Assessing Logistic and Poisson Regression Model in Analyzing Count Data

Ijomah, M. A., Biu, E. O., & Mgbeahurike, C.
University of Port Harcourt,
Choba, Rivers State
zubikeijomahs@yahoo.com, emmanuelbiu@yahoo.com

Abstract
An examination of the relationship between a response variable and several predictor variables were considered using logistic and Poisson regression. The methods used in the analysis were descriptive statistics and regression techniques. This paper focuses on the household utilized/ not utilizes primary health care services with a formulated questionnaire, which were administered to 400 households. The statistical Softwares used are Microsoft Excel, SPSS 21 and Minitab 16. The result showed that the Logistic regression model is the best fit in modelling binary response variable (count data); based on the two assessment criteria employed [Akaike Information Criterions (AIC) and Bayesian Information Criterions (BIC)].

Keywords: *Binary response variable, model selection criteria, Logistic and Poisson regression model*

1. Introduction

As a statistical methodology, regression analysis utilizes the relation between two or more quantitative variables, that is, a response variable can be predicted from the other(s). This methodology is widely used in business, social, behavioural and biological sciences among other disciplines, Michael *et al*. (2005). The two types of regression are Linear and Nonlinear regression.

The different types of linear regression are simple and multiple linear regression (Nduka, 1999) while the Nonlinear regressions are log-linear, quadratic, cubic, exponential, Poisson, logistic and power regression. Notably, our interests in this research are the Poisson regression and Logistic regression.

Poisson regression is useful when the outcome is a count. It is used to estimate rates or counts comparing different exposure groups in the same way that logistic regression is used to estimate odds ratios comparing different exposure groups.

Also, the logistic and Poisson regression are used to determine which variables are important and what is the direction of the effect for each variable. These models allow analysts to take account of the knowledge present in a set of observations between the dependent variable and independent variables (Armstrong, 2012).

The general form of Poisson regression model is similar to logistic regression and multiple regression models. Applications of the Poisson distribution can be found in many fields related to counting such as; 1) Telecommunication; 2) Biology; 3) Radioactivity; etc.

Similarly, logistic regression is an essential model considered for use when the response variable is binary with two possible outcomes, such as financial status of firm (profit or loss), blood pressure: (High or low) etc. Both models are appropriate for analyzing

data arising from either observational or experimental studies (Michael *et al*, 2005). This paper work considered Poisson and Logistic regression models because the response outcomes obtained are discrete (or binary response variable).

The aim of this study is to assess Poisson and logistic regression analysis using data on Primary Health Service and to determine which regression model is appropriate to investigate the response variable effects on more than one predictive variable (i.e. Logistic or Poisson regression analysis). The objectives are; 1) To estimate a suitable Poisson and logistic regression model. 2) To determine the odd ratio for a unit change in the predictor. 3) To compare the model and parameter estimates of Poisson and logistic regression.

This study focuses on finding out the effects of response variable (when the response variable is a binary count variable) on the predictors; using data on Primary Health Care Service in Choba (Obio/Akpor) Local Government Area, Rivers State. Therefore, it is necessary to report the limitations of the study as a way of pointing out the extent to which the finding may be generalized: (1) The sample is limited to 400 (four hundred) Households in Choba (Obio/Akpor) Local Government Area; (2) The sampling is irrespective of social stratification; and (3) The study is limited to logistic and Poisson regression. The target population of this study was four hundred (400) household. However, four hundred and twenty (420) questionnaires were administered to households and based on the questionnaires retrieved (or returned), 400 questionnaires were considered for this study. The remaining part of the paper is organized as follows: Section two provides related literature followed by the description of the two methods (logistic and Poisson regression) and data used in the Section three. The numerical analysis and results in section four. Section five concludes and provides recommendation.

2. Review of Poisson and Logistic Regression

Greene (2003) said that Poisson regression may be appropriate when the dependent variable is a count, for instance event such as the arrival of telephone call at a call centre. The events must be independent in the sense that the arrival of one call will not make another more or less likely, but the probability per unit time of event is understood to be related to covariates such as time of day.

Berk (2003) stated that count data are common in criminological research. When the response variable is a count, one option is to employ Poisson regression as a special case of the generalized linear model. Poisson formulation is relatively simple to interpret because the right hand side is the familiar linear combination of predictors and because when exponential, the regression coefficients are interpreted as multipliers. Poisson regression applications have been published by a number of respected criminologist; Paternoster and Brame (1997), Sampson and Laub (1997); Osgood (2000).

According to Yanqiu *et al.* (2009), Poisson regression was used to study time trends and regional differences in maternal mortality (RMM) in China from 2000-2005 and found that RMM declined by an average of 5% per year. Here, Poisson regression model is used to examine the incidence of maternal mortality at the hospital. The Poisson model assumes that the variance of the count data is equal to the mean (Agresti, 2007). The coefficients of the Poisson regression model are estimated using the maximum likelihood techniques. The deviance (likelihood ratio) test statistic, G^2, is used to assess the adequacy of the fitted model.

According to Michael *et al.* (2005), a Poisson regression is useful when the outcome is a count, with large-count outcomes being rare event. For instance, the number of times a household shops at a particular supermarket in a week is a count, with a large number of shopping trips to the store during the week being a rare event.

However, according to Kleinbaun (1994), logistic regression is identified as the most popular method used in analyzing epidemiological data when the outcome variable is binary. The response variable is coded with the value 0 or 1 and it is used in categorical data.

Logistic regression provides a method for modeling a binary response variable. For example, we may wish to investigate how death (1) or survival (0) of patients can be predicted by a level of one or more metabolic markers. Logistic regression makes no assumption about the distribution of the independent variables. The relationship between the predictor and response variable is not a linear function in logistic regression. Despite this, there are many distribution functions that have been proposed for use in the analysis of a dichotomous variable. Cox and Snell (1989) discussed some of these.

Although the statistical properties of linear regression models are invariate to the (unconditional) means of the dependent variable, the same is not true for binary dependent variable model. The mean of binary variable is the relative frequency of event in data, which in addition to the number of observations, constitutes the information content of the data set.

According to McCaullagh and Nelder (1992), a logistic regression is considered as a parametric model and is a form of generalized linear model. This is because the probability distribution for the response variable is specified as well as the error terms. Logistic regression makes use of several predictor variables which may be categorical or numerical. The odds ratio is usually of interest in a logistic regression due to its ease of interpretation. Odds ratio is a statistic that measures the odds of an events compared to the odds of another event [for 2 x 2 contingency table, the odds ratio is a measure of association (Agresti, 2007)]. Combination of the odds and the logistic regression leads to the interpretation of any logistic regression result (Hosmer and Lemeshow, 1989).

The Logistic regression model has been used in many discipline including medical studies; Devita *et al.* (2008). It has been used in the social research [Ingeles *et al*. (2009)] also an important tool at the commercial applications and in Medical studies. The dependents variable of the logistic model is classified into two basic types (Afifi *et al*. 2004).

A large sample size is needed for testing of hypothesis in logistic regression since it does not require much assumption for the hypothesis to be accurate. This is because of the nature of probabilities which logistic regression principles are based. A logit transformation is used [Grimms and Yarnold, (1995); Grizzle *et al*. (1969)].

In logistic regression interpretation, two other similar statistically equivalent tests have been suggested. These are the Wald Test and score Test. The assumptions needed for these tests are the same as those of the likelihood ratio test. The Wald test is obtained by comparing the maximum likelihood estimate of the slope parameter, β_1 to an estimate of its standard error.

Hauck and Donner (1997) examined the performance of the Wald test and likelihood ratio test. They found that Wald test behaved in an aberrant manner, often failing to reject the null hypothesis when the coefficient was significant. Therefore, they recommended the likelihood ratio.

Jennings (1986) has also looked at the adequacy of inference on logistic regression based on Wald statistic. His conclusions are similar to those of Hauck and Donner (1997). Both the likelihood ratio test (G) and the Wald test (W), require the computation of maximum likelihood estimate for β_1.

However, Christensen (1997) gave the following warning about the Hosmer and Lemeshow (2000) goodness of fit test; if too few groups are used to calculate the statistics (<5) it will always indicate that the model fits the data. That is why Hosmer and Lemeshow

(2000) advocated that before finally accepting that a model fits, an analysis of the individual residuals and relevant diagnostic statistics be performed.

Hosmer and Lemeshow (2000), highlighted that it is possible to construct a model that fits the data (good estimation of the relationship between response and explanatory variables) but is a poor predictive model.

According to Michael *et al* (2005), logistic regression is an important nonlinear regression model and could be considered for use when the response variable is qualitative with two possible outcomes, such as financial status of firm (sound status, headed towards insolvency) or blood pressure status (high blood pressure, low blood pressure). Logistic nonlinear regression model is appropriate for analyzing data arising from either observational studies or from experimental studies (such as in this study).

The two main uses of logistic regression include:

1. The first is the prediction of group membership. Since logistic regression calculates the probability of success over the probability of failure, the results of the analysis are in the form of odds ratio.
2. Logistic regression also provides knowledge of the relationships and strengths among the variables. In any regression problem, the key quantity is the mean value of the outcome variable, given the value of the independent variable. The quantity is called the conditional mean and will be expressed as "E(Y/X)"; where Y denotes the outcome variable and X denotes a value of the independent variable and with dichotomous data, the conditional mean must be greater than or equal to zero and less than or equal to 1[$0 \leq E(y/x) \leq 1$].

There are two primary reasons for choosing the logistic distribution namely, (a) the mathematical point of view, it is an extremely flexible and easily used function. (b) It tends itself to a clinically meaningful interpretation. Testing for significance in logistic regression, the overall significance is based upon the value of G test statistic and this is commonly referred to as the deviance statistic. Interpreting a regression equation involves relating the independents variables to the dependent variable that the equation was developed to answer. However, with logistic regression, it is difficult to interpret the relation between the independent variables and that probability that Y = 1 directly because the logistic regression equation is non-linear. However, Statisticians have shown that the relationship can be interpreted indirectly using a concept called the odds ratio. The odds in favour of an event occurring is defined as the probability that the event will occur divided by the probability that the event will not occur. In logistic regression, the event of interest is always Y = 1.

Logistic regression forms a best fitting equation or function using the maximum likelihood method, which maximizes the probability of classifying the observed data into the appropriate category given the regression coefficients.

Therefore, the first goal of Poisson and logistic regression analysis is the statistical significance of certain variable and how they affect the response, whereas the latter is more concerned with the ability to accurately and efficiently predict the response. Hosmer and Lemeshow (2000) highlighted that it is possible to construct a model that fits the data (good estimation of the relationship between response and explanatory variables) but is a poor predictive model. In this research, we compare the model and parameter estimates of Poisson and logistic regression analysis; then determine which of the regression model fits the data set considered.

3. Methods of evaluation

The methods considered in this study are logistic and Poisson regression models

where the error terms are normally distributed and the response outcomes are discrete [Michael, et al., (2005) and Pregibon, (1981)].

3.1 Mathematical Expression of Logistic and Poisson Regression Models

3.1.1 Logistic Regression models with Binary Response Variable

Consider the simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \qquad Y_i = 0, 1 \tag{3.1}$$

where the outcome Y_i is binary, taking on the value of either 0 or 1. The expected response $E(Y_i)$ has a special meaning in this case. Since $E(\varepsilon_i) = 0$ we have:

$$E(Y_i) = \beta_0 + \beta_1 X_i \tag{3.2}$$

Consider Y_i to be a discrete random variable for which we can state the probability distribution as follows:

Table 3.1: Probability of Y_i

Y_i	**Probability**
1	$P(Y_i = 1) = \pi_i$
0	$P(Y_i = 0) = 1 - \pi_i$

Thus, π_i is the probability that $Y_i = 1$ and $1 - \pi_i$ is the probability that $Y_i = 0$. By the definition of expected value of a random variable in Equation (3.2), we obtain

$$E(Y_i) = 1(\pi_i) + 0\,(1 - \pi_i) = \pi_i = P(Y_i = 1) \tag{3.3}$$

Equaling Equation (3.2) and (3.3), we thus have

$$E(Y_i) = \beta_0 + \beta_1 X_i = \pi_i \tag{3.4}$$

Then, the logistic mean response function is

$$E(Y_i) = \pi_i = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)} \tag{3.5}$$

- **Likelihood Estimation**

Let the discrete random variable Y_i be Bernoulli random variable, and each Y_i observation is an ordinary Bernoulli random variable where:

$$\left.\begin{aligned} P(Y_i = 1) &= \pi_i \\ P(Y_i = 0) &= 1 - \pi_i \end{aligned}\right\} \tag{3.6}$$

Then, its probability distribution is representing as follows:

$$f_i(Y_i) = \pi_i^{Y_i} (1 - \pi_i)^{1 - Y_i}, \quad Y_i = 0, 1, \; i = 1, 2, \ldots, n \tag{3.7}$$

Note that $f_i(1) = \pi_i$ and $f_i(0) = 1 - \pi_i$. Hence, $f_i(Y_i)$ simply represents the probability that $Y_i = 1$ *or* 0. Since the Y_i observation are independent. Their joint probability function is:

$$L(Y_1, Y_2, \ldots, Y_n) = \prod_{i=1}^{n} f_i(Y_i) = \prod_{i=1}^{n} \pi_i^{Y_i} (1 - \pi_i)^{1 - Y_i} \tag{3.8}$$

Taking logarithm of Equation (3.8), then the joint probability function:

$$\log_e L(Y_1, Y_2, \ldots, Y_n) = \log_e \prod_{i=1}^{n} \pi_i^{Y_i} (1 - \pi_i)^{1 - Y_i} = \sum_{i=1}^{n} \left[Y_i \log_e \left(\frac{\pi_i}{1 - \pi_i} \right) \right] + \sum_{i=1}^{n} \log_e (1 - \pi_i) \tag{3.9}$$

Since $E(Y_i) = \pi_i$ for a binary variable, it follows from Equation (3.5) that:

$$1-\pi_i = \left[1 + \exp\left(\beta_0 + \beta_1 X_i\right)\right]^{-1} \quad (3.10)$$

Furthermore, from Equation (3.5), we obtain

$$\log_e\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 X_i \quad (3.11)$$

Hence, Equation (3.9) can be expressed as follows:

$$\log_e L\left(\beta_0, \beta_1\right) = \sum_{i=1}^{n}\left[Y_i\left(\beta_0 + \beta_1 X_i\right)\right] - \sum_{i=1}^{n}\log_e\left(1+\left(\beta_0 + \beta_1 X_i\right)\right) \quad (3.12)$$

where $L\left(\beta_0, \beta_1\right)$ replaces $L\left(Y_1, Y_2, ..., Y_n\right)$, to show explicitly that this function is now viewed as the likelihood function of the parameter to be estimated, given the sample observation.

- **Maximum Likelihood Estimation**

The maximum likelihood estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$ in the simple logistic regression model are those values of β_0 and β_1 that maximize the log-likelihood function in Equation (3.13). Computer intensive numerical search procedures are therefore required to find the maximum likelihood estimates of b_0 and b_1. Once the maximum likelihood estimates b_0 and b_1 are found, we substitute these values into the response function in Equation (3.5) to obtain the fitted response function. We shall use $\hat{\pi}_i$ to denote the fitted value for the i^{th} case.

$$\hat{\pi}_i = \frac{\exp\left(b_0 + b_1 X_i\right)}{1 + \exp\left(b_0 + b_1 X_i\right)} \quad (3.13)$$

The fitted logistic response function is as follow:

$$\hat{\pi} = \frac{\exp\left(b_0 + b_1 X\right)}{1 + \exp\left(b_0 + b_1 X\right)} \quad (3.14)$$

If we utilize the logit transformation in (3.4), we can express the fitted response function in (3.14) as follow:

$$\pi' = b_0 + b_1 X \quad (3.15)$$

where

$$\pi' = \log_e\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) \quad (3.16)$$

Equation (3.15) is called the filled logit response function. Once the fitted logistic response function has been obtained, the next steps are to examine the appropriateness of the fitted response function and, if the fit is good to make a variety of inferences and predictions.

Multiple Logistic Regression Model

The simple logistic regression model (3.15) is easily extended to more than one predictor variable. In fact, several predictor variables are usually required with logistic regression to obtain adequate description and useful predictions.

In extending the simple logistic regression model, we simply replace $\beta_0 + \beta_1 X_1$ in Equation (3.14) by $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_{p-1} X_{p-1}$. To simplify the formula, we use matrix notation and the following three vectors:

$$\beta_{p\times 1}=\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ .\\ .\\ .\\ \beta_{p-1}\end{bmatrix}\quad X_{p\times 1}=\begin{bmatrix}1\\ X_1\\ X_2\\ .\\ .\\ .\\ X_{p-1}\end{bmatrix}\quad X_{i\,p\times 1}=\begin{bmatrix}1\\ X_{i1}\\ X_{i2}\\ .\\ .\\ .\\ X_{i\,p-1}\end{bmatrix} \tag{3.16}$$

We have

$$X'\beta=\beta_0+\beta_1X_1+\beta_2X_2+\ldots+\beta_{p-1}X_{p-1} \tag{3.17}$$

$$X_i'\beta=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\ldots+\beta_{p-1}X_{i\,p-1} \tag{3.18}$$

From Equation (3.17) and (3.18), the simple logistic function (3.5) extends to the multiple logistic response function as follows:

$$E(Y)=\frac{\exp(X'\beta)}{1+\exp(X'\beta)} \tag{3.19}$$

and the equivalent simple logistic response form Equation (3.10) extend to:

$$E(Y)=[1+\exp(-X'\beta)]^{-1} \tag{3.20}$$

Similarly, the logit transformation (3.11):

$$\pi'=\log_e\left(\frac{\pi}{1-\pi}\right) \tag{3.21}$$

Now leads to the logit response function, or linear predictor:

$$\pi'=X'\beta \tag{3.22}$$

The multiple logistic regression model can be stated as follows:

$$E(Y_i)=\pi_i=\frac{\exp(X_i'\beta)}{1+\exp(X_i'\beta)} \tag{3.23}$$

Note: when the logistic regression model contains only qualitative variables, it is often referred to as a log-linear model.

However, fitting of model utilize the method of maximum likelihood to estimate the parameter of the multiple logistic response function (3.23). The log-likelihood function for simple logistic regression in (3.12) extends directly for multiple logistic regression:

$$\log_e L(\beta)=\sum_{i=1}^{n}[Y_i(X_i'\beta)]-\sum_{i=1}^{n}\log_e(1+(X_i'\beta)) \tag{3.24}$$

Numerical search procedures are used to find the values of $\beta_0,\beta_1,\ldots,\beta_{p-1}$ the maximize $\log_e L(\beta)$. These maximum likelihood estimates will be denoted by $b_0,b_1,\ldots,b_{p-1}$. Let b denote the vector of the maximum likelihood estimates:

$$b_{p\times 1}=\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ . \\ . \\ . \\ b_{p-1} \end{bmatrix} \quad (3.25)$$

The fitted logistic response function and fitted values can be expressed as follows:

$$\hat{\pi}=\frac{\exp\left(\mathrm{X}'b\right)}{1+\exp\left(\mathrm{X}'b\right)}=\left[1+\exp\left(-\mathrm{X}'b\right)\right]^{-1} \quad (3.26)$$

$$\hat{\pi}_i=\frac{\exp\left(\mathrm{X}_i'b\right)}{1+\exp\left(\mathrm{X}_i'b\right)}=\left[1+\exp\left(-\mathrm{X}_i'b\right)\right]^{-1} \quad (3.27)$$

where

$$\mathrm{X}'b=b_0+b_1\mathrm{X}_1+b_2\mathrm{X}_2+\ldots+b_{p-1}\mathrm{X}_{p-1} \quad (3.28)$$

$$\mathrm{X}_i'b=b_0+b_1\mathrm{X}_{i1}+b_2\mathrm{X}_{i2}+\ldots+b_{p-1}\mathrm{X}_{i\ p-1} \quad (3.29)$$

In this paper, we shall rely on standard statistical packages for Logistic regression to conduct the numerical search procedures for obtaining the maximum likelihood estimates, such as SPSS 21 and Minitab 16.

3.1.2 Poisson Regression Model

Poisson distribution outcomes are counts $\left(\mathrm{Y}_i=0,1,2,\ldots,\right)$, with a large count or frequency being a rare event. The Poisson probability distribution is as follows:

$$f\left(\mathrm{Y}\right)=\frac{\mu^{\mathrm{Y}}\exp\left(-\mu\right)}{\mathrm{Y}!};\ \mathrm{Y}=0,1,2,\ldots \quad (3.30)$$

where

$f\left(\mathrm{Y}\right)$ denotes the probability that the outcome is Y and $\mathrm{Y}!=\mathrm{Y}(\mathrm{Y}-1)\ldots 3.2.1$.

The mean and variance of the Poisson probability distribution are:

$$\mathrm{E}(\mathrm{Y})=\mu \quad (3.31)$$

$$\sigma^2(\mathrm{Y})=\mu \quad (3.32)$$

Note that the variance is the same as the mean. Hence, if the number of trips follows the Poisson distribution and the mean number of store trips for a family with three children is larger than the mean number of trips for a family with no children, the variance of that distribution of outcomes for the two families will also differ. At times, the count response Y will pertain to different units of time or space, then Poisson probability distribution is expressed as follow:

$$f\left(\mathrm{Y}\right)=\frac{(t\mu)^{\mathrm{Y}}\exp\left(-t\mu\right)}{\mathrm{Y}!};\ \mathrm{Y}=0,1,2,\ldots \quad (3.33)$$

where

μ denote the mean response for Y for a unit of time or space (e.g. one month).

t denote the number of units of time or space to which Y corresponds.

Y is the number of store trips during the month (e.g. Y is the number of store trips during one week where the unit time is one month)for a unit of time or space. Note: all

response Y_i pertains to the same unit of time or space.

Poisson Regression Model: Like any nonlinear regression model, can be stated as follows:

$$Y_i = E(Y_i) + \varepsilon_i \;\; ; \;\; i = 0,1,2,... \tag{3.34}$$

The mean response for the i[th] case, to be denoted now by μ_i for simplicity, is assumed as always to be a function of the set of predictor variables: $X_1, X_2, ..., X_{p-1}$. We use the notation $\mu(X_i, \beta)$ to denote the function that relates the mean response μ_i to X_i, the values of the predictor variables for case i, and β, the values of the regression coefficients. The commonly used functions for Poisson regression are:

$$\mu_i = \mu(X_i, \beta) = X_i'\beta \tag{3.35}$$

$$\mu_i = \mu(X_i, \beta) = \exp(X_i'\beta) \tag{3.36}$$

$$\mu_i = \mu(X_i, \beta) = \log_e(X_i'\beta) \tag{3.37}$$

In all three cases, the mean response μ_i must be nonnegative. Since the distribution of the error terms ε_i for Poisson regression is a function of the distribution of the response Y_i which is Poisson regression model in the following form:

$$\mu_i = \mu(X_i, \beta) \tag{3.38}$$

where Y_i are independent Poisson random variables with expected values μ_i.

The most commonly used response function is $\mu_i = \exp(X_i'\beta)$, also used in this study.

-Maximum likelihood Estimation

For Poisson regression model (3.38), the likelihood function is as follows:

$$L(\beta) = \prod_{i=1}^{n} f_i(Y_i) = \prod_{i=1}^{n} \frac{(\mu(X_i, \beta))^Y \exp(-\mu(X_i, \beta))}{Y_i!} \tag{3.39a}$$

or

$$L(\beta) = \frac{\prod_{i=1}^{n} (\mu(X_i, \beta))^Y \exp(-\mu(X_i, \beta))}{\prod_{i=1}^{n} Y_i!} \tag{3.39b}$$

Once the functional form of $\mu(X_i, \beta)$ is chosen, the maximization of (3.39a) or (3.39b) produces the maximum likelihood of the likelihood function:

$$\log_e L(\beta) = \sum_{i=1}^{n} Y_i \log_e[\mu(X_i, \beta)] - \sum_{i=1}^{n} \mu(X_i, \beta) - \sum_{i=1}^{n} \log_e(Y_i!) \tag{3.40}$$

Numerical search procedures are used to find the maximum likelihood estimates $b_0, b_1, ..., b_{p-1}$. Iteratively reweighted least squares can again be used to obtain these estimates. We also rely on standard statistical software packages specifically designed to handle Poisson regression to obtain the maximum likelihood estimates.

After the maximum likelihood estimates are been found, we can obtain the fitted response function and fitted values using Equation (3.41) and (3.42):

$$\hat{\mu} = \mu(X, b) \tag{3.41a}$$

$$\hat{\mu}_i = \mu(X_i, b) \tag{3.41b}$$

From the three functions in (3.19) to (3.21), the fitted response functions and fitted values are:

$$\mu = X'\beta: \qquad \hat{\mu} = X'b \qquad \hat{\mu}_i = X_i'b \qquad (3.42c)$$

$$\mu = \exp(X'\beta): \qquad \hat{\mu} = \exp(X'b) \qquad \hat{\mu}_i = \exp(X_i'b) \qquad (3.42d)$$

$$\mu = \log_e(X'b): \qquad \hat{\mu} = \log_e(X'b) \qquad \hat{\mu}_i = \log_e(X_i'b) \qquad (3.42e)$$

Model Development: Model development for a Poisson regression model is carried out in a similar fashion to that logistic regression, conducting tests for individual coefficients or group of coefficients based on the likelihood ratio test Statistic G^2 in (3.30) . For Poisson regression model (3.40), the model deviance is as follows:

$$DEV(X_1, X_2, \ldots, X_{p-1}) = -2\left[\sum_{i=1}^{n} Y_i \log_e\left(\frac{\hat{\mu}_i}{Y_i}\right) + \sum_{i=1}^{n}(Y_i - \hat{\mu}_i)\right] \qquad (3.43)$$

where $\hat{\mu}_i$ is the fitted value for the ith case according to (3.41b). The deviance residual for the ith case is:

$$dev_i = \pm\left[-2Y_i \log_e\left(\frac{\hat{\mu}_i}{Y_i}\right) - 2(Y_i - \hat{\mu}_i)\right]^{-\frac{1}{2}} \qquad (3.44)$$

The sign of the deviance residual is selected according to whether $Y_i - \hat{\mu}_i$ is positive or negative. Index plots of the deviance residuals and half-normal probability plots with simulated envelopes are useful for identifying outliers and checking the model fit. Note that if $Y_i = 0$, the term $\left[Y_i \log_e\left(\hat{\mu}_i / Y_i\right)\right]$ in (3.43) and (3.44) equals 0.

3.2 Model Specification

Hence, two class of models (logistic and Poisson regression) are defined as

1) Logistic regression model:

$$\pi = \beta_0 + X'\beta \qquad (3.45)$$

and

2) Poisson regression model:

$$\mu_i = \mu(X_i, \beta) = \exp(X_i'\beta) \qquad (3.46)$$

where

π is the probability of a success, X is vector of predictor variables and β is a vector of unknown coefficients associated with the predictors for Logistic regression model.

For Poisson regression model; $\mu(X_i, \beta)$ denotes the function that relates the mean response μ_i to X_i, the values of the predictor variables for case i, and β is the values of the regression coefficients.

However, the link function $g(\pi)$ can be expressed as

Logit: $$g(\pi) = \log\left(\frac{\pi}{1-\pi}\right) \qquad (3.47)$$

and the odds of success are

$$\frac{\pi}{1-\pi} = \exp(\beta_0 + \beta_1 X) \qquad (3.48)$$

Note that Equation (3.48) is the binary logistic regression model with one covariate or factor.

For multiple logistic regression model with more than one covariate, the probability event is

$$\pi = \frac{\exp\left(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p\right)}{1 + \exp\left(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p\right)} \quad (3.49)$$

for Logistic regression model, while

$$\mu_i = \mu(X_i, \beta) = \exp(X_i'\beta) = \exp\left(\beta_0 + X_1\beta_1 + \ldots + X_p\beta_p\right) \quad (3.50)$$

for Poisson regression model; we have β_0 = constant, β_i = coefficients, and $X_i = i^{th}$ predictors (or exponential of the i^{th} predictors).

To test whether several $\beta_k = 0$, or relate to the response variables, the following techniques are employed; Likelihood ratio test statistic G^2, Odd ratio, Wald test (z-test) and Model selection criteria: Akaile Information Criterion (AIC) and Bayesian Information Criterion (BIC):

3.3 Likelihood ratio test Statistic G^2

To test whether a subset of the X variables in a multiple logistic regression model can be dropped, that is, in testing whether the associated regression coefficients $\beta_k = 0$. The test procedure employed in this research is the general linear test procedure for Maximum likelihood estimation, the test is called the likelihood ratio test. It is based on comparison of full and reduced models. The full logistic model with response function:
For logistic regression model:

$$\pi = \left[1 + \exp\left(-X'\beta_F\right)\right]^{-1} \quad (3.51)$$

and Poisson regression model:

$$X'\beta_F = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_{p-1} X_{p-1} \quad (3.52)$$

3.4 Odd ratio

The Odd ratio is useful in interpreting the relationship between a predictor and response. The odds ratio (q) can be any nonnegative number. The odds ratio = 1 serves as the baseline for comparison. If q = 1, it indicates there is no association between the response and predictor. Also, if q > 1 the odds of success are higher for the reference level of the factor (or for higher levels of a continuous predictor). Then, if q < 1, the odds of success are less for the reference level of the factor (or for higher levels of a continuous predictor). Values farther from 1 represent stronger degrees of association.

Illustration: For the binary logistic regression model with one covariate or factor, the odds of success are:

$$\frac{\pi}{1-\pi} = \exp\left(\beta_0 + \beta_1 X\right) \quad (3.53)$$

Equation (3.53) is the binary logistic regression model with one covariate or factor. The exponential relationship provides an interpretation for β_0 : The odds increase multiplicatively by e^{β_1} for every one-unit increase in X. The odds ratio is equivalent to exp(β_1). For example, if β_1 is 0.75, the odd ratio is exp(0.75), which is 2.11. This indicates that there is 111% increase in the odds of success for every one unit increase in X.

3.5 Wald test (Z^* - test)

A large-sample test of a regression parameter can be constructed based on the hypotheses, such that

$$H_0 : \beta_k = 0 \quad against \quad H_0 : \beta_k \neq 0 \tag{3.56}$$

an appropriate test statistic is:

$$Z^* = \frac{b_k}{S\{b_k\}} \quad k = 0,1,\ldots, p \tag{3.55}$$

and the decision rule is:

$$If \ \left|Z^*\right| \leq Z_{\left(1 - \alpha/2\right)}, accept \ H_0, \ otherwise \ reject \ H_0$$

where Z is a standard normal random variable and $S\{b_k\}$ is the estimated approximate standard deviation of b_k obtained from Equation (3.51) and (3.52).

3.6 Criteria for Model Selection

The Model selection criteria considered in this research are (1) Akaike Information Criterions (AIC) and (2) Bayesian Information Criterions (BIC)

3.6.1 Akaike Information Criterions (AIC)

The general form for calculating AIC

$$AIC=-2\times In(Likelihood) + 2\times p \tag{3.56}$$

where

In is the natural logarithm
(Likelihood) is the value of the likelihood
P is the number of parameter in the model.

AIC can be calculated using residual sum of squares from regression (Henry, 2010):

$$AIC= n\times In(RSS/n) +2\times p \tag{3.57}$$

where

n is the number of data points (observations)
RSS is the residual sum of squares

AIC requires a bias- adjustment small sample sizes. If ratio of $\frac{n}{K} < 40$, then use bias – adjustment:

$$AIC_C = n\times In \text{ (likelihood)} +2p + \frac{(2p(p+1))}{(n-p-1)} \tag{3.58}$$

Note that the parameters are defined as above in Equation (3.56). Also, that as the size of the dataset, n, increases relative to the number of parameters, p; the bias-adjustment term on the right becomes very small. Therefore, it is recommended that we always use the small sample adjustment.

3.6.2 Bayesian Information Criterions (BIC)

The general form for calculating BIC

$$BIC_C = n\times In(likelihood)) + n\times In(n) + [Inn] \times p \tag{3.59}$$

Note: all parameters are defined as Equation (3.56); and the small values of BIC and AIC model will be chosen as the best model for the selected models (or selected as the suitable model among the selected models) [Schwarz, 1978].

4. Numerical Analysis and Results

This study focuses on the household utilized and non-utilized primary health care services in Choba, Obio/Akpor LGA of Rivers State. The data used for analysis were obtained through a constructed questionnaire (Appendix A). The extracted data from the

administered questionnaire comprises responses from four hundred households visited. The target population of this study was four hundred (400) household. However, four hundred and twenty (420) questionnaires were administered to households and based on the questionnaires that were returned, 400 were considered for this study. The statistical Software used are Microsoft Excel, SPSS 21 and Minitab 16.

In this study, the response (Y) denotes the probability distribution $P(Y_i = 1) = \pi_i$ or $P(Y_i = 0) = 1 - \pi_i$ [Response Variable (Y), representing if household utilized and non-utilizes primary health care services two or more times in the last one month] and the predictors (X) are described as follows:

AMO (X_1): Availability of Medical Officer
EDUC(X_2): Educational Years
PRIMED(X_3): Average price of medication (naira)
DISTANCE(X_4): Average driving distance (mins)
RY(X_5): Average Monthly Income of respondents (naira)

From any set of p predictors, we have 2^p alterative models can be constructed. It is based on the fact that each predictor can either be excluded or included from the model (Christensen, 1997). Therefore, we have p=5, then 2^5 =32 different possible subset models that can be formed from the pool of five variables (X), such that $Y_i = \beta_0 + e_i$. That is there are regression models with five variables (X_1, X_2, X_3, X_4 , X_5), with two variables [X_1 and X_2, X_1 and X_3, X_1 and X_4, X_1 and X_5, X_2 and X_3, X_2 and X_4, X_2 and X_5, X_3 and X_4, X_3 and X_5, then X_4 and X_5], and so on. To choose the best model from the selected models, we used the two common Criteria for Model Selection developed to compare the selected suitable models.

4.1 Hypothesis

H_0: The model adequately describes the data

against

H_1: The model does not adequately describe the data

4.2 The Procedure for Data Analysis

We used the goodness-of-fit tests, wald test (Z^* -test) and if p-value is less than accepted α -level, the test would reject the null hypothesis of an adequate fit. Our interest is to investigate the use of (if household utilized and non-utilized) primary health care services two or more times in the last one month as follows:

1. The availability of medical officer and educational Years (or level) upon the response variable (Model A).
2. Availability of medical officer and the Average price of medication upon the response (Model B).
3. Availability of medical officer and average driving distance upon the response (Model C).
4. Availability of medical officer and average real income (or monthly income of respondents) upon the response (Model D).
5. Availability of medical officer, education years and average driving distance upon the response (Model E).
6. Availability of medical officer, Education years, average monthly income of respondents, average driving distance and average real income (or monthly income of respondents) upon the response (Model F).

we compared this models to identify which model is suitable between logistic and

Poisson regression models, using the household utilized/ non utilizes primary health care services in Choba, Obio/Akpor LGA of Rivers State data obtained. Furthermore, we test whether $\beta_k = 0$, or relate to the response variable, using the following techniques; Likelihood ratio test Statistic G^2**,** Odd ratio, Wald test (z-test) and two model selection criteria: Akaike Information criterion (AIC) and Bayesian Information criterion and (BIC) discussed in Chapter three, Section 3.4 to 3.6.

4.3 Results

This section is divided into four parts; 1) Descriptive statistics of the respondents profile in the questionnaire; 2) Logistic Regression Model; 3) Poisson Regression Model; and 4) Comparison of the estimated models parameters and identification of the optimal model using the two criterions considered for models selection (i.e. Logistic and Poisson regression Models).

4.4 Descriptive statistics

Table 4.1: Summary of Number of Household size and Sex of the Respondents who utilized/non utilized the primary health care services

Household size	Frequency	Percent
2	79	19.75
3	78	19.50
4	93	23.25
5	79	19.75
6	71	17.75
Sex	Frequency	Percent
F	383	95.75
M	17	4.25

Table 4.1 illustrates that household with 6 members has the minimum percentage with the value of 18%; while the household with 4 members has the maximum value of (23%), indicating that 23% of the respondents' families have 2 children alongside the husband and wife. It is further seen from the Table 4.1 that, 96% (383) of the respondents were females and 4% (17) were males. Suggesting that majority of respondents who visit hospitals are females.

4.5 Logistic Regression

In Section 4.2; we considered six models to be built. The log-likelihood function (maximum likelihood estimators) of the Logistic regression is used to estimate the parameters denoted as $\beta_0, \beta_1, ..., \beta_p$ as described in Section 3.2 Equation (3.49).

4.5.1 Fitted Logistic Regression for Model A: Y versus A.M.O. and Education

When p=2, X_1=A.M.O, X_2 =Education

Table 4.2: Estimated Coefficients, p-values and odds ratios for Model A

Predictor	**Coefficients**	**P-values**	**Odds Ratio**
Constant	1.32609	0.000	
A.M.O.	0.96622	0.000	2.63
Education Years	-0.08238	0.005	0.92

4.5.2 Fitted Logistic Regression for Model B: Y versus A.M.O, Primed

Similarly, the log-likelihood function (maximum likelihood estimators) of the Logistic regression model is obtained as above.

When p=2, X_1=A.M.O, X_3=Primed.

Table 4.3: Estimated Coefficients, p-values and odds ratios Model B

Predictor	Coefficients	P-values	Odds Ratio
Constant	0.692676	0.001	
A.M.O.	0.901548	0.000	2.46
Primed	-0.0001587	0.219	1.00

4.5.3 Fitted Logistic Regression Model C: Y versus A.M.O., Distance (mins)

When p=2, X_1 = A.M.O, X_4 = Distance (mins)

Table 4.4: Estimated Coefficients, p-values and odds ratios Model C

Predictor	Coefficients	P-values	Odds Ratio
Constant	0.813482	0.002	
A.M.O.	0.901657	0.000	2.46
Distance (mins)	-0.0076804	0.194	0.99

4.5.4 Fitted Logistic Regression Model D: Y versus A.M.O., RY (N0.000)

When p=2, X_1=A.M.O, X_5 = RY (N0.000)

Table 4.5: Estimated Coefficients, p-values and odds ratios Model D

Predictor	Coefficients	P-values	Odds Ratio
Constant	0.622565	0.013	
A.M.O.	0.898790	0.000	2.46
RY (N0.000)	-0.0000040	0.665	1.00

4.5.5 Fitted Logistic Regression Model E: Y versus A.M.O., Education, Distance (mins)

When p=3 X_1=A.M.O, X_2=Education, x_4=Distance (mins)

Table 4.6: Estimated Coefficients, p-values and odds ratios Model E

Predictor	Coefficients	P-values	Odds Ratio
Constant	1.54634	0.000	
A.M.O.	0.969482	0.000	2.64
Education	-0.0803738	0.007	0.92
Distance (mins)	-0.0067063	0.263	0.99

4.5.6 Fitted Logistic Regression Model F; Y versus A.M.O., Education, Primed, Distance (mins), RY

When p=5, X_1=A.M.O, X_2=Education, X_3=Primed, x_4=Distance, X_5=RY

Table 4.7: Estimated Coefficients, p-values and odds ratios Model F

Predictor	Coefficients	P-values	Odds Ratio
Constant	1.72150	0.000	
A.M.O.	0.972932	0.000	2.65
Education	-0.0811425	0.006	0.92
Primed	-0.0001570	0.246	1.00
Distance (mins)	-0.0064727	0.280	0.99
RY (N0.000)	-0.0000011	0.911	1.00

The fitted Logistic response function and the fitted values (estimates for the Model A) in Table 4.2 can be expressed as;

$$\hat{\pi}_i = \frac{\exp(1.326 + 0.966x_{i1} - 0.0823x_{i2})}{1 + \exp(1.326 + 0.966x_{i1} - 0.0823x_{i2})} \quad (4.1)$$

From Table 4.2, it is seen that both availability of Medical Officers (A.M.O) and the level of Education played major roles on the respondents' utilization of the primary healthcare services, since both estimated parameters has significant effect. While the odds ratio suggests that in a facility where there are medical officers, there is the likelihood of having at least 3 patients at any time with one being educated.

Similarly, the fitted Logistic response function and the fitted values (estimates for the Model B to F) in Table 4.3 to 4.7 are expressed as;

$$\hat{\pi}_i = \frac{\exp(0.693 + 0.902\ x_{i1} - 0.00016\ x_{i3})}{1 + \exp(0.693 + 0.902\ x_{i1} - 0.00016\ x_{i3})} \quad (4.2)$$

$$\hat{\pi}_i = \frac{\exp(0.813 + 0.902x_{i1} - 00768x_{i2})}{1 + \exp(0.813 + 0.902x_{i1} - 00768x_{i2})} \quad (4.3)$$

$$\hat{\pi}_i = \frac{\exp(0.623 + 0.899x_{i1} - 0.00004x_{i5})}{1 + \exp(0.623 + 0.899x_{i1} - 0.00004x_{i5})} \quad (4.4)$$

$$\hat{\pi}_i = \frac{\exp(1.546 + 0.969x_{i1} - 0.0804x_{i2} - 0.00671x_{i4})}{1 + \exp(1.546 + 0.969x_{i1} - 0.0804x_{i2} - 0.00671x_{i4})} \quad (4.5)$$

$$\hat{\pi}_i = \frac{\exp(1.722 + 0.973x_{i1} - 0.0811x_{i2} - 0.000157x_{i3} - 0.00647x_{i4} - 0.000001x_{i5})}{1 + \exp(1.722 + 0.973x_{i1} - 0.0811x_{i2} - 0.000157x_{i3} - 0.00647x_{i4} - 0.000001xi5)} \quad (4.6)$$

In Table 4.3, it is observed that availability of Medical Officers (A.M.O) played major role on the respondents' utilization of the primary healthcare services, with p-value < 0.05 and has a significant effect. However, the average price of medication including transport cost (Primed), does not have effect on the healthcare service utilization and p-value >0.05 (0.22).

In Table 4.4, availability of Medical Officers (A.M.O) determines the respondents' utilization of the primary healthcare services, with p-value (0.000) < 0.05 and hence, has a significant effect, while the Average Distance (mins), does not have an effect on the services with p-value of (0.194) > 0.05

In Table 4.5, availability of Medical Officers (A.M.O) determines the respondents' utilization of the primary healthcare services, with p-value (0.000) < 0.05 while the Average income RY (N0.000), does not have effect on the healthcare services with p-value of (0.665) > 0.05.

Table 4.6 shows availability of Medical Officers (A.M.O) and Level of Education determines the utilization of the primary healthcare services, with p-value < 0.05 (0.000 and 0.007 respectively) while the distance does not have effect with p-value of (0.263) > 0.05. It implies that availability of Medical Officers and level of education are the major determining factors on the use of primary healthcare services. While the odds ratio suggests that in a facility where there are Medical Officers, there is the likelihood of having at least 3 patients at any time with one being educated and one because of the distance (min) who might not visit the facility.

Table 4.7 shows that availability of Medical Officers (A.M.O) and Level of education determines the utilization of services with p-values < 0.05 (0.000 and 0.006 respectively), while distance, primed and RY have no effect with p-values > 0.05. It suggests that availability of Medical Officers and level of education are the major determining factors on the use of primary healthcare services.

- The selected six models summary for Logistic regression using the two criterion

Table 4.8: The rank of AIC and BIC values of the Six Logistics models selected

Model	Predictors	Rank	AIC	BIC
A	**X_1 , X_2 (A.M.O, Education)**	**1**	**447.313**	**459.287**
B	X_1 , X_3 (A.M.O, Primed)	6	453.681	465.655
C	X_1 , X_4 (A.M.O, Distance)	4	453.510	465.485
D	X_1 , X_5, (A.M.O, RY)	5	454.993	466.967
E	X_1 , X_2, X_4 (A.M.O, Education, Distance)	2	448.071	464.037
F	X_1 , X_2, X_3 X_4, X_5 (A.M.O, Education, Primed, Distance, RY)	3	450.566	474.515

In Table 4.8, the AIC and BIC with the least values is Model A, therefore Model A is the best model using the two models selection criterion considered and can be expressed as Equation 4.1; *where*, $\hat{\pi}_i$ = Y is response variable (if household utilized primary health care services or not) and (X_1=A.M.O, X_2= Education).

4.6 Poisson Regression

Considering the Poisson regression, the log-likelihood function for Poisson regression to estimate the maximum likelihood estimators denoted as $\beta_0, \beta_1, ..., \beta_p$ and it is obtained as described in Equations 3.50.

4.6.1 Fitted Poisson Regression Model A; Y versus A.M.O., Education

When p=2 (X_1=A.M.O, X_2=Education),

Table 4.9: Estimated Coefficients, Wald Chi-Square and Significance Values Model A

Predictors	Coefficients	Wald Chi-Square	Sig.
(Intercept)	-.260	2.406	.121
Education	-.022	2.024	.155
AMO	.259	4.547	.033

4.6.2 Fitted Poisson Regression Model B, Y versus A.M.O., Primed
When p=2 (X_1=A.M.O, X_3=Primed)

Table 4.10: Estimated Coefficients, Wald Chi-Square and Significance Values Model B

Predictors	Coefficients	Wald Chi-Square	Sig.
(Intercept)	-.420	13.464	.000
AMO	.246	4.136	.042
Primed	-4.190E-005	.387	.534

4.6.3 Fitted Poisson Regression Model C; Y versus A.M.O., Distance (mins),
When p=2, X_1=A.M.O, X_4=Distance (mins)

Table 4.11: Estimated Coefficients, Wald Chi-Square and Significance Values Model C

Predictors	Coefficients	Wald Chi-Square	Sig.
(Intercept)	-.388	7.241	.007
AMO	.246	4.131	.042
Distance(min)	-.002	.430	.512

4.6.4 Fitted Poisson Regression Model D; Y versus A.M.O., RY (0.00),
When p=2 X_1=A.M.O, X_5=RY (0.00),

Table 4.12: Estimated Coefficients, Wald Chi-Square and Significance Values Model D

Predictors	Coefficients	Wald Chi-Square	Sig.
(Intercept)	-.439	10.315	.001
AMO	.246	4.139	.042
RY	-1.010E-006	.046	.831

4.6.5 Fitted Poisson Regression Model E; Y versus A.M.O, Education, Distance when p=3, X_1=A.M.O, X_2=Education, X_4=Distance (mins)

Table 4.13: Estimated Coefficients, Wald Chi-Square and Significance Values Model E

Predictors	Coefficients	Wald Chi-Square	Sig.
(Intercept)	-.206	1.128	.288
AMO	.258	4.528	.033
Education	-.021	1.896	.169
Distance (mins)	-.002	.302	.582

4.6.6 Fitted Poisson Regression Model F; Y versus A.M.O., Education, Primed, Distance, RY

When p=5, X_1=A.M.O, X_2=Education, X_3= Primed, X_4=Distance (mins), X_5=RY

Table 4.14: Estimated Coefficients, Wald Chi-Square and Significance Values Model F

Predictors	Coefficients	Wald Chi-Square	Sig.
(Intercept)	-.164	.564	.453
AMO	.259	4.546	.033
Education	-.021	1.925	.165
Primed	-4.151E-005	.357	.550
Distance(Min)	-.002	.276	.599
RY	-2.051E-007	.002	.966

The fitted Poisson response function and the fitted values (estimates for the Model A to F) in Table 4.9 to 4.14 are expressed as;

$$\hat{\mu}_i = \exp(-0.260 - 0.022x_{i1} + 0.259x_{i2}) \quad (4.7)$$

$$\hat{\mu}_i = \exp(-0.420 + 0.246x_{i1} - 4.19\times10^{-5}x_{i3}) \quad (4.8)$$

$$\hat{\mu}_i = \exp(-0.388 + 0.246x_{i1} - 0.002x_{i4}) \quad (4.9)$$

$$\hat{\mu}_i = \exp(-0.439 + 0.246x_{i1} - 1.01\times10^{-6}x_{i5}) \quad (4.10)$$

$$\hat{\mu}_i = \exp(-0.206 + 0.258x_{i1} - 0.021x_{i2} - 0.002x_{i4}) \quad (4.11)$$

$$\hat{\mu}_i = \exp(-0.164 + 0.259x_{i1} - 0.021x_{i2} - 4.15\times10^{-5}x_{i3} - 0.002x_{i4} - 2.05\times10^{-7}x_{i5}) \quad (4.13)$$

From Table 4.9, the Poisson regression done shows only availability of Medical Officers (A.M.O) has mean effect on the utilization of primary healthcare facility with significance value of 0.033 < 0.05; while education level was seen not to have any effect, hence, insignificant.

In Table 4.10, it was discovered that only availability of Medical Officers (A.M.O) has mean effect on the utilization of primary healthcare facility with significance p-value of 0.042< 0.05; while the Average price of medication including transport cost (Primed), was seen not to have any effect, hence, insignificant.

In Table 4.11, it was also discovered that only availability of Medical Officers (A.M.O) that has mean effect on the utilization of primary healthcare facility with significant value of 0.042< 0.05; while, Distance was seen to have no effect, hence, it is said not be significant.

Table 4.12 also shows that only availability of Medical Officers (A.M.O) has mean effect on the utilization of primary healthcare facility with significance p-value of 0.042< 0.05; while, RY (the Average income (N0.000)), had no effect, hence, insignificant.

In Table 4.13, it was also discovered that only availability of Medical Officers (A.M.O) that has mean effect on the utilization of primary healthcare facility with its significance p-value of 0.033< 0.05; while, Distance and Education were seen to have no effect, hence, insignificant.

Finally in Table 4.14, availability of Medical Officers (A.M.O) only determines the respondents' utilization of the primary healthcare facilities, with p-value < 0.05 (p=0.033). Hence, AMO have a significant effects, while the others has no effect on the healthcare facilities utilization, since their p-values are greater than 0.05.

- **The selected six models summary for Poisson regression using the two criterion**

Table 4.15: The rank of AIC and BIC values of the Six Poisson models selected

Model	Predictors	Rank	AIC	BIC
A	**X_1 , X_2 (A.M.O, Education)**	**1**	**768.157**	**780.131**
B	X_1 , X_3 (A.M.O, Primed)	3	769.814	781.788
C	X_1 , X_4 (A.M.O, Distance)	2	769.769	781.743
D	X_1 , X_5, (A.M.O, RY)	5	770.159	782.219
E	X_1 , X_2, X_4 (A.M.O, Education, Distance)	4	769.851	785.817
F	X_1 , X_2, X_3 X_4, X_5 (A.M.O, Education, Primed, Distance, RY)	6	773.452	797.401

From Table 4.15, the best model is Model A, using the AIC_P and BIC_P criterion and can expressed as Equation 4.7; where, $\hat{\mu}_i$ = Y is response variable (if household utilized primary health care services or not) and (X_1=A.M.O, X_2=Education).

Table 4.16: Summary of the Coefficients and p-values of Six Logistics and Poisson Regression Models

S/No.	Model	Model Variables	Estimated Coefficients (P-values at 5%)		Level of Significant
			Logistic Regression Model (LM)	**Poisson Regression Model (PM)**	
1	**A**	**Constant X_1=A.M.O, X_2=Edu. (Y , X_1, X_2)**	$\beta_0 = 1.326$ (0.000) $\beta_1 = 0.966$ (0.000) $\beta_2 = -0.082$ (0.005)	$\beta_0 = -0.26$ (0.121)** $\beta_1 = 0.259$ (0.033) $\beta_2 = -0.022$ (0.155)**	**β_0 β_1 β_2 are sig. for (LM) while only β_1 is sig. for (PM)**
2	B	Constant X_1=A.M.O, X_3=Primed (Y, X_1, X_3)	$\beta_0 = 0.693$ (0.001) $\beta_1 = 0.902$ (0.000) $\beta_3 = -0.00016$ (0.219)**	$\beta_0 = -0.420$ (0.00) $\beta_1 = 0.246$ (0.042)** $\beta_3 = -4.19\times10^{-5}$ (0.534)**	β_0 β_1 are sig. for both
3	C	Constant X_1=A.M.O, X_4=Distance (Y, X_1,X_3)	$\beta_0 = 0.813$ (0.002) $\beta_1 = 0.702$ (0.000) $\beta_4 = -0.0076$ (0.194)**	$\beta_0 = -0.388$ (0.007) $\beta_1 = 0.246$ (0.042)** $\beta_4 = -0.002$ (0.512)**	β_0 β_1 are sig. for both
4	D	Constant X_1=A.M.O, X_5=RY (Y, X_1,X_4)	$\beta_0 = 0.623$ (0.013) $\beta_1 = 0.899$ (0.000) $\beta_4 = -4.01\times10^{-6}$ (0.665)**	$\beta_0 = -0.439$ (0.001) $\beta_1 = 0.246$ (0.042) $\beta_4 = -1.01\times10^{-6}$ (0.831)**	β_0 β_1 are sig. for both
5	E	Constant X_1=A.M.O, X_2=Edu. X_4=Distance (Y, X_1, X_2, X_3)	$\beta_0 = 1.546$ (0.000) $\beta_1 = 0.969$ (0.000) $\beta_2 = -0.080$ (0.007) $\beta_4 = -0.0067$ (0.263)**	$\beta_0 = -0.206$ (0.288)** $\beta_1 = 0.258$ (0.033) $\beta_2 = -0.021$ (0.169)** $\beta_4 = -0.002$ (0.582)**	β_0 β_1 β_2 are sig. for (LM) while only β_1 is sig. for (PM)

6	F	Constant X_1=A.M.O, X_2=Edu. X_3=Primed X_4=Distance X_5=RY (Y, X_1, X_2, X_3, X_4, X_5)	$\beta_0 = 1.722$ (0.000) $\beta_1 = 0.973$ (0.000) $\beta_2 = 0.0297$ (0.006) $\beta_3 = -1.57\times10^{-4}$ (0.246) ** $\beta_4 = -0.0065$ (0.280) ** $\beta_5 = -1.1\times10^{-6}$ (0.911) **	$\beta_0 = -0.164$ (0.453) ** $\beta_1 = 0.259$ (0.033) $\beta_2 = -0.021$ (0.165) ** $\beta_3 = -4.15\times10^{-5}$ (0.550) ** $\beta_4 = -4.01\times10^{-6}$ (0.665) ** $\beta_5 = -2.05\times10^{-7}$ (0.966) **	β_0 β_1 β_2 are sig. for (LM) while only β_1 is sig. for (PM)

Footnote: ** p-values greater than the appropriate critical value (0.05) is not significant and the bold Model is the optimal model identified.

Table 4.17: Comparison of the rank of AIC and BIC of the Six Logistics and Poisson regression models

Rank	**Model: Predictors**	**Logistic regression**			**Poisson regression**	
		AIC	**BIC**	**Predictors**	**AIC**	**BIC**
1	**Model A:** Y, X_1 , X_2 (A.M.O, Education)	**447.313**	**459.287**	**Model A:** X_1 , X_2 (A.M.O, Education)	**768.157**	**780.131**
2	**Model E:** Y, X_1 , X_2, X_4 (A.M.O, Education, Distance)	448.071	464.037	**Model C:** X_1 , X_4 (A.M.O, Distance)	769.769	781.743
3	**Model F:** Y, X_1 , X_2, X_3 X_4, X_5 (A.M.O, Education, Primed, Distance, RY)	450.566	474.515	**Model B:** X_1 , X_3 (A.M.O, Primed)	769.814	781.788
4	**Model C:** Y, X_1 , X_4 (A.M.O, Distance)	453.510	465.485	**Model E:** X_1 , X_2, X_4 (A.M.O, Education, Distance)	769.851	785.817
5	**Model D:** Y, X_1 , X_5, (A.M.O, RY)	454.993	466.967	**Model D:** X_1 , X_5, (A.M.O, RY)	770.159	782.219
6	**Model B:** Y, X_1 , X_3 (A.M.O, Primed)	456.680	465655	**Model F:** X_1 , X_2, X_3 X_4, X_5 (A.M.O, Education, Primed, Distance, RY)	773.452	797.401

Footnote: the bold Model is the optimal model identified.

The Logistic regression AIC and BIC values are 447.331 and 459.287 respectively for the Model A (i.e. A.M.O and Education) and it is the lowest of all the models selected suggesting that Model A is the best model. Also, the Poisson regression AIC and BIC values are 768.157 and 780.131 respectively for the Model A (i.e. A.M.O and Education) and it's the lowest of the entire model suggesting that Model A is the best model.

4.8 Result Discussion

From Table 4.16 and Table 4.17, both regression techniques identified Model A (i.e. A.M.O and Education) as the best model for predictive of the response variable (if household utilized primary health care services or not).

In Table 4.17, both the AIC and BIC values of Model A had the lowest of the entire model suggested in the Logistic and Poisson regression models done.

Table 4.16 shows the coefficients and p-values of Six Logistic and Poisson regression models selected. The best Logistic and Poisson regression model identified is Model A, where the explanatory variables are A.M.O and Education.

The p-values of A.M.O and Education are 0.000 and 0.005 respectively for the Logistic regression Model A; shows the coefficients are significant at 5%. However, the Poisson regression Model A; shows that the A.M.O variable is the only independent variable that is significant with p-value of 0.033. Therefore, the best Logistic regression model identified is when the explanatory variables are A.M.O and Education. However, the Poisson regression model also identified the same model as the best model, but the A.M.O is the only significant independent variable. In addition, since the AIC and BIC values of Logistic regression is smaller than that of Poisson regression, it suggests that Logistic regression model is the best fit in modeling count data.

5. Conclusion

The descriptive statistics shows that a household with 6 members has the minimum percentage with the value of 18% while a household with 4 members has the maximum value of 23%. The best logistic regression model identified is when the explanatory variables are A.M.O and Education (since, there are coefficients that are significant at 5% with p-value 0.000 and 0.005 respectively. However, the Poisson regression model also identified the same model as the best model, but the A.M.O is the only significant independent variable with p-value 0.033.

The AIC and BIC for both Logistic and Poisson regression identified the model to be the best, when predictors are A.M.O and Education. Conclusively, the AIC and BIC of Logistic is smaller than Poisson regression model, suggesting that Logistic regression model is the best fit in modeling binary response variable count data [or the result showed that the Logistic regression model is the best fit in modeling binary response variable in form of a count data; based on the two assessment criteria employed Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)]. Therefore, the findings are

1) Was able to identify a suitable Logistic and Poisson regression model from the count data observation where the response is binary.
2) The performance of the Logistic and Poisson regression model estimated coefficients have been compared, using AIC and BIC for Criterions.
3) The research identified Logistic regression model to be more suitable for the observation considered.

5.1 Recommendation

This research recommended Logistic regression model in analyzing count data, with

binary response based on the findings of the research. This paper identified Logistic regression model to be more suitable where the response is binary.

6. References

Afifi, W. A. Dillow, M. R. and Morse, C. (2004). *Examining Predictors and Consequences of Information seeking in Close Relationships*, Personal relationships, 11, 429-449.

Agresti, A. (2007), *An Introduction to Categorical Data Analysis*, Second Edition, Wiley, Inc., New York.

Arlitt, M.F. and Williamson, C.L., (1997), "*Internet Web Servers Work Load Characterization and Performance Implications",* IEEE/ACM Transactions on Networking. 5(5):631.

Armstrong, J. S. (2012), *Illusions in Regression Analysis*, International Journal of forecasting. 28 (3), 689.

Berk, R. A. (2003), *Regression Analysis: A Constructive Critique*, CA: Sage Publications, Newbury Park.

Berk, R. and MacDonald, J. (2008), "*Overdispersion and Poisson Regression*" *Journals of* Quantitative Criminology 24(3), 269 – 284.

Cannizzaro, F, Greco, G.; Rizzo, S. Sinagra, E. (1978) ,"*Result of the Measurement carried in order to Verify the Validity of the Poisson-Exponential Distribution in Radioactive Decay Event*". The international Journal of Applied Radiation and Isotopes 29(11), 649.

Christensen, R. (1997), *Log-linear Models and Logistic Regression,* Second edition. Springer-Verlag, New York.

Cox, D. R. and Snell, E. J. (1989), *Analysis of Binary Data* (second edition), Chapman & Hall/CRC.

Devita, V. T., Hellman, S. L. and Rosenberg's, S. A. (2008), *Principles and Practice of Oncology*, Volume 2.

Freeman, D. H. and Holford, T. R. (1980), *Summary Rates*, Biometrics 36, 195 – 205.

Frome, E. L. and Beauchamp, J. J. (1968), *Maximum Likelihood Estimation of Survival Curve Parameters,* Biometrics 24, 595 – 605.

Frome, E. L. and DuFrain, R. R, (1982), *Analysis of Cytogenesis Close- Response Data Using a Model Derived from the Theory of Dual Radiation Action*, Abstract. Biometrics 38, 11-17.

Greene, W. H. (2003), *Econometric Analysis;* New York University, Upper Saddle River, New Jersey, Fifth Edition.

Grimms, L. G. and Yarnold, P. R. (1995), *Reading and Understanding Multivariate Statistics*, American Psychological Association, Washington, D.C.

Grizzle, J. E., Starmer, C. F. and Koch, G. G. (1969), *Analysis of Categorical data by Linear Models,* Biometrics, 25, 489-504.

Hauck, W. W, and Donner, A., (1997), *Wald's test as Applied to Hypothesis in Logit Analysis.* Journal of the American Statistical Association, 72,851 – 853.

Henry, G. A. (2010). *Comparison of Akaike Information Criterion (AIC) and Bayesian Criterion (BIC) in Selection of an Asymmetric Price Relationship*. Journal of Development and Agricultural Economics, Vol 2 (1), page 001-006.

Hosmer, D. W., and Lemeshow, S. (2000), *Applied Logistic Regression,* Second Edition, John Wiley and Sons Inc., New York.

Ingeles, C. J.; Garcia-Fernandez, J. M. Castejon, J. L.; Valle Antonio, B. D. and Marzo, J. C (2009), *Reliability and Validity Evidence of Score on the Achievement Goal Tendencies:* Questionnaire in a sample of Spanish students of compulsory secondary education, Psychology in the school, Vol. 46. 1048 – 1060, Wiley Periodicals, Inc;

A Wiley company.

Jennings, D. E. (1986), *Judging Inference Adequacy in Logistic Regression,* Journal of the American statistical Association, 81,471 – 476.

King, G. (1989), *Variance Specification in Event Count Models from Restrictive Assumptions to a Generalized Estimator*, American Journal of Political Science. 762 – 784.

Kleinbaum, D. G. (1994), *Logistic Regression*, A self-learning text. Springer - Verlag, New York, 104 – 119.

McCaullagh, H. and Nelder, J. N., (1992), *Generalized linear Models* (2nd Edition). Chapman and Hall, Madras.

Michael, H. K.; Christopher, J. N., John N., and William. L (2005). *"Applied Linear Statistical Model",* fifth Ed.; 555-623., McGraw Hill International, New York.

Nduka, E.C., (1999), *Principles of Applied Statistics 1*, Crystal Publishers, Okigwe.

Nixon D.C., (1991), *Event count Models for Supreme Court dissents*, Political Methodology 4, 11-14.

Osgood, W. (2000), *"Poisson-based Regression Analysis of Aggregate crime Rates".* Journal of Quantitative Criminology 16, 21 – 43.

Paternoster R, and Brame R. (1997), *"Multiple routes to delinquency: A test of developmental and general theories of crime*", Journal of Criminology 35, 45-84.

Pregibon, D., (1981), *Logistic Regression Diagnostics* Analysis of Statistics 9, 705-724.

Sampson, R. J. and Laub, J. H. (1997), *A Life-Course Theory of Cumulative Disadvantage and the Stability of Delinquency*, Pp. 133-161 in *Developmental Theories of Crime and Delinquency*. (*Advances in Criminological Theory*, Volume 7), edited by Terence P. Thornberry. New Brunswick, NJ: Transaction

Schwarz, G. E. (1978). "*Estimating the dimension of a model"*, Annals of Statistics, 6 (2): 461–464.

Yanqiu, G., Ronsmans, C. and Lin, A. (2009), *Time Trends and Regional Differences in Maternal Mortality in China from 2000 to 2005*, Peking University Health Science Center, Beijing, China, Bull World Health Organ 87:913–920.

APPENDIX A
QUESTIONNAIRE
Title: ASSESSING LOGISTIC AND POISSON REGRESSION MODEL IN ANALYZING COUNT DATA
[In Determinants of primary healthcare Service Variables in Choba, Rivers State (Obio/Akpor Local Government Area)]
Instruction: Please tick ☐ and fill in the appropriate answer where necessary

Part A

1. Sex: Male ☐ Female ☐
2. Household size (Number of individual) ……

Part B

3 If your household utilizes primary healthcare service two or more times in last one month. Yes ☐ No ☐
4. What is your average monthly income …….. (Naira)
5. What is the total price of medication including transportation cost in Naira paid by household per visit to a primary healthcare centre…….. (Naira).
6. What is the distance to the nearest primary healthcare centre…… (in Minutes.)
7. What is your educational qualification
   Primary Education ☐ Secondary Education ☐ Tertiary Education ☐
8. Availability of medical officer in the nearest primary healthcare centre.
   Yes ☐ No ☐

Definition of Variables

y = 1, if household utilized primary health care services two or more times in the last one month.
y = 0, if household does not utilize primary health care services.
Sex: Female=F, Male=M
Houssehold Size: Number of Individuals
Ry = Average monthly income of respondents (naira)
PRIMED = Average price of medication including transport cost in naira paid by household per visit to a primary health care centre.

Distance (Mins) = Average driving distance to a nearest primary health centre in minutes.
EDUC = Level of education attained.
6 = Completed primary education
12 = Completed secondary education
16 = Completed tertiary education

HSIZE = Average Household Size
AMO = Indicator variable for availability of medical officer in the nearest primary health care centre.
AMO = O if no medical doctor is available in the primary health centre nearest to the household.
AMO = 1 if one or more medical doctors is available in the primary health centre nearest to the household.
Source: Researcher's Field work, 2016.