GENERALIZED LINEAR MODELS ON ROAD TRAFFIC CRASH ALONG KEFFI-LAFIA ROAD FROM 2006 TO 2015

Supported by tetfund TERTIARY EDUCATION TRUST FUND

M. O. Adenomon* and I. Ayuba
Department of Statistics, Nasarawa State University, PMB 1022, Keffi, Nigeria
*Corresponding author: admonsagie@gmail.com

Received: February 12, 2018 **Accepted:** September 18, 2018

Abstract: Poisson and Negative Regression analyses are techniques used to model dependent variable that describe count data. The aim of this paper is to study the relationship between death from road crashes and, total cases of road crashes and time on Keffi-Lafia road. To achieve this, secondary data was obtained from RS 4.3 sector command, Nasarawa State on annual road traffic crash for the period of 2006 to 2015 on Keffi-Lafia road. The results from Poisson and Negative Binomial regression models revealed a positive significant relationship between total cases and Death from road crashes from 2006-2015 on Keffi-Lafia road. In addition, there is a significant negative trend of about 7.5% annual in the deaths from road crashes from 2006- 2015 on Keffi-Lafia road and there was no evidence of over dispersion, the Poisson regression model was more robust than the negative binomial regression model. This paper suggest that public enlightenment and sensitization of road users should be intensify by FRSC, as this can further reduce the rate of deaths from accident from 7.5% to 50% on Keffi-Lafia road.
Keywords: Regression, Poisson, Negative Binomial, Road Crash, Count Data

Introduction

Road traffic crashes in Nigeria have been identified as one of the major causes of deaths in the country. It is classified as the second major cause of death in the country following malaria (FRSC, 2009). According to the Statistics on road accidents in Nigeria for the year 2009 disseminated by Federal road safety commission, department of policy, research and statistics, shows that there were 34,632 road traffic crashes for the year 2009 (FRSC, 2009). There were total of 66,924 casualties with 19,188 of them losing their lives, while 47,736 sustained serious injuries. This reveals that there was an average of 53 deaths every day in Nigeria which was caused by road traffic crashes. The most dangerous part of it all is that most of the people who are killed by road traffic crashes are those in the age group that constitute the work force of this nation. It is in this regard that more attention needs to be placed on the studies on road traffic crash and its impact on human lives and properties in Nigeria.

As a result of these tremendous effects of Road traffic crashes on human lives, properties and the environment, many researches have come out with the causes, effects and recommendations to road traffic crashes in Nigeria and the world. These causes include drinking of alcohol, machine failure and over speeding (Ukwu and Adedokun, 2016). Yet every year the Federal road safety commission, National Bureau of Statistics and other organizations would report an increase in Road traffic crashes (NRTR, 2004). Hence the need to analyze the Road Traffic Crashes data statistically to check whether there is any evidence of increasing road traffic crashes as years go by resulting to large number of people losing their lives.

Researchers have been modelling vehicular crashes with different models in various parts of the world. However, it is extremely difficult to just apply models which have worked somewhere to data obtained from different country due to the variations in the various factors pertaining in different countries (Fletcher *et al*., 2006). There has not been much statistical research in the field of road traffic crash in Nigeria. This might have been as a result of the inadequate information available on road traffic crashes and its impact on human live and properties in the country. These road traffic crashes have killed a lot of people in this country and as such it is described as one of the major causes of death in Nigeria. The causes of death of casualties in road traffic crashes have been associated with secondary collision, improper handling of casualties and inadequate emergency services in the country. In the publication of the policy, research and statistics of FRSC, that there are certain factors which could be possible contributors of the death of casualties in road traffic crashes in Nigeria. These factors include location of the road traffic crash, the age and sex of the casualty, the type of vehicle involved in the road traffic crash, nature of the road, weather conditions, the days of the week, the time of the day, and the ownership of the vehicle involved in the road traffic crash.

The purpose of this study is to estimate the trend of deaths resulting from road traffic crash along Keffi-Lafia road and to investigate the relationship between cases of road crashes and death resulting from road crashes using Poisson and negative binomial multiple regression models.

The following are Review of Road Crashes in Nigeria: Oyedepo and Makinde (2010) examined accident data on the 52 km Akure-Ondo carriage way and spot speed data between 2002 and 2007. Results from multiple regression analysis revealed a correlation coefficient (R) value of 70.70% and the coefficient of determination (R^2) of 49.70%, respectively. Kareem *et at*. (2012) study data on casualties of the road accident in Nigeria using various linear and non linear regression models. The results obtained showed that the fatalities of the road accidents in Nigeria have a non-linear relationship with population of the road users and time. Abdulkabir *et al.* (2015) study the trend of road traffic accident in Nigeria using Ibadan, oyo state as a case study. They employed time series analysis and their results revealed an increase in the number of accident occurring in the future. Nwankwo and Nwaigwe (2016) analyzed data on road traffic crashes from Anambra State Command of the FRSC using Possion regression, Negative binomial regression and Generalized Poisson regression model. Using the Akaike Information Criterion (AIC), Negative binomial regression model provide the best estimates of road traffic crashes in Anambra State, Nigeria. Oyenuga *et al.* (2016) examined monthly accident data on Oyo-Ibadan express road. Using SARIMA model revealed a positive trend of accident rate on Oyo-Ibadan express road. Adenomon *et al.* (2017) investigated the trend in vehicular accident cases in Nigeria using annual data from 1995 to 2015. Their results using Poisson and negative binomial regression model revealed that accident cases in Nigeria increase by 0.545 percent annually.

It is observed that if nothing is done globally to curtail the rampant nature of road traffic crashes and most especially the causes of deaths due to casualties can increased by the year 2020, while 1.9 million people will be killed by road traffic crashes in the world (WHO, 2004).

Model Specification

(A) Poisson Regression

The classical Poisson and negative binomial regression models for count data belong to the family of generalized linear models (Zeileis *et al*., 2008).

Poisson distribution is the probability distribution of the number of occurrences, X, of some random event, in an interval of time or space. The density function is given by $Pr(X=x)=\frac{\ell^{-\lambda}\lambda^{x}}{x!};x=0,1,...$ the mean and the variance of the distribution are both λ. The skewness of the distribution is $1/\sqrt{\lambda}$ and its kurtosis is $3+1/\lambda$. Poisson regression is a method for the analysis of the relationship between an observed count with a Poisson distribution and a set of explanatory variables (Everitt, 2002).

Poisson models are widely used in the regression analysis of count data and as a basis for categorical data analysis (Dean and Lawless, 1989).

Suppose the observed counts n_i ($n_i \geq 0$) follow a Poisson distribution with parameter λ. Then from Poisson distribution properties, we have $E(n_i)=\lambda$ and $Var(n_i)=\lambda$

We assume here that observed counts occur over a fixed interval, and because these counts are nonnegative. The Poisson regression model is utilized and defined in terms of log of expected counts $(\hat{m}_i)$ as: $\ell_i = X'\beta$ where the **X** represents the explanatory variables (Lawal, 2003).

In this work we will look at the simple case of a Poisson regression as $\log_e(Y)=\beta_o+\beta_1X_1$. In other words, it expresses the log outcome rate as a linear function of a set of predictors. In other words, the typical Poisson regression model expresses the log outcome rate as a linear function of a set of predictors (Adenomon *et al*., 2017).

(B). Negative binomial model

The Negative Binomial (NB) model can be obtained from the mixture of Poisson and Gamma distributions and is expressed as

$$P(y_i/x_i)=\frac{\Gamma(y_i+\alpha^{-1})}{y_i!\Gamma(\alpha^{-1})}\left[\frac{1}{1+\alpha\theta_i}\right]^{\alpha^{-1}}\left[\frac{\alpha\theta_i}{1+\alpha\theta_i}\right]^{y_i}$$

where y_i is the number of road accidents for a road segment i and θ_i is the mean or expected number of persons killed in a road accidents per period, which can be expressed as;

$$\theta_i=\exp(x_i^{'}\beta)$$

The conditional mean and variance of the Negative Binomial distribution are;

$E(y_i/x_i)=\theta_i$ and $Var(y_i/x_i)=\theta_i(1+\alpha\theta_i)>E(y_i/x_i)$, respectively.

Hence the NB model is over-dispersed and allows extra variation relative to the traditional Poisson model. It has more desirable properties than the Poisson model (Chin and Quddus, 2003).This is because the variance of the Negative Binomial is significantly greater than the mean.

The NB model, α represents the dispersion parameter which allows or indicates the degree of over-dispersion. For instance, if $\alpha=0$, the NB model reduces to the traditional Poisson model.

The Log-likelihood function of the NB model is obtained from the following equation

$$\ell=\sum_{i=1}^{n}\left\{\left[\sum_{i=1}^{y-1}\ln(y_i+\alpha^{-1})-\ln y_i!+\alpha^{-1}[\ln(\alpha^{-1})-\ln(\alpha^{-1}+e^{x_i\beta})]+y_i[x_i\beta-\ln(\alpha^{-1}+e^{x_i\beta})]\right]\right\}$$

In order to estimate β and α as in the Poisson model, the iteration procedure or method of Newton Raphson is applied (Lee and Mannering, 2002).

(C). Model assessment: Goodness of fit

After fitting the Poisson regression model, it is very important to check the overall fit as well as the quality of the fit of the respective models. The quality of the fit between the observed values (y) and the predicted values $(\hat{\theta})$ can be measured by the various test statistics, but one useful statistics is called the deviance goodness of fit test which is defined as;

$$D(y:\hat{\theta})=-2\sum_{i=1}^{n}y_i\ln\left(\frac{y_i}{\hat{\theta}_i}\right)-(y_i-\hat{\theta}_i)$$

Where y_i is the number of events (observed values or counts), n is the number of observations and $(\hat{\theta})$ represents the fitted means (predicted values) of the models.

For a better model, one must expect a smaller value of the Deviance. Hence the smaller the value of the deviance obtained for the specific model the better the model.

Material and Methods

Secondary data was obtained from RS 4.3 sector Command Nasarawa State. The data for annual road traffic crash were extracted for the period of 2006 to 2015 for Keffi-Lafia road. The data used in this paper are presented in table 1 and in bar charts in Figs. 1 and 2.

Table 1: A table showing data of road traffic crashes from 2006-2015 on Keffi-Lafia Road

Year	Fatal Cases	Serious Cases	Minor Cases	Total Cases	No. of Persons Killed	No. of Persons Injured	Total Causality
2006	14	43	16	73	53	109	162
2007	9	39	14	62	42	101	143
2008	11	32	13	56	38	99	137
2009	13	44	17	74	41	102	143
2010	15	47	12	74	38	83	121
2011	9	27	7	43	19	93	112
2012	13	17	6	36	21	187	208
2013	28	124	17	169	49	451	500
2014	27	96	6	129	46	353	399
2015	29	92	19	140	54	465	519

Source: RS4.3 Sector command, Nasarawa state

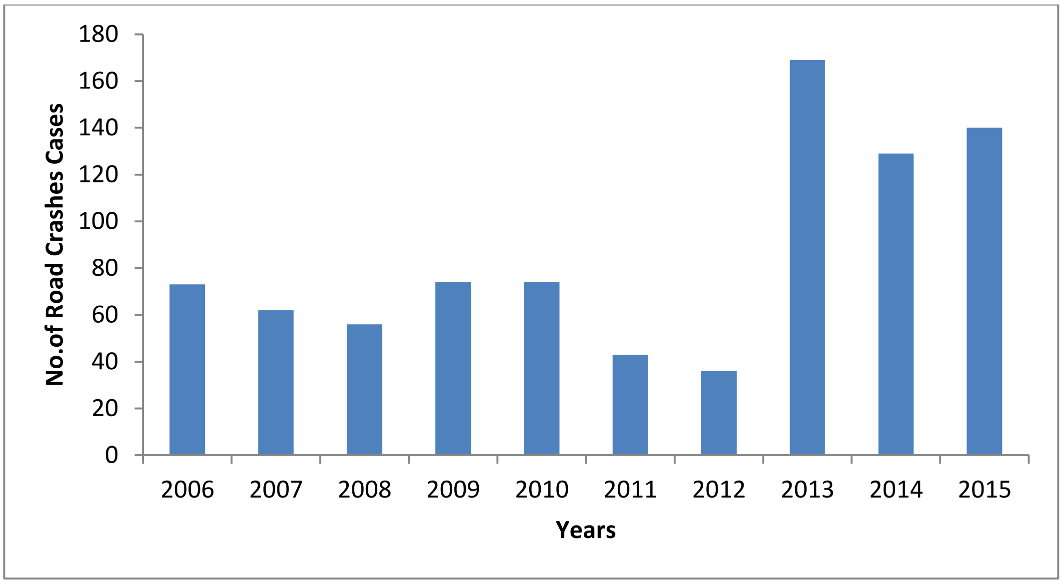

Fig. 1: Graph showing number of total crashes cases from 2006-2015 on Keffi-Lafia Road

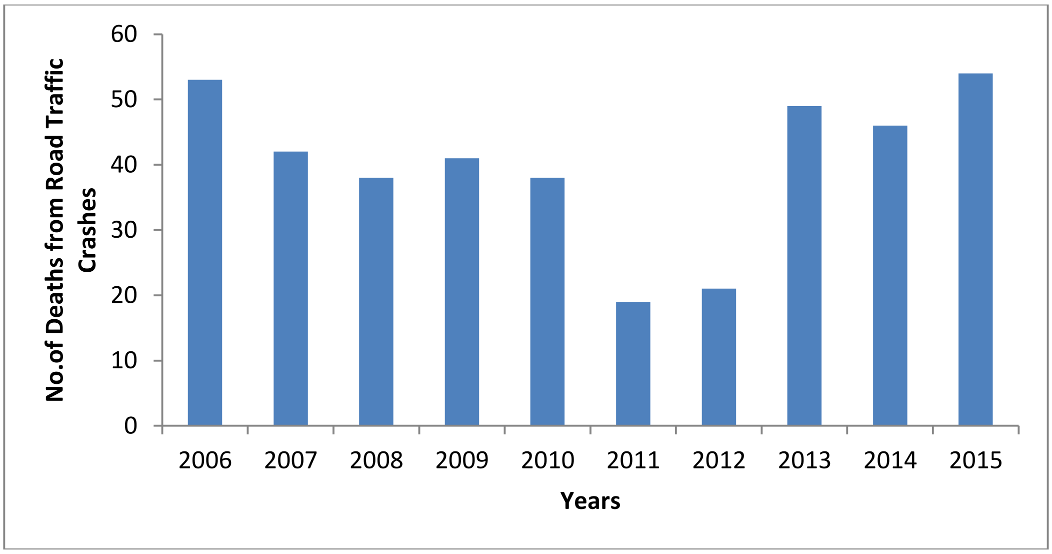

Fig. 2: Graph showing number of death cases from 2006-2015 on Keffi-Lafia road

Table 1 above revealed a total causalities of 519 in 2015 while a total causalities of 112 in 2011. From the Fig. 1, we observed that year 2013 has the highest number of road traffic cases of 169 persons followed by year 2015 and 2014 which is 140 and 129, respectively. This is an indication that road traffic crashes on Keffi-Lafia road is still on the increase. In Fig. 2, It is observed that the death cases in year 2015 is the highest with 54 persons killed followed by 2006 and 2013 with 53 and 49 persons killed, respectively.

Results and Discussion

The data on Keffi-Lafia Road Traffic Crashes was obtained from RS 4.3 Sector command Nasarawa state was analyzed using STATA 12 statistical software.

When there is no effect of cases and time the expected death from Road Traffic Crash on Keffi-Lafia Road is about four (4) persons.

If cases of Road Traffic Crash increase by 1 unit, it will lead to 0.8% increase in log death resulting from Road Traffic Crashes. Furthermore, log death resulting from Road Traffic Crashes decrease by 7.5% annually on Keffi-Lafia Road. The coefficients in the model (cases, time and constant) are significantly different from zero, as (p values < 0.05).

The goodness of fit test for the Poisson regression model revealed that death from Road Traffic Crash on Keffi- Lafia road follows the Poisson distribution (i.e. p value > 0.05). This implies that there is no problem of overdispersion.

Table 2: Results from Poisson regression analysis

```
poisson Death Cases Time

Poisson regression                              Number of obs   =         10
                                                LR chi2(2)      =      26.63
                                                Prob > chi2     =     0.0000
Log likelihood = -32.074754                     Pseudo R2       =     0.2934

------------------------------------------------------------------------------
      Deaths |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       Cases |   .0079345   .0015659     5.07   0.000     .0048655    .0110035
        Time |  -.0746927   .0229759    -3.25   0.001    -.1197245   -.0296608
       _cons |   3.388758   .1197852    28.29   0.000     3.153983    3.623533
------------------------------------------------------------------------------
. estat gof
         Deviance goodness-of-fit =  9.328707
         Prob > chi2(7)           =    0.2299

         Pearson goodness-of-fit  =  9.106805
         Prob > chi2(7)           =    0.2451
```

The Poisson Regression Model is given as; $\log_e$(Death) = 3.3888 + 0.0079cases -0.0747time

Table 3: Results from negative binomial regression analysis

```
. nbreg Death Cases Time, dispersion(mean)

Negative binomial regression                    Number of obs   =         10
                                                LR chi2(2)      =      14.27
Dispersion     = mean                           Prob > chi2     =     0.0008
Log likelihood = -32.074439                     Pseudo R2       =     0.1820

------------------------------------------------------------------------------
      Deaths |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       Cases |   .0079487    .001674     4.75   0.000     .0046678    .0112296
        Time |  -.0747937   .0234395    -3.19   0.001    -.1207344   -.0288531
       _cons |   3.388034   .1238451    27.36   0.000     3.145302    3.630766
-------------+----------------------------------------------------------------
    /lnalpha |  -8.246334   40.15083                     -86.94051    70.44784
-------------+----------------------------------------------------------------
       alpha |   .0002622   .0105283                      1.75e-38    3.94e+30
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) = 6.3e-04 Prob>=chibar2 = 0.490
```

The negative binomial regression model for the data is given as; $\log_e$(Death) = 3.3880 + 0.0079cases – 0.0748time – 8.2483Inalpha

The result from the negative binomial regression model is similar to the result from Poisson regression. Actually, this is so because there is no problem of over dispersion in the dataset used in the project. Another result from analysis revealed that the values of the standard error of the coefficient in the model for Poisson regression are less than the values of the standard error of the coefficient in the binomial regression model. This implies that if count data follows the Poisson distribution there is some evidence that the Poisson regression model will be more efficient than negative binomial regression.

Upon analysis of the data on Road Traffic Crashes on Keffi-Lafia Road, we discover that the results from the Poisson Regression model revealed that the expected log Death of road traffic crash on Keffi-Lafia road if there is no effect of total cases of road crash and time is about four persons. And if the total cases increase by one unit it will lead to 0.8% increase in log death resulting from Road Traffic Crash, while death on road crash on Keffi-Lafia road decreased by 7.5% annually. The reduction of the trend of log death from road crash on Keffi-Lafia road may be due to the frequent sensitization of the negative impact of the road traffic crash by the FRSC to drivers that ply the road. The result of this study on trend of log death from road crash on Keffi-Lafia road contradicts the study of Adenomon *et al.* (2017). Furthermore, the test of over dispersion on the Poisson regression analysis revealed the absences of over dispersion since p-values > 0.05 and this revealed the appropriateness on the use of Poisson regression in this study. The absences of over dispersion in the data set negates the observations by Famoye *et al.* (2004)

and Nwankwo and Nwaigwe (2016) who noted that Poisson regression model always suffer from over dispersion. This justifies the fact that the data used in this study does not suffer from overdispersion.

Conclusion and Recommendations

This study concludes that there is a significant positive relationship between total cases and log death from road crashes from 2006-2015 on Keffi-Lafia road. Similarly the study revealed significant negative trend in log deaths from road crash from 2006-2015.

Looking at road traffic crashes, the state of road and the number of people who are killed via road traffic crashes on Keffi - Lafia road, it is recommended that;

1) Public enlightenment and sensitization of road users on the rules of driving, as well as severe punishment should be administer for road traffic offenders
2) Due to frequent road traffic crashes on our roads, the government should look into the poor state of the country's roads; this has been observed as a major cause of road traffic crashes. Effort should be concentrated on the maintenance of our roads as it is being championed by FERMA
3) The road traffic crash data base of the country should be expanded to include more variables so that researcher could really determine the actual factors contributing to casualties and death in road traffic crashes.
4) Finally, institutions that enforce road traffic regulations should do well to apply the law.

Conflict of Interest

Authors declare that there are no conflicts of interest.

References

Abdulkabir M, Tunde RS & Edem UA 2015. Trend Analysis on Road Accident in Nigeria. *Science Innovation*, 3(5): 52-57.

Adenomon MO, Yahaya HU & Tela MN 2017. Modelling Vehicular Accident Cases in Nigeria: Evidence from Count Data Time Series Analysis. Proceeding of the 1st FNAS National Conference, Nasarawa State University, Keffi. 19th – 22nd March, 2017, pp. 240-246.

Cameron AC & Trivedi PK 1998. Regression Analysis of Count Data. Econometric Society Monographs, No. 30, Cambridge University Press.

Chin HC & Quddus MA 2003. Modelling count data with excess zeros: An empirical application to traffic accidents. *Sociological Methods and Research*, 32: 90-116. http://dx.doi.org/10.1177/0049124103253459.

Dean C & Lawless JF 1989. Tests for detecting over dispersion in Poisson regression models. *J. Am. Stat. Assoc.*, 84(406): 467-472.

Everitt BS 2002. The Cambridge Dictionary of Statistics (2nd ed.). New York: Cambridge University Press.

Famoye F, John TW & Karan PS 2004. On the generalized Poisson regression model with an application to accident data. *Journal of Data Science*, 2: 287-295.

Fletcher D, Hanton S & Mellalieu SD 2006. An organisational stress review: Conceptual and theoretical issues in competitive sport. In Hanton D & Mellalieu (Eds), Literature Reviews in Sport Psychology. Hauppauge, NY: Nova Science, pp. 321-373.

FRSC 2009. Federal Road Safety Corps (FRSC) Annual Report 2009. Nigeria.

Kareem B, Oke PK & Lawal AS 2012. Modelling fatalities of road accidents in Nigeria. *Research in Logistics & Production*, 2(3): 259-271.

Lawal HB 2003. Categorical Data Analysis with SAS and SPSS Application. London: Lawrence ERlbawn Association Publisher.

Lee J & Mannering F 2002. Impact of road side features on the frequency and severity of run-off-road accidents: An empirical analysis. *Accident Analysis and Prevention*, 34(2): 149-161.

NRTR 2004. Annual Report, National Road Traffic Regulations, Nigeria.

Nwankwo CH & Nwaigwe GI 2016. Statistical model of road traffic crashes data in Anambra State, Nigeria: A Poisson regression approach. *Eur. J. Stat. & Prob*., 4(1): 41 – 60.

Oyedepo OJ & Makinde OO 2010. Accident prediction model for Akure-Ondo carriageway, Ondo State Southwest Nigeria using multiple linear regression. *African Research Review*, 4(2): 30-49.

Oyenuga IF, Ayoola FJ & Shittu OI 2016. Statistical analysis of pattern on monthly reported road accidents in Nigeria. *Sci. J. Appl. Math. & Stat.*, 4(4): 119-128.

Ukwu J & Adedokun A 2016. Federal Road Safety Commission , Department of Policy, Research and Statistic Publication, Nigeria.

WHO 2004. World Report on Road Traffic Injury Prevention: Summary. World Health Organization (WHO), Geneve, Switzerland.

Zeileis A, Kleiber C & Jackman S 2008. Regression models for count data in R. *Journal of Statistical Software*, 27(8): 1-25.