

On the Performance of the Poisson, Negative Binomial and Generalized Poisson Regression Models in the Prediction of Antenatal Care Visits in Nigeria

Oyindamola B. Yusuf^{1,*}, Linda O. Ugalahi²

¹Department of Epidemiology and Medical Statistics, University of Ibadan, Oyo State, Nigeria (Senior lecturer in Medical Statistics)

²Department of Epidemiology and Medical Statistics, University of Ibadan, Oyo State, Nigeria (Postgraduate trainee in Medical statistics)

Abstract The number of visits to an antenatal care (ANC) facility is a count variable and it is often modeled by the Poisson regression, however it is sometimes erroneously used in situations where over dispersion (i.e. the variance of the response variable exceeds the mean) occurs. Negative Binomial and Generalised Poisson regression models are alternative models for estimating regression parameters in the presence of over dispersion. This analysis compared Poisson, Negative Binomial and Generalized Poisson regression models to determine the best statistical model which describes the utilisation of ANC visits. A nationally representative sample of women within reproductive age (15-49 years) within households in communities was obtained from the 2013 National Demographic and Health Survey (NDHS) dataset. The number of ANC visits was the outcome variable and other explanatory variables include age, region, employment status, wealth index, husband's/partner's employment status, husband's/partner's educational level, place of residence and place of ANC. The Poisson, Negative Binomial and Generalized Poisson regression models were fitted to the data at 5% significance level. The best model was selected based on the values of $-2\log L$, AIC and BIC selection test/criteria. Of the three regression models, Generalized Poisson regression had the least $-2\log L$, AIC and BIC. Age (35-49 years) (IRR = 1.142, 95% CI: 1.053, 1.240) and rural place of residence (IRR=0.910, 95% CI: 0.874, 0.947) were some of the significant predictors of the number of ANC visits. In the presence of over dispersion, Generalised Poisson Regression was the best regression model in identifying factors associated with the number of antenatal care visits.

Keywords Over dispersion, Model comparison, Number of ANC visits

1. Introduction

Maternal mortality is an important pointer to maternal health and well-being of a country. The health/ survival of a child and the welfare of a family is affected by maternal mortality and morbidity [1]. Maternal mortality is defined by the World Health Organization as "the death of a woman while pregnant or within 42 days of a termination of a pregnancy irrespective of the duration and the site of the pregnancy, from any cause related or aggravated by the pregnancy or its management but not from accidental and incidental causes" [2].

The Millennium Development Goal 5 (MDG 5) focuses on improving maternal health and reducing maternal mortality by 75 percent by 2015 from what it was in 1990. Subsequently in 2006, the introduction of a second target was included which is to achieve universal access to health

facilities by reproductive age women (15-49yrs) by 2015 [1].

In 2008, the estimated number of women getting prenatal care was about 58 percent in Nigeria according to the World Bank, (2013) [3]. These are the women who attended at least one visit connected to or during pregnancy and were attended to by a skilled health worker. The women in urban areas utilized ANC visits more compared to those in rural areas [4].

Utilization of ANC facility is measured by the number of visits a woman makes to such a facility. The number of ANC visits is a count variable which takes non negative integer values. Poisson regression is widely used to model count outcomes however it is sometimes erroneously used in situations where over dispersion occurs i.e. variance of the response variable exceeds the mean. It is inappropriate to use Poisson regression as it underestimates the standard error and exaggerates the significance of regression parameters [5] [6]. Due to the heterogeneous nature of data, over dispersion usually occur in such count outcome variables. There is more variability around the model's fitted values than is consistent with a Poisson formulation

* Corresponding author:

boyusuf@comui.edu.ng (Oyindamola B. Yusuf)

Published online at <http://journal.sapub.org/ajms>

Copyright © 2015 Scientific & Academic Publishing. All Rights Reserved

i.e. the mean does not equal the variance; over dispersion [7]. To correct this problem, the negative binomial Regression and generalised Poisson regression which are more special cases of the generalised linear model were used to model number of ANC visits.

In this analysis, we applied the Poisson, Negative Binomial and Generalised Poisson regression models to the National Demographic and Health Survey (NDHS 2013) dataset to determine which model best fits the nature of the outcome variable; number of visits to an antenatal facility. It also suggests the Negative Binomial and Generalised Poisson regression models as alternative models for handling over dispersion. Furthermore, these models were compared to ascertain which model best describes over dispersion. The objective of this study was to determine the best statistical model that describes the utilization of antenatal care visits.

2. Methodology

This study is a secondary analysis of the National Demographic and Health Survey (NDHS) dataset, a nationally representative survey conducted in 2013. Administratively, Nigeria has 36 states and Federal Capital Territory (FCT). These were divided into local government areas (LGA) and each LGA was further divided into smaller localities. The 36 states were regrouped by geopolitical location into six zones and using the 2006 Population census implementation, each locality was subdivided into Enumeration Areas (EAs). A complete list of the EAs served as the sample frame of the survey. The sampling technique for the 2013 NDHS was a stratified sample, selected at random in three stages from the sampling frame. At the first stage; each state was stratified into urban and rural areas; this resulted in a list of localities. Second stage; one enumeration area was randomly selected from a selected locality with equal probability selection, the resulting list of households served as sampling frame for the selection of households in the third stage. At the third stage; 45 households were selected in every urban and rural cluster through equal probability systematic sampling using the household listing. More details can be obtained from the NDHS report, 2013 [8].

Simple summary statistics (percentage for categorical variables or mean for continuous variables) for all independent variables was also performed. Sample mean and sample variance of the dependent variable (number of antenatal care visits) was calculated to check for the presence of over dispersion or under dispersion. Pearson chi-square and Deviance was utilized to evaluate the presence of over dispersion. These ascertained whether the data followed a standard Poisson, Negative Binomial or the Generalised Poisson regression models. The Poisson, Negative Binomial and Generalised Poisson regression parameters were

estimated, 95% confidence intervals and incidence rate ratios (IRR) were reported. SPSS version 20 and STATA version 12 were used for analysis.

2.1. The Poisson Regression

Poisson regression is used in situations where the dependent variable is a count variable; the counts are positive integers and has a mean greater than zero. A Poisson distribution is derived as a limiting form of binomial distribution when the number of trials becomes large and the probability of success is small. Poisson regression model expresses the logarithm outcome rate as a linear function of a set of predictors. It makes an important assumption that; for a sample of observations, the mean and the variance of the distribution are equal.

A random variable Z is said to have a Poisson distribution with parameter θ if it takes integer values; $0, 1, 2 \dots \infty$.

The probability distribution is represented as

$$\Pr\{Z = z\} = \frac{\theta^z e^{-\theta}}{z!} \quad (2.1)$$

With mean and variance represented as;

$E(Z) = \theta$ and $\text{var}(Z) = \theta$ respectively.

Writing the parameter θ as a log linear model:

$$\text{Log } \theta_i = x\beta_i \quad (2.2)$$

The Poisson regression which was fitted to the number of visits made to antenatal care facility, θ was expressed for the ten independent variables as shown below, where the explanatory variables were: ($X_1 = \text{Age}$, $X_2 = \text{Region}$, $X_3 = \text{Residence}$, $X_4 = \text{Educational Level}$, $X_5 = \text{Religion}$, $X_6 = \text{wealth index}$, $X_7 = \text{occupation}$, $X_8 = \text{place of antenatal}$, $X_9 = \text{husband's/partner's educational level}$, $X_{10} = \text{husband's/partner's occupation}$).

$$\text{Log } (\theta) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} \quad (2.3)$$

The parameter β which represents the expected change in the logarithm of the mean per unit change in the predictor X_i can be estimated by the maximum likelihood method.

2.2. The Negative Binomial Regression

Under the Poisson regression the mean θ_i , is assumed to be homogenous within the classes, however by defining a specific distribution for θ_i , the classes become heterogeneous. Negative binomial regression model is a more generalised model than the Poisson regression. Making the assumption that the mean θ_i , follows a gamma distribution with mean $E(\theta_i) = \mu_i$ and variance $\text{Var}(\theta_i) = \mu_i v_i^{-1}$. Consider the Poisson distribution $Z_i | \theta_i$ with a conditional mean $E(Z_i | \theta_i) = \theta_i$; the marginal distribution can be shown to follow a Negative Binomial distribution [7].

The probability density function of a negative binomial distribution is represented as:

$$\Pr[Z_i = Z_i] = \sum \Pr(Z_i = Z_i | \theta_i) f(\theta_i) d\theta_i = \frac{\Gamma(Z_i + V_i)}{\Gamma(Z_i + 1)\Gamma(V_i)} \left(\frac{v_i}{v_i + \mu_i} \right)^{v_i} \left(\frac{\mu_i}{v_i + \mu_i} \right)^{z_i} \quad (2.4)$$

Where;

Mean is $E(Z_i) = \mu_i$ and variance is $\text{var}(Z_i) = \mu_i + \mu_i^2 v_i^{-1}$,

V_i is the dispersion parameter.

The maximum log likelihood function of the distribution is expressed as;

$$\log(\beta, \alpha) = \sum_i \left(\sum_{r=1}^{Z_i-1} \log(1 + \alpha r) \right) - Z_i \log(\alpha) - \log(Z_i!) + Z_i \log(\alpha \mu_i) - (Z_i + \alpha^{-1}) \log(1 + \alpha \mu_i) \quad (2.5)$$

The parameters (β, α) , can be estimated by partial differentiation of the maximum log likelihood function with respect to β and α .

The negative binomial regression does not make the assumption of equality of the mean and variance but it corrects for over dispersion that arises when the variance is greater than the conditional mean.

2.3. The Generalised Poisson Regression

The Generalised Poisson regression model adjusts for both under dispersion and over dispersion properties in a dataset. For the dependent variable; number of visits to antenatal care, represented by Z_i , the probability distribution function of a generalised Poisson distribution is given as

$$f(Z_i, \theta_i, \alpha) = \left(\frac{\theta_i}{1 + \theta_i} \right)^{Z_i} \frac{(1 + \alpha Z_i)^{Z_i - 1}}{Z_i} \exp \left(- \frac{\theta_i (1 + \alpha Z_i)}{1 + \alpha Z_i} \right) \quad (2.6)$$

Where; $z = 0, 1, 2 \dots \infty$

The standard Poisson regression model is a special form of the generalised Poisson regression model, when α is zero, the probability function of the generalised Poisson random variable reduces to the Poisson probability function. The positive value of α in the equation above indicates over dispersion while a negative value of α indicates under dispersion property of the distribution.

The mean of the dependent variable is related to the independent variables through the link function; $\theta_i = (x_i \beta)$, which is a simple linear model. This model has the disadvantage of θ_i assuming any real value but a Poisson mean assume only count non negative values. To correct this problem the logarithm of the linear model is used; this gives a link log function:

$\text{Log}(\theta_i) = (x_i \beta)$. Taking the exponential of the model we have; $\theta_i = \exp(x_i \beta)$.

In the link function, x_i is $(k-1)$ dimensional vector of explanatory variables and β is a k dimensional vector of regression parameters and α is a dispersion parameter. The mean and variance of Z_i are given by; $E(Z_i | X_i) = \theta_i$, and $V(Z_i | X_i) = \theta_i(1 + \alpha \theta_i)^2$ respectively [7] [5].

In the generalised Poisson regression model, the parameters (β, α) can be estimated by taking the derivatives of the log likelihood function of the model; this means partial differentiation with respect to β and α of the logarithm function of equation (2.7) below;

$$\log L(\alpha, \beta, Z_i) = \sum_{i=1} \left\{ Z_i \log \left(\frac{\theta_i}{1 + \alpha \theta_i} \right) + (Z_i - 1) \log(1 + \alpha Z_i) - \frac{\theta_i (1 + \alpha Z_i)}{1 + \alpha \theta_i} - \log(Z_i!) \right\} \quad (2.7)$$

2.4. Model Selection

Model selection was used to determine the best statistical model which best approximates reality given the set of data and also minimize loss of information. The following goodness of fit tests were utilized in this analysis for model selection.

2.4.1. Chi-square -2log Likelihood Statistic

The maximized likelihood, L , for a given model is the value of the likelihood function when the parameters are substituted with their maximum likelihood estimates and the statistic $-2\log L$ was used to compare models. It is useful for comparing models fitted to the same set of data as the value

of L depends on the number of observations in the data and $-2\log L$ is a measure of agreement between the model and the data. The larger the maximum likelihood the better the agreement between the model and observed data, but the smaller the value of $-2\log L$ the better the model.

2.4.2. Akaike Information Criterion

This method of model selection is based on a relationship between maximum likelihood estimation and Kullback-Leibler information. This information criterion was developed by Akaike. It is derived under the assumption that the operating models belong to the approximating family.

$$AIC = -2\log L(\theta) + 2K$$

Where $L(\theta)$ is the maximized likelihood function and K is the number of estimated parameters included in the model (number of variables plus the intercept). The log likelihood of the model of the data is the overall fit of the model. The smaller value of the log likelihood indicates a worse fit model. However after comparing different models the model with the minimum AIC value is the best model.

In AIC the compromise occurs between maximized log likelihood $-2 \log L(\theta)$; which is the lack of fit component and k , the number of free parameters estimated within the model which is a measure of the compensation for the bias in the lack of fit when maximum likelihood estimators are used. The term $2k$ is the penalty term, the reason for using the penalty term is to prevent over fitting [9].

AIC uses the log likelihood which is the probability of obtaining the chosen data under the given model; hence it makes sense to choose a model that makes the probability as large as possible. The logarithm does not affect the value but the negative sign does, it means minimizing the value of the statistics.

2.4.3. Algorithm for AIC Cross Validation

The following are the algorithm used for the AIC cross validation:

- 1 Assess AIC for all models.
- 2 Identify the model with the smallest AIC value denoted by AIC_{\min} . This is the best model.
- 3 To further compare the models, the AIC difference will be calculated for each model; $D_i = AIC_i - AIC_{\min}$
- 4 Compute the relative likelihood, $\exp\left[-\frac{1}{2}D_i\right]$ for each model
- 5 Compute Akaike weights w_i for each model, these are normalized relative likelihoods

$$w_i = \frac{\exp\left[-\frac{1}{2}D_i\right]}{\sum_{i=1}^m \exp\left[-\frac{1}{2}D_i\right]}$$

These weights can be interpreted as probabilities; the

probability that the given model is the best model.

2.4.4. Bayesian Information Criterion (BIC)

BIC is similar to AIC; the difference is in the second term which depends on the sample size n .

$$BIC = -2\log(L) + p \log(n)$$

Where L denotes the log likelihood, p the number of parameters and n the number of rating classes or number of observations utilized in the model, the smaller the BIC, the better the model. The derivation of BIC assumes equal priors on each model and uninformative priors on the parameters, given each model [10].

The goal of BIC is to find the best model for prediction using highest posterior probability while the goal of AIC is to identify the model that most plausibly generated the data. Both AIC and BIC can be used whether the models are nested or not.

Information criterion supply information on the strength of evidence for each model, it does not make use of a significance level rather it is based on maximum likelihood. It has a high potential of selecting the best model as it is independent of the order in which the models are computed. In this study all selection models were utilized, this was to ensure agreement between the three model selection methods in order to select the best model.

3. Results

3.1. Poisson Regression Results

Table 1 shows the regression parameters for Poisson regression. The deviance for the model was about 11.547 times larger than the degree of freedom which indicated the possible existence of over dispersion.

The number of ANC visits was higher by approximately 9.5% among respondents aged 20-34 years compared to respondents aged less than 20 years (IRR= 1.095, 95% CI: 1.062, 1.130). Number of ANC visits among respondents from the south east, was higher by 29.9% (IRR=1.299, 95% CI: 1.272, 1.327), compared to number of ANC visits by respondents in the north central zone. However respondents from the north east and north west had lower number of ANC visits by approximately 10.6% (IRR=0.894, 95% CI: 0.875, 0.913) and 32.7% (IRR=0.673, 95% CI: 0.658, 0.688) respectively compared to respondents from north central.

Respondents in rural residence had a 10.5% decrease (IRR=0.895, 95% CI: 0.884, 0.907) in number of ANC visits compared to respondents in urban residence. The number of ANC visits increased by 5.2% among women with secondary education compared to women with no education (IRR=1.052, 95% CI: 1.031, 1.074). Respondents in the richest wealth quintile had a 32.5% increase in number of ANC visits compared to respondents in the poorest wealth quintile (IRR=1.325, 95% CI: 1.286, 1.366).

Table 1. Parameter estimates for Poisson regression

Parameters	β	IRR	SE	p-value	95% CI for IRR	
					Lower bound	Upper bound
Constant	2.417	11.212	0.372	<0.001	10.506	11.965
Age (yrs) less than 20*						
20-34	0.091	1.095	0.017	<0.00	1.062	1.130
35-49	0.168	1.182	0.019	<0.00	1.145	1.221
Region North central*						
North east	-0.112	0.894	0.010	<0.001	0.875	0.913
North west	-0.396	0.673	0.008	<0.001	0.658	0.688
South east	0.262	1.299	0.014	<0.001	1.272	1.327
South south	0.473	1.605	0.016	<0.001	1.573	1.637
South west	0.479	1.615	0.015	<0.001	1.587	1.644
Residence Urban*						
Rural	-0.111	0.895	0.006	<0.001	0.884	0.907
Educational level No education*						
Primary	0.009	1.009	0.010	0.361	0.990	1.029
Secondary	0.051	1.052	0.011	<0.001	1.031	1.074
Higher	-0.008	0.992	0.013	0.559	0.966	1.019
Religion Christianity*						
Islam	-0.0185	0.982	0.007	0.014	0.967	0.996
Traditional	-0.174	0.840	0.029	<0.001	0.786	0.898
Wealth Quintile Poorest*						
Poorer	0.068	1.070	0.014	<0.001	1.042	1.099
Middle	0.074	1.076	0.015	<0.001	1.048	1.106
Richer	0.180	1.197	0.017	<0.001	1.164	1.231
Richest	0.282	1.325	0.020	<0.001	1.286	1.366
Employment Status Unemployed*						
Employed	-0.035	0.966	0.007	<0.001	0.953	0.979
Place of ANC At a home*						
Government facility	-0.156	0.856	0.012	<0.001	0.831	0.880
Private facility	-0.075	0.927	0.014	<0.001	0.900	0.955
Husband's/partners educational level No education*						
Primary	0.001	1.001	0.011	0.939	0.980	1.022
Secondary	-0.004	0.996	0.010	0.715	0.976	1.017
Higher	0.058	1.060	0.016	<0.001	1.035	1.085
Husband's/partners employment status Unemployed*						
Employed	-0.229	0.795	0.016	<0.001	0.764	0.828

*reference category, AIC = 192516.95, BIC =192702.9, -2logL =192466.95

3.2. Negative Binomial Regression Results

Table 2 shows the regression parameters for Negative Binomial regression, the ratio of deviance to degree of

freedom for the model is about 0.70 (approximately 1), this indicates an adjustment for over dispersion. Alpha of the model had a value significantly greater than zero ($\alpha=0.653$, 95% CI = 0.637, 0.671), this showed that the

data was better modelled by negative binomial regression due to the presence of over dispersion. The estimates of the standard error were slightly larger than those of the Poisson regression.

The number of ANC visits was higher by approximately 8.2% among respondents aged 20-34 years compared to

respondents aged less than 20 years (IRR= 1.082, 95% CI: 1.000, 1.170). The number of ANC visits in the different geopolitical zones increased by 32% for south east (IRR=1.324, 95% CI: 1.244, 1.410) compared to the north central zone.

Table 2. Parameter estimates for Negative Binomial Regression

Parameters	B	IRR	Standard error	p-value	95% CI For IRR	
					Lower bound	Upper bound
Constant	2.400	11.005	1.130	<0.001	8.999	13.458
Age (yrs)						
Less than 20*						
20-34	0.078	1.082	0.043	0.048	1.000	1.170
35-49	0.141	1.154	0.048	<0.001	1.063	1.253
Region						
North central*						
North east	-0.114	0.892	0.025	<0.001	0.845	0.943
North west	-0.380	0.683	0.020	<0.001	0.646	0.723
South east	0.282	1.324	0.423	<0.001	1.244	1.410
South south	0.480	1.615	0.050	<0.001	1.521	1.715
South west	0.507	1.659	0.044	<0.001	1.574	1.748
Residence						
Urban*						
Rural	-0.098	0.906	0.018	<0.001	0.872	0.941
Educational level						
No education*						
Primary	0.011	1.011	0.026	0.679	0.961	1.063
Secondary	0.045	1.047	0.029	0.106	0.990	1.106
Higher	0.007	1.001	0.040	0.866	0.932	1.087
Religion						
Christianity*						
Islam	-0.003	0.996	0.022	0.869	0.954	1.041
Traditional	-0.272	0.765	0.071	0.004	0.639	0.917
Wealth Quintile						
Poorest*						
Poorer	0.064	1.067	0.034	0.043	1.002	1.135
Middle	0.050	1.053	0.035	0.119	0.987	1.124
richer	0.133	1.145	0.041	<0.001	1.067	1.229
Richest	0.241	1.276	0.052	<0.001	1.178	1.381
Employment Status						
Unemployed*						
Employed	-0.037	0.964	0.019	0.054	0.928	1.001
Place of ANC						
At a home*						
Government health facility	-0.154	0.857	0.042	<0.001	0.780	0.943
Private health facility	-0.069	0.934	0.047	0.173	0.846	1.031
Husband's/partners educational level						
No education*						
Primary	0.005	1.005	0.028	0.867	0.952	1.061
Secondary	0.027	1.026	0.028	0.339	0.973	1.083
Higher	0.072	1.073	0.035	0.029	1.007	1.144
Husband's/partners employment status						
Unemployed*						
Employed	-0.206	0.813	0.058	0.004	0.707	0.935

*reference category, AIC = 83845.849, BIC = 84039.233, -2logL = 83793.848

Table 3. Parameter estimates for Generalised Poisson regression

Parameters	β	IRR	Standard error	p-value	95% CI for IRR	
					Lower bound	Upper bound
Constant	2.385	10.861	1.231	<0.001	8.698	13.562
Age (years)						
Less than 20*						
20-34	0.075	1.078	0.043	0.057	0.998	1.165
35-49	0.133	1.142	0.477	<0.001	1.053	1.240
Region						
North central*						
North east	-0.115	0.891	0.025	<0.001	0.844	0.941
North west	-0.378	0.686	0.019	<0.001	0.649	0.725
South east	0.290	1.336	0.045	<0.001	1.251	1.428
South south	0.481	1.618	0.054	<0.001	1.516	1.727
South west	0.520	1.682	0.049	<0.001	1.589	1.781
Residence						
Urban*						
Rural	-0.095	0.910	0.019	<0.001	0.874	0.947
Educational level						
No education*						
Primary	0.011	1.011	0.026	0.663	0.961	1.064
Secondary	0.043	1.044	0.030	0.135	0.987	1.104
Higher	0.014	1.013	0.042	0.742	0.934	1.100
Religion						
Christianity*						
Islam	0.003	1.003	0.024	0.896	0.958	1.050
Traditional	-0.303	0.738	0.067	<0.001	0.619	0.882
Wealth Quintile						
Poorest*						
Poorer	0.065	1.067	0.033	0.035	1.005	1.133
Middle	0.044	1.045	0.034	0.172	0.981	1.114
richer	0.117	1.124	0.041	<0.001	1.047	1.206
Richest	0.226	1.254	0.052	<0.001	1.155	1.361
Employment status						
Unemployed*						
Employed	-0.038	0.963	0.019	0.059	0.926	1.001
Place of ANC						
At a home*						
Government health facility	-0.0155	0.857	0.047	0.005	0.769	0.954
Private health facility	-0.0654	0.937	0.054	0.254	0.837	1.048
Husband's/partners educational level						
No education*						
Primary	0.004	1.004	0.028	0.877	0.951	1.060
Secondary	0.040	1.041	0.028	0.144	0.987	1.098
Higher	0.076	1.079	0.035	0.021	1.012	1.151
Husband's/partners employment status						
Unemployed*						
Employed	-0.194	0.824	0.067	0.018	0.702	0.967

*reference category, AIC = 81282.048, BIC = 81475.433, -2logL = 81230.048

Table 4. Model comparison using AIC cross validation algorithm

Model	LogL	K	AIC	D_i	Relative likelihoods	Akaike weights W_i
PR	-96233.475	35	192516.95	111286.902	0.000	0.000
NBR	-41896.924	36	83845.849	2615.801	0.000	0.000
GPR	-40615.024	26	81230.048	0.000	1.000	1.000
					Sum=1	

Respondents in rural areas had a 10.5% decrease (IRR=0.895, 95% CI: 0.884, 0.907) in the number of ANC visits compared to respondents in urban areas. Respondents in the richest wealth quintile had a 27% increase in the number of ANC visits compared to respondents in the poorest wealth quintile (IRR=1.276, 95% CI: 1.178, 1.381). Respondents whose husbands/partners had higher education had a 7.3% increase in the number of ANC visits compared to respondents whose husband/partner had no education (IRR=1.073, 95% CI: 1.007, 1.144).

3.3. Generalised Poisson Regression Results

Table 3 shows the regression parameters for Generalised Poisson regression. The positive value of phi ($\phi=0.182$; 95% CI: 0.178, 0.186) from the model indicated an adjustment for Poisson over dispersion. The estimates of the standard error were similar to those of the Negative Binomial regression, although slightly higher.

The number of ANC visits was higher by approximately 14% among respondents aged 35-49 years compared to respondents aged less than 20 years (IRR = 1.142, 95% CI: 1.053, 1.240). The number of ANC visits among respondents increased by 33.6% for south east (IRR = 1.336, 95% CI: 1.251, 1.428), compared to respondents in the north central zone.

Respondents in rural areas had about 9% decrease (IRR = 0.910, 95% CI: 0.874, 0.947) in the number of ANC visits compared to respondents in urban areas. Respondents in the richest wealth quintile had a 25% increase in the number of ANC visits compared to respondents in the poorest wealth quintile (IRR = 1.254, 95% CI: 1.155, 1.361). Respondents whose husband/partner had higher education had a 7.9% increase in the number of ANC visits compared to respondents whose husbands/partners had no education (IRR=1.079, 95% CI: 1.012, 1.151).

3.4. AIC Cross Validation

Table 4 shows the comparison of the three models PR, NBR and GPR utilising the difference (D_i) between the models and the model with the least value, hence the relative likelihood were computed, this led to Akaike weights computation. Using this cross validation of AIC, GPR model would be ranked best model 100% of the time when comparing these three models i.e. PR, NBR and GPR.

4. Discussion

The standard errors from Poisson regression are smaller than those of Negative Binomial regression, the relatively larger values of NBR standard errors led to some insignificant regression parameters. Similarly larger standard errors of the parameters in the GPR model led to more insignificant regression parameters. The large standard error in NBR and GPR shows that in the presence of over dispersion, the Poisson regression overstates the significance of the regression parameter and the significance of the evaluation factors. This is compatible with findings from other studies [11] [5] [7] [12]. Generalised Poisson regression was the best model selected. This was inferred from the values of the model selection test/criteria utilized. These are $-2\log L$, AIC and BIC. All these criteria established the GPR as the best model because it had the smallest value of all three selection criteria. The cross validation of AIC agrees with the selection of GPR as the best model for count data in the presence of over dispersion.

This study demonstrated that GPR model is the best model to determine the factors that predict the number of antenatal care visits to a health care facility, when there is an indication of the presence of over dispersion. It is recommended that objective criteria should be used to select appropriate statistical models for analysing count data in the presence of over dispersion.

REFERENCES

- [1] S. H. Adamu, Utilisation of Maternal Health Care Services in Nigeria: An Analysis of Regional Differences in the Patterns and Determinants of Maternal Health Care Use, 2011. Available at: <http://mph/MPHQuantitativeDisertation>.
- [2] WHO. Making pregnancy safer: the critical role of the skilled attendant. A joint statement by WHO, ICM and FIGO. Geneva, Switzerland WHO. 2000, Available at: <http://whqlibdoc.who.int/publications/2004/9241591692.pdf>.
- [3] World Bank. Pregnant women receiving prenatal care (%). The World Bank Data. Available at: <http://data.worldbank.org/indicator>.
- [4] B. I. Babalola, T. O. Adeyoju, and A. Makumi, Determinants Of Urban-Rural Differentials in Antenatal care utilisation in

- Nigeria, presented at Population Association of America New Orleans, April 2013.
- [5] N. Ismail and A.A. Jemain, "Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models," *Casualty Actuarial Society Forum Winter*, 2007 pp. 103–158.
- [6] G. Sileshi, "Selecting the right statistical model for analysis of insect count data by using information theoretic measures". *Bulletin of Entomological Research*, 2006 vol. 96 pp.479-488.
- [7] F. Sadia, Performance of Generalized Poisson Regression Model and Negative Binomial Regression Model in case of Over-dispersion Count Data, 2013, pp. 558–563.
- [8] National Population Commission. Nigeria Demographic and Health Survey: Preliminary Report, 2012. Available from: <http://dhsprogram.com/pubs/pdf/PR41/PR41.pdf>
- [9] H. Bozdogan, "Akaike's Information Criterion and Recent Developments in Information Complexity," *Journal of Mathematical Psychology*, 2000 vol. 44 pp. 62–91.
- [10] W. Zucchini, "An Introduction to Model Selection." *Journal of Mathematical Psychology*, 2000 vol. 44 pp. 41–61.
- [11] B.E.L. Piza, Using Poisson and Negative Binomial Regression Models to Measure the Influence of Risk on Crime Incident Counts. 2012 Available at: <http://rutgerscps.weebly.com/uploads/2/7/3/7/27370595/countregressionmodels.pdf>.
- [12] M.M. Islam, M. Alam, M. Tariqzaman, M.A. Kabir, R. Pervin, M. Begum, et al. Predictors of the number of under-five malnourished children in Bangladesh: application of the generalized poisson regression model. *BMC Public Health*, 2013, vol. 13 pp. 11. Available at: <http://www.pubmedcentral.nih.gov>.