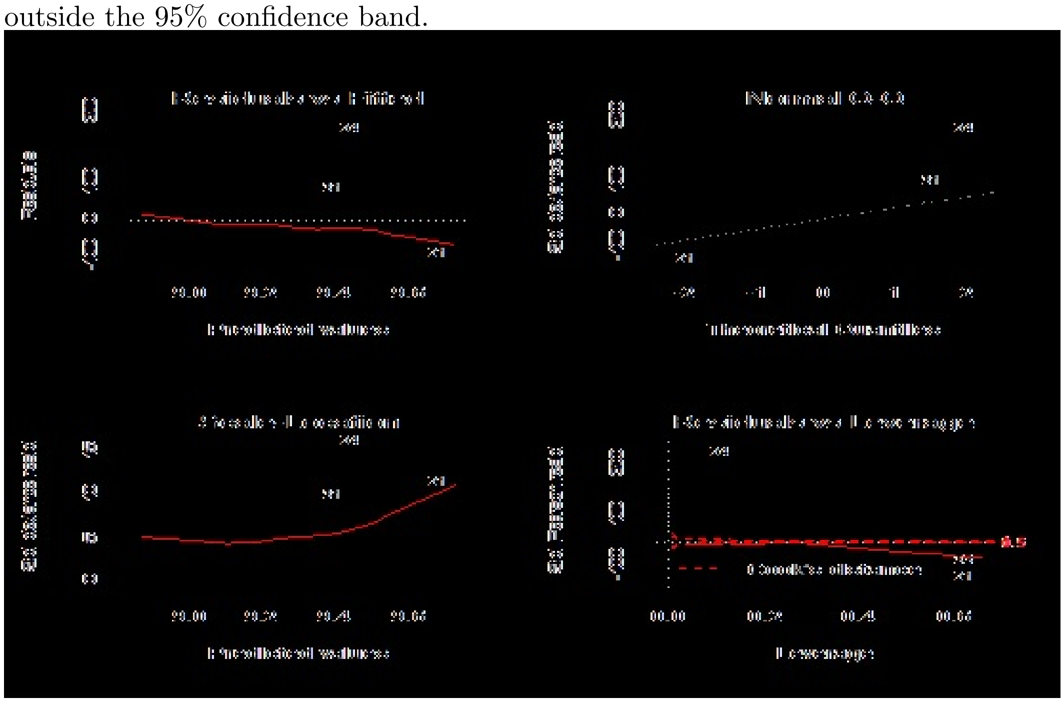


Figure 5: A Plot of Poisson Model for Estimating Maternal Mortaliy in Nigeria. Figure 5 above is a plot of the residual versus fitted model, Normal quartile, scale-location and the residual versus leverage of the Poisson model. The fitting of this model using R statistical package, resulted in a deviance value of 109589, with 31 degrees of freedom, and the ratio of deviance to degree of freedom is greater than 2. Which indicate the existence of overdispersion in the data recommended by Lindsey (1999). This shows that the maternal mortality of each state is not independent from each other. This is due to fact that the geographical locations within the state have some common characteristics. This model does not properly model the overdispersion in the data therefore the Poisson fitted model is not the best to estimate maternal mortality hence we use the alternatives which are the negative binomial, quasi-likelihood, and double Poisson overdispersion model. Also from Table 4 there exist a strong relationship between the independent variables and the dependent variable; the number of maternal mortality within the period of study.

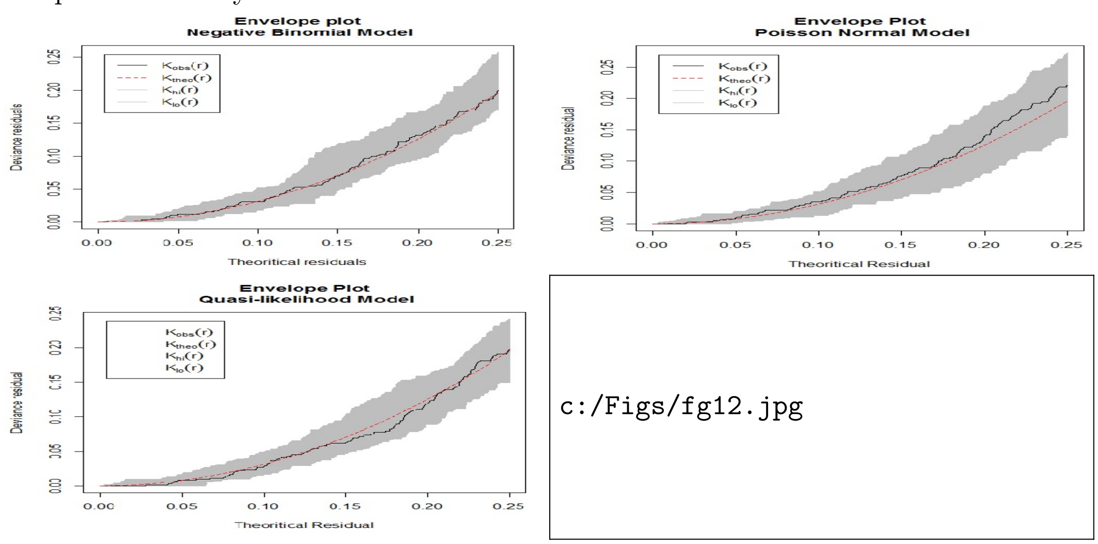


Figure 6: Envelope Plot of theoretical 95% confidence bands for Negative binomial, Poisson normal Quasi-likelihood,and the Double poisson Models.

The envelope plot in Figure 6 shows that the overdispersion in the data as seen in the fitted poisson model envelope plot (Figure 4) does not exist by using the overdispersion model. Also all the model plotted fall within the 95% confidence bands following the normal distribution but the negative binomial model will be the best since it has most of it observed values matching with the theoretical values.

Comparing the different models in Table 5 below, using the AIC, the negative binomial model $AIC_{NB} = 738.92$, quasi-likelihood model $AIC = N/A$, Poisson Normal Model $AIC_{PN} = 15857.6$ and double Poisson overdispersion $AIC_{DP} =$

749.801, it is observed that $AIC_{NB} < AIC_{DP} < AIC_{PN}$ it since the negative binomial model has the minimum AIC, it would be considered a good model to estimate the numbers of maternal mortality during the period under the study.

Table 5: Estimates for the five fitted models for the number of maternal mortality between 2001 and 2008.

		Intercept	**MRW**	**ANTE**	**TET**	**DHP**	**DHF**
POISSON	ESTIMATE	9.184e+00	1.300e-05	4.557e-04	-2.852e-03	5.211e-03	-3.983e-03
AIC=110008	SE	3.129e-03	1.539e-07	3.376e-05	6.029e-05	8.264e-05	7.764e-05
BIC=110018.1	*p*-value	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16
Neg Binomial	ESTIMATE	9.167e+00	1.710e-05	-1.224e-04	-2.311e-03	5.364e-03	-4.105e-03
AIC=738.92	SE	1.410e-01	7.843e-06	1.558e-03	2.643e-03	3.816e-03	3.668e-03
BIC=750.198	*p*-value	<2e-16	0.0293	0.9374	0.3820	0.1599	0.2631
$\widehat{\tau}$= 5.1482							
Quasi-like	ESTIMATE	9.167e+00	1.710e-05	-1.224e-04	-2.311e-03	5.364e-03	-4.105e-03
AIC=N/A	SE	2.098e-01	1.167e-05	2.318e-03	3.931e-03	5.676e-03	5.455e-03
BIC=N/A	*p*-value	<2e-16	0.153	0.958	0.561	0.352	0.457
$\widehat{\sigma}$= 0.4299							
PoissonNormal	ESTIMATE	9.220e+00	2.468e-05	-2.639e-03	2.113e-03	-1.006e-04	1.072e-04
AIC=15857.6	SE	3.178e-03	1.809e-07	3.646e-05	6.439e-05	8.828e-05	8.367e-05
BIC=15868.88	*t*-value	2901.552565	136.382974	-72.384555	32.807897	-1.139873	1.282097
$\widehat{\sigma}$= 0.7623							
DoublePoisson	ESTIMATE	9.203e+00	1.167e-05	1.042e-03	-3.847e-03	7.346e-03	-6.144e-03
AIC=749.801	SE	1.624e-01	8.676e-06	1.630e-03	2.947e-03	3.835e-03	3.522e-05
BIC=759.466	*p*-value	0.0000	0.1786	0.5224	0.1918	0.0554	0.081
$\widehat{\propto}_1$= - 0.002							

5. Summary of Findings

(1) This study has shown that some overdispersion models such as the negative binomial model, quasi-likehood model, Poisson normal model and double Poisson are good model to fit the data.

(2) This study also shows that the negative Binomial model is the best fitted model suitable for estimating infant and maternal mortality in Nigeria; in view of the fact that it has the least Akaike information criterion of the data collected from NBS for estimating maternal mortality in Nigeria.

5.1. Conclusions and Recommendations.

5.1.1. *Infant Mortality.* In order to determine the factors that had a significant effect on the mortality of children under 5 years for the 8-year period 2000–2007, we concluded that the Poisson model was not appropriate. And that, it resulted in incorrect inferences about the regression parameters in the model, suggesting, for example, that the percentage of pregnant women who received anti tetanus injection (PWT) was a negative factor that increased infant mortality.

Finally, the proposed generalization of the Poisson normal model, using a Bayesian approach, was also fitted and conclusions were not quite similar to the selected negative binomial model. That is, the average number of infants immunized with the first & third doze of antigen (OPV3 & MV), had a positive effect and average number of infants immunized with the second doze of antigen (DPT3), had a negative effect on the number of children who died in the 8-year period 2000–2007. But it provided a much better fit in terms of AIC value, so that its proposal and fitting are well motivated and justified.

5.1.2. *Maternal Mortality.* The use of some overdispersion models as suitable models to fit in data for evaluating maternal mortality in Nigeria had been investigated in this work. It was shown that maternal mortality rates for each state in Nigeria are not independent from each other, this suggest that the residuals variability increases with the mean number of maternal death in the period under study. This increase depends on the number of women that received antenatal care which reveal an indication that the women are literate and hence, they are aware of the necessary requirements to avoid maternal death during and after pregnancy.
Finally, the best fitted model for evaluating maternal mortality in Nigeria employing the data constrained from NBS covering 8yrs (2001-2008) is the Negative Binomial model with the least AIC. However, more studies are therefore proposed to determine the fitting of some of these overdispersion models with several variance functions.

References

[1] drián, Q.S, Edilberto, C.C, Vicente, N.A (2012). Estimating infant mortality in Colombia: some overdispersion modelling approaches, Journal of Applied Statistics, 39:5, 1011-1036
[2] tkinson, A. C. (1985). Plots, Transformations and Regression, Clarendon Press, Oxford.
[3] reslow, N.E. (1990). Test of hypotheses in overdispersed Poisson regression and other quasi-likelihood models, J. Amer. Statist. Assoc. 85(**??**), pp. 565–571.
[4] ox, D.R. (1983). Some remarks on overdispersion, Biometrika 70(**??**), pp. 269–274.
[5] ean, C.D. (1992). Testing for overdispersion in Poisson and binomial regression models, J. Amer. Statist. Assoc. 87(**??**), pp. 451–457.
[6] fron, B. (1986). Double exponential families and their use in generalized linear regression, J.Amer. Statist. Assoc. 81(**??**), pp. 709–721.
[7] ordon, K.S. (1989). Generalized linear models with varying dispersion, J. R. Stat. Soc. Ser. B 51(**??**), pp. 47–60.
[8] inde, J. and Demétrio C.G.B. (1998). Overdispersion: Models and estimation, Comput. Stat. Data Anal. 27, pp. 151–170.
[9] beh, C. C. (2008). Is poor maternal mortality index in Nigeria a problem of care utilization? A case study of Anambra State. Afr. J. Reprod. Health, 12(**??**): 132-140.
[10] nam, S. N. and Khan, S. (2002), Importance of Antenatal Care in reduction of Maternal Morbidity and Mortality. Department of Community Health Sciences, Ziauddin Medical University, Karachi.
[11] ee, Y and Nelder, J.A (2002). Two ways of modelling overdispersion in non-normal data. Journal of Royal Statistical Society (Applied Statistics), vol. 49 Issue 4.
[12] indsey, J.K. (1999). The use of corrections for overdispersion, Appl. Stat. 48(**??**), pp. 553–561.
[13] argolin, B.H., Kaplan, N. and Zeiger, E. (1981), Statistical analysis of the Ames Salmonella/ mocrosome test, Proc. Natl. Acad. Sci. 76, pp. 3779–3783.
[14] athew, J.L. (2004). Effect of maternal antibiotics on breast feeding infants, Postgrad. Med. J. 80, pp. 196–200.

[15] cDonagh, M. (1996), Is antenatal care effective in reducing maternal morbidity and mortality? *Health Policy and Planning* 11,1-15.
[16] ondal, S.K.(1997), Utilization of antenatal care services in Rajasthan: Observations from NFEIS. J. Fam. Welfarc,1997; 43 28-33.
[17] ational Bureau of statistics, Nigeria (2010). GeoHive – Nigeria Population Statistics, http://www.xist.org/
[18] elder, J. and Wedderburn, R.W.M. (1972). Generalized linear models, J. R. Stat. Soc. Ser. A 135(**??**), pp. 370–384.
[19] elder, J.A. and Pregibon, D. (1987). An extended quasi-likelihood function, Biometrika 74(**??**), pp. 221–232.
[20] lalekan, A. U., Mubashir, B. U. and Ismail, Y. (2008), A population-based study of effect of multiple birth on infant mortality in Nigeria. BMC Pregnancy and Childbirth, www.biomedcentral.com/1471-2393/8/41
[21] numere, O. (2010). Averting maternal mortality in Nigeria. thewillnigeria.com. Retrieved on 19/09/2010.
[22] aula, A.P. (2004). Modelos de Regressao con Apoio Computacional. Universidade de Sao Paulo, Sao Paulo, 1st ed., pp. 7-15.
[23] chall, R. (1991). Estimation in generalized linear models with random effects, Biometrika; pp 719-727.
[24] edderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method, Biometrika 61(**??**), pp. 439–447.
[25] illiams, D.A. (1982). Extra-binomial variation in logistic linear models, Appl. Stat. 31(**??**), pp. 144–188.

Appendix

Here, the instructions, data and packages in R used to fit the models presented in this paper are given.

```
## Infant Mortality Models##
##
# Poisson model #
Pmodel=glm( AND  NLB + OPV3 + DPT3 + MV + PWC + PWT, family=poisson)
#
summary(Pmodel)
#
# Negative binomial model #
library(MASS)
PmodelBN=glm.nb( AND  NLB + OPV3 + DPT3 + MV + PWC + PWT)
#
summary(PmodelBN)
#
# Quasi-likelihood model #
PmodelQL=glm( AND  NLB + OPV3 + DPT3 + MV +
PWC + PWT, family=quasi(link=log,variance="mu2"))
#
summary(PmodelQL)
```

```
#
# Poisson normal model #
library(statmod) ; library(npmlreg)
PmodelPN=alldist( AND  NLB + OPV3 + DPT3 + MV + PWC + PWT, random=1, family=poisson(link
#
summary(PmodelPN)
#
# Double Poisson model #
library(dglm)
PmodelED=dglm( AND  NLB + OPV3 + DPT3 + MV + PWC + PWT, dformula=PWC, family=poisson)
#
summary(PmodelED)#
> datos
```