Zimbabwe Journal of Science & Technology pp 100 – 110 Vol.14 [2019]
e-ISSN 2409-0360
Zimbabwej.sci.technol

STATISTICAL ANALYSIS OF ROAD TRAFFIC ACCIDENTS FATALITY USING POISSON AND NEGATIVE BINOMIAL REGRESSION (CASE STUDY OF OSUN STATE FRSC)

Aladeniyi O. B., *Bodunwa O. K and Adesina, A. S
Department of Statistics, Federal University of Technology,
Corresponding Author's e-mail; okbodunwa@futa.edu.ng

ABSTRACT

Poisson regression and negative binomial regression methods of estimation are the two important estimation methods used/employed in estimating count data. In this work, we compared the two methods in order to determine the better for modeling data with over dispersion. The Osun state FRSC sector command data of Road Traffic Accidents (RTAs) fatality was used for the analysis. The dispersion parameter was used in detecting over dispersion in the data. The results of the Poisson regression indicated that there was over dispersion in the data. Therefore, negative binomial regression analysis was to compared with the Poisson regression model. The Akaike Information Criterion (AIC) of the two methods was also compared to identify the better for fitting model. The lower of the AICs gives the better fit model. The negative binomial regression provides a better fit with relatively low AIC. The conclusion is that Poisson regression underestimates the standard errors of the estimated coefficients when the variance-mean ratio is greater than one making it difficult to detect insignificant regressor(s). Thus, negative binomial regression provides a better fit with relatively low AIC.

KEYWORDS: Poisson Regression, Negative Binomial Regression, Road Traffic Accidents, Federal Road Safety Commission
Received: 28.05.19 **Accepted:** 27.07.20

1.0 INTRODUCTION

Road Traffic Accident (RTA) is collisions between vehicles and pedestrians; between vehicles and animals; or between vehicles and geographical or architectural obstacles. Road traffic accidents in Nigeria have been identified as one of the major causes of deaths in the country. Aside from Boko Haram crisis, accidents are currently by far the main cause of violent death in Nigeria. It is classified as the second major cause of death in the Country. The WHO adjudged Nigeria the most dangerous Country in Africa with 33.7 deaths per 100,000 people every year. According to their report, one in every four road accident deaths in Africa occurs in Nigeria. Similarly, one out of 50 accidents on Nigerian roads occurs in Osun State. The WHO survey and the FRSC report of 5,693 fatal road accidents in 2014 leave no doubt about the dangerous situation on Nigeria roads. Road traffic accidents have become a normal and regular phenomenon in Nigeria. According to research carried out by World Health Organization (WHO), out of 193 Countries ranked, Nigeria has the second-highest rate of road accidents in the world. Researchers have been modeling vehicular accidents with crash prevention 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 are many factors that can be responsible for this occurrence such as type of vehicle, nature of the road, vehicle age and days of casualty contribute to casualty survival in road accidents in the State and have still not been considered. It is therefore important to statistically analyze the accident data to ascertain the truth or otherwise of these possibilities. When it is confirmed that there is a relationship, then it will be prudent to apply some mathematical and statistical models such as Poisson regression and/or

Negative Binomials to fit a model to the accident data for better prediction for decision making. Poisson regression analysis is a technique used to model dependent variables that describe count data (Cameron et al, 1998). It is often applied to study the occurrence of small number of counts as a function of a set of predictor variables in experimental and observational study in many disciplines, including Economy, Demography, Psychology, Biology and Medicine (Gardener and Trivedi, 1995).

Miauo (1994) compared Poisson and Negative Binomial regressions since they both cater for the distributional properties of accident data. Many other researchers have also assessed the use of linear regression models for road accident models and confirmed the limitations in such models,

Ghee *et al*. (1997) stated that fatality rate is defined as the number of injury accidents occurring per annum per million vehicle kilometer travelled. But since there is no much reliable accident data base in developing countries and much information required to compute this type of fatality rate, Ghee *et al*. (1997) defined the fatality rate for road accidents in a given country to be measured in respect of the number of persons killed through road accidents per 10,000 licensed

vehicles in a country. As population increases and the number of licensed vehicles in developing Countries (Nigeria inclusive and Osun State as a unit) are rising rapidly.

However, Jacobs and Aeron-Thomas (2000) suggested that this index cannot be used to compare accident fatality rates of different countries (Nigeria inclusive and Osun State as a unit) since the countries may vary in terms of population and total vehicles which ply their roads. He then proposed a model which assessed the relationship between fatalities, population and motorization of the country. This model supported the Smeed Formula for international comparisons of accident fatalities. Smeed (1938, 1968). Smeed in 1938 used accident data from different countries and proposed the formula

$$\frac{D}{N} = 0.0003\left(\frac{N}{P}\right)^{-0.67}$$

where D is annual number of fatalities from road accidents, N is number of vehicles in use and P is population. This model was confirmed Smeed (1968). Kim et al (2005) used generalised log-linear models and Garber and Wu (2001) applied stochastic models in fitting models to road accidents data. However, one should be more cautious of the use of the Poisson and the Negative binomials since the estimation of the various parameters could be misleading, Miauo (1994). However, in Nigeria much work has not been done in this regard but researchers such as Salifu (2004) used the generalized linear model to predict road accidents in unsignalised urban junctions, Afukaar et al (2007) and Ackaah and Salifu (2012) applied Poisson, Negative Binomial and Log-Linear regression models to accident data.

2.0 Material and Methodology

Secondary data were used. The data were collected from the state secretariat, the Federal Road Safety Commission (FRSC) Osun sector command. This study considered accident data for ten year period from 2005 to 2014. The number of people killed by road accident was used as the response variable in all models and the other variables such as age of casualty, the day and month the accident occurred which resulted in the death of the people, vehicle type and road user class as the explanatory variables. The Negative Binomial distribution was used to correct the error of over dispersion in the data in situations where the result of the Poisson regression model shows over dispersion. The various models obtained when number of people killed by road accidents was regressed on factors such as age, the day the accident which killed the people occurred and the year that the accident occurred.

In spite of its recent wide application, Poisson regression model remains partly poorly known, especially if compared with

other regression techniques, like linear, logistic and Cox regression models.

2.1 POISSON REGRESSION AND NEGATIVE BINOMIAL DISTRIBUTION

In spite of its recent wide application, Poisson regression model remains partly poorly known, especially if compared with other regression techniques, like linear, logistic and Cox regression models. The Poisson regression model assumes that the sample of n observations, are observations on independent Poisson variables Y_i with mean u_i

If this model is correct, the equal variance assumption of classic linear regression is violated, since the Y_i have means equal to their variances.

So we fit the generalized linear model,

$$log(u_i) = x_i^{'}\beta \quad (1)$$

We say that the Poisson regression model is a generalized linear model with Poisson error and a log link, so that

$$u_i = exp(\beta) \quad (2)$$

This implies that one unit increases in an x_i are associated with a multiplication of u_i by exp(βi).

The major assumption of the Poisson model is equation (3)

$$E\ (y_i \mid x_i) = u_i = e^{xi^{'}\beta} = Var\ (y_i \mid x_i) \quad (3)$$

Implying that the conditional mean function equate the condition variance function.

This is very restrictive. If $(y_i \mid) < Var\ (y_i \mid x_i)$ then we speak about overdispersion, and when $E\ (y_i \mid x_i) > Var\ (y_i \mid x_i)$ we say we have underdispersion. The Poisson model does not allow for over or underdispersion. A richer model is obtained by using the negative binomial distribution instead of the Poisson distribution.

Instead of equation

$$P\ \{Y = y_i\} = \frac{e^{-\mu}\ \mu^{yi}}{y!} \quad (4)$$

we then use

$$P\left(Yi = \frac{yi}{\beta}, xi\right) = \frac{\Gamma\ (\theta + yi)}{\Gamma\ (yi+1)\Gamma\ (\theta)}\left(\frac{\lambda_i}{\lambda_i + \theta}\right) yi\ (1 - \frac{\lambda_i}{\lambda_i + \theta})\theta \quad (5)$$

This negative binomial distribution can be shown to have conditional mean λi and conditional variance (1 + $\eta^2 \lambda i$) with $\eta^2 := \frac{1}{\theta}$. The parameter η^2 is not allowed to vary over the observations. As before, the conditional mean function is modeled as

$$E\ (y_i \mid x_i) = u_i = e^{xi^{'}\beta} \quad (6)$$

The conditional variance function is then given by

$$Var\ (y_i \mid x_i) = e^{xi^{'}\beta}\ (1 + \eta^2 e^{xi^{'}\beta}) \quad (7)$$

Using maximum likelihood, we can then estimate the regression parameter, and also the extra parameter η. The parameter η measures the degree of over (or under) dispersion. The limit case

η = 0 corresponds to the Poisson model.

3.0 RESULTS AND DISCUSSION

3.1 Annual Distribution of Fatality of Road Accidents in Osun State

The Table 1 below shows the total number of people killed in road accident annually from 2005-2014 in years, days and ages/stages respectively.

Table 1: Total Number of people killed In Road Accidents

		Victim	%			Victim	%
Year	2005	85	3.8	**Day**	Monday	372	16.5
	2006	172	7.6		Tuesday	281	12.4
	2007	266	11.8		Wednesday	230	10.2
	2008	266	10.0		Thursday	301	13.3
	2009	390	17.3		Friday	423	18.7
	2010	379	16.8		Saturday	330	14.6
	2011	252	11.2		Sunday	322	14.3
	2012	116	5.1	**Age(year)**	0-2	69	3.1
	2013	192	8.5		3-12	471	20.8
	2014	181	8.0		13-19	222	9.8
					20-45	1209	53.5
					>45	288	12.7

The most significant feature of Table 1 is that the number of people who were killed by road accidents in Osun State seems to be increasing as years go by until 2012 when it started coming down. In 2005 there were 85 people who were killed in road accidents, this was increased to 172 in 2006 and in 2007 the number rose to 266. By 2009, the number had risen to 390. However, there were sharp decreases in 2012, 2013 and 2014 with the number of who were killed in road accidents being 116,192 and 181 respectively. From the table, one observes that Friday has the highest number (423) of people who were killed by road accidents from 2005 to 2014 in Osun state which represent 18.7% of those killed by road accidents. This was followed by Monday which had 372 people constituting 16.5% of those killed by road accidents. Saturday and Sunday recorded the same number of people who were killed in road accidents within the 10 year period. Wednesday recorded 230 which was the least number of people who were killed in road accident representing 10.2%. It is clear from the table that the youth is the most vulnerable group to road accident. With 1209 recorded casualties. Followed by Children which recorded 471 casualties from 2005-2014. This result is not surprising since research has shown that most people who are at risk in road accidents are in the ages between 15 and 44 years. The age group which recorded the least number of deaths is

the babies with just 69 casualties for the years under consideration.

3.2 Modeling the Number of People Killed in Road Accidents in Osun State.

3.2.1 The days in which people were killed in road accident

The days in which people were killed in road accident from 2005 to 2014 was modeled using Poisson regression and are presented in the Table 2.

Table 2: Parameter Estimates of the Poisson Model for Days of Fatality

Coefficients	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.97973	0.07700	51.683	< 2e-16 ***
wkdayMonday	-0.12848	0.07108	-1.808	0.070678**
wkdayTuesday	-0.40902	0.07696	-5.315	1.07e-07 ***
Wkdaywednesday	-0.60929	0.08193	-7.437	1.03e-13
wkdayThursday	-0.34026	0.07541	-4.512	6.41e-06
wkdayFriday	-0.34026	0.00531	-1.529	0.006741 *
wkdaySaturday	-0.24828	0.07345	-3.380	0.000724 ***
year2006	-0.11501	0.08800	-1.307	0.191225 ****
year2007	-0.03301	0.08814	-0.374	0.708062 ***
year2008	-0.04025	0.08829	-0.456	0.648461 **
year2009	-0.06978	0.08891	-0.785	0.432569 *
year2010	-0.25435	0.09308	-2.733	0.006284 **
year2011	-0.27253	0.09352	-2.914	0.003568
year2012	-0.43667	0.09780	-4.465	8.01e-06 ***
year2013	-0.78183	0.10860	-7.199	6.05e-13 ***
year2014	-0.66659	0.10470	-6.367	1.93e-10 ***

Table 2 shows the parameter estimates of the selected model. The AIC of this model was 489.61; the null deviance was 294.483 on 69 degrees of freedom and residual deviance of 90.604 on 54 degrees of freedom following the chi-square distribution ($\chi 2$) with one degree of freedom. The dispersion parameter was found to be 1.9931 and p-value of 1.959054e-07 which indicates that the model is significant. However, one assumption of Poisson distribution which is the equality of the mean

and variance which means that the dispersion parameter should always be closer to 1 has been violated. The dispersion parameter of the above model is far greater than 1, an indication of over dispersion in the data. This means that the parameters of the model have been over estimated and will not give a true reflection of number of people likely to be killed through road accidents in a given day of the week for a particular year. To eliminate this error, Negative Binomial regression was used and the result is shown in the Table 3.

Table 3: Parameter Estimates of the Negative Binomial Regression Model for Days of Fatality

Coefficients	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.98228	0.08458	47.080	< 2e-16 ***
wkdayMonday	-0.12838	0.07774	-1.652	0.09863 *
wkdayTuesday	-0.40939	0.08316	-4.923	8.52e-07 ***
Wkdaywednessday	-0.61076	0.08780	-6.956	3.50e-12 ***
wkdayThursday	-0.34001	0.08171	-4.161	3.17e-05 ***
wkdayFriday	-0.002368	0.06543	-2.2871	0.00324****
wkdaySaturday	-0.24915	0.07992	-3.117	0.00183 **
year2006	-0.11407	0.09586	-1.190	0.23404 **
year2007	-0.03322	0.09633	-0.345	0.73022 **
year2008	-0.04445	0.09654	-0.460	0.64526 ***
year2009	-0.07234	0.09708	-0.745	0.45616 **
year2010	-0.25904	0.10096	-2.566	0.01030 *
year2011	-0.27445	0.10131	-2.709	0.00675 **
year2012	-0.43783	0.10526	-4.160	3.19e-05
year2013	-0.78292	0.11535	-6.787	1.14e-11 *
year2014	-0.66944	0.11175	-5.991	2.09e-09 **

In Table 4, comparison of the assessment criteria for Poisson and Negative Binomial regression models were looked into. The dispersion parameter for the negative binomial model was 1.8948 which is acceptable when modeling a count data set with variance greater than the mean. Also, it could be observed from the above table that the AIC of the negative binomial regression model is 489.61 which is smaller than that of Poisson regression model of AIC 490.97 an indication of a better model. The R package takes the first category in a data as the base level by default and as such Sunday and

2005 were picked as the base levels for comparison in the analysis of the parameter estimates in the model. From table 3 the intercept was found to be 3.98228 which was very significant at 95% significance level. All the days were significant in the model at 5% α-level but the most significant day was found to be Friday with coefficient of -0.02368. This means that the expected number of people killed through accident on Friday was e-0.2368=0.997635 more than that of Sunday for every year. The day with the least significance in the model was Wednesday which had a parameter estimate of -0.61076 indicating a reduction of e−0.61076=0.542928 which means that the expected number of people who were killed on Wednesday was 7.8% lower for each year.

Table 4: Parameter Comparison between Poisson and Negative Binomial Regression for Goodness of Fit Test of Days of Fatality

Assessment Parameter	Poisson Regression Model	Negative Binomial Regression Model
Null Deviance	294.483	257.713
Degree of Freedom	69	69
Residual Deviance	90.604	80.055
Degree of Freedom	54	54
Dispersion parameter	1.9931	1.8948
AIC	490.97	489.61

From Table 3 the expected number of people killed in road accident for the year 2005 on Sundays is e3.98228 = 53.6392. This means that the expected number of people who were killed in 2005 on Sundays only was 54. From the Table, the fitted model for the accident data using day and year can be formulated as shown in equation 9 below.

$$\log(mean\_killed) = 3.98228 - 0.12838D_1 - 0.40939D_2 - 0.61076D_3 - 0.34001D_4 - 0.00237D_5 - 0.24915D_6 - 0.11407Y_1 - 0.03322Y_2 - 0.04445Y_3 - 0.07234Y_4 - 0.25904Y_5 - 0.27445Y_6 - 0.43783Y_7 - 0.78292Y_8 - 0.66944Y_9 \quad (9)$$

where $D_1, D_2, \cdots, D_6$ represent Day1 (Monday), Day2 (Tuesday),…, Day6(Saturday) respectively and $Y_1, Y_2, \cdots, Y_9$ denote 2006,2007,…,2014 in that order.

For instance, to obtain the expected number of people killed in any day in a given year, the day and the year are represented by 1 in the equation and the rest are represented by 0. For example the expected number of people killed on Monday in the year 2006 is given by $e^{3.98228-0.12838(1)-0.11407(1)} = 284.3350$. That is, the expected number of people killed on Mondays in the year 2006 was 46.7367.

3.2.2 Age Group/ Stages of Life of those killed through Road Accidents

This section talks about the age group or stages of life of people that were affected in road accidents.

Table 5: The Parameter Estimates of Poisson Model for Age Group/Stages of life of those who were killed by Road Accident in Osun State.

Parameter	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	5.04819	0.06181	81.668	
Stage children	-0.94269	0.05432	-17.356	< 2e-16 ***
Stage teenager	-1.69487	0.07302	-23.212	< 2e-16 ***
Stage youth	-1.69487	0.07302	-23.212	< 2e-16 ***
Stage aged	-1.43459	0.06557	-21.879	< 2e-16 ***
Year2006	0.03046	0.08228	0.370	< 2e-16 ***
Year2007	0.08242	0.08125	1.014	0.711236
Year2008	-0.28883	0.08957	-3.224	0.001262 **
Year2009	-0.16799	0.08661	-1.940	0.052427
Year2010	0.56738	0.09745	-5.822	5.81e-09 ***
Year2011	0.30735	0.09005	-3.413	0.00643 ***
Year2012	-0.40547	0.09269	-4.375	1.22e-05 ***
Year2013	-0.56738	0.09745	-5.822	5.81e-09 ***
Year2014	-0.66269	0.10051	-6.593	4.31e-11 ***

The Table 5 presents the parameter estimates of the Poisson model. The AIC of this model was 363.69 with 63 degrees of freedom whereas the dispersion parameter was found to be 2.7715 and which indicates that the model is significant at 5% $\alpha-$level. However, the assumption of equal variance to the mean in Poisson distribution has been violated since the dispersion parameter is not approximately equal to1. This means that the parameters of the model have been over estimated and the standard errors have been under estimated and will not give a true reflection of the model for the age/stages of life of people of fatality via road accidents in Osun State from 2001 to 2010. To address this error, Negative Binomial regression will be employed to check if the dispersion parameter can be reduced further as shown in table 6 below.

Table 6: The Parameter Estimates of Negative Binomial Model for Age Group/Stages of Life of those who were killed by Road Accident in Osun State.

Coefficients	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	5.062699	0.091942	55.064	< 2e-16 ***
Stage.children	-0.932718	0.074756	-12.477	< 2e-16 ***
Stage.teenager	-1.685458	0.089330	-18.868	< 2e-16 ***
Stage.youth	0.855907	0.134107	-21.296	< 2e-16 ***
Stage.aged	-1.427494	0.083387	-17.119	< 2e-16 ***
year2006	-0.014729	0.119000	-0.124	0.901492
year2007	-0.005467	0.118850	-0.046	0.963311
year2008	-0.249727	0.123164	-2.028	0.042602 *
year2009	-0.169906	0.121666	-1.396	0.162565
year2010	-0.519607	0.128945	-4.030	5.59e-05 ***
year2011	-0.386175	0.125942	-3.066	0.002167 **
year2012	-0.419624	0.126667	-3.313	0.000924 ***
year2013	-0.558935	0.129888	-4.303	1.68e-05 ***
year2014	-0.707095	0.133694	-5.289	1.23e-07 ***

Table 7: Parameter Comparison between Poisson and Negative Binomial Regression for Goodness of Fit Test for Age/Stages of Life

Assesssment parameter	Poisson Regression Model	Negative Binomial Regression Model
Null Deviance	1804.97	1108.216
Degree of Freedom	49	49
Residual Deviance	562.21	48.515
Degree of Freedom	36	36
Dispersion parameter	2.7715	2.4436

AIC	363.69	358.37

From the Table 7 it is observed that the dispersion parameter for the negative binomial regression model is 2.4436which smaller than that of Poisson, an indication of a good sign of reduced over dispersion in the data as compare to Poisson regression.
By default, R package takes the first category in a data as the base level and as such stagechildren and the year 2005 were picked as the base levels for comparison in the analysis of the parameter estimates in the model as shown in table 6. The intercept was found to be 5.062699 and very significant at 95% significant level. With the exception of people from teenager which was not significant in the model, the rest were all significant in the model at 5% α –level. It is observed that teenager who were killed via road accidents was significantly smaller with parameter estimate -1.685458 having expected value of $e^{-1.685458}$ = 0.18525. The most significant age group in the model was found to be 26-35 years with parameter estimate of 0.855907 and log expected value of $e^{0.855907}$ = 2.3535. This means that the expected number of people killed through accident in the age group of 26-35yrs were truly the most vulnerable to fatal accident.
The model for the above table is presented in equation 10 below

Log $(mean\_killed)$ = 5.062699 -0.932718K_1 – 1.685458K_2+ 0.855907K_3 - 1.427494K_4 - 0.014729Y_1 - 0.005467Y_2 – 0.249727Y_3 - 0.169906Y_4 - 0.519607Y_5 - 0.386175Y_6 - 0.419624Y_7 - 0.558935Y_8 - 0.707095Y_9

(10)

Where $K_1, K_2, \cdots, K_4$ represent stagechildren, …,stage.aged respectively and $Y_1, Y_2, \cdots, Y_9$ denotes 2006,2007,…,2014 in that order.

CONCLUSION

This research aimed at modeling the number of people who are killed by road accidents in Osun State.Based on the results from the analysis above, it can be concluded that
The Poisson regression underestimates the standard errors of the estimated coefficients when the variance-mean ratio is greater than one, therefore making it difficult to detect insignificant regresssor(s).
The negative binomial regression provides a better fit with relatively lower AIC than poisson regression when the variance-mean ratio is greater than one.
The results also revealed that the most affected people who die through road accidents in Osun State are the youth (that is, the breadwinners of families).This gives us a cause to worry since the work force of the country is being killed by road accidents. Out of 2259 people who are killed via road accidents, 1209 of them are youth that is,people in the age group of 26-40years, 471 are children that is, people in the age group 3-12years, and the least of all are the babies that is, people in the age group 0-2year(s).
The preliminary analysis of the number of people who are killed at any given day in the week showed that most people are killed on Fridays in road accidents and Wednesday is the day which recorded the least number of people killed by road accident. This result was verified using both poisson and negative binomial regression models. It was discovered that 423 people were killed on Fridays as the highest and lowest recorded as Wednesday with 230 people killed for the ten years under consideration. It was found out from the Poisson mean analysis that the expected number of people who are killed in road accidents on Monday, Tuesday, Thursday, Saturday and Sunday is not different from the grand mean of the number of people killed by road accident every year.

REFERENCES

Afukaar F.K., and Debrah E.K., (2007). Accident prediction for signalized intersections in

Ghana. Ministry of transportation: Nation Accident management project, report. Final report

Ackaah, Williams; Salifu, Mohammed (2011): Crash prediction model for two-lane rural highways in the Ashanti region of Ghana. In *IATSS Research* 35 (1), pp. 34–40.

Cameron, C., Trivedi, P., (1998). *Regression Analysis of count data*. Cambridge University Press, New York.

Fletcher J.P., Baguley C.J., sexton B., and Done S. (2006) Road accident modeling for highway development counties. Main Report Trials in India and Tanzania. Report No PPRO95.

Garber, N. J., Wu L., (2001) Stochastic model relating to crash probabilities with geometric and corresponding traffic characteristics data, UVACTS-5: 15-74

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). *Regression analyses of counts and rates: Poisson, overdispersedpoisson, and negative binomial models.* Psychological Bulletin, 118, 392–404

Ghee C., Silcock D., Astrop A., and Jacobs G. (1997). Socio-economic aspects of road accidents in developing countries, Transport Research Laboratory Report 247:29.

Jacobs, G.; Aeron-Thomas, A. (2000): Africa Road Safety Review, Final Report: U.S. Department of Transportation.

Kim S.H., Chung S.B., and Song K.S. (2005). Development of an accident prevention model using GLM (generalized log -lirear model) and EB Method a case of seal. Journal of the Eastern Asia society for Transportation studies 6:3669-36682

Miaou S.P. (1994). The relation between track accident and geometric design of road section: Poisson vesrus negative binomial regression. Accident Analysis and Prevention. 26(4) :471-482.

Salifu, Mohammed; Ackaah, Williams (2012): Under-reporting of road traffic crash data in Ghana. In *International journal of injury control and safety promotion* 19 (4), pp. 331–339.

Smeed R.J. (1968). Variation in the pattern of accident rates in different countries and their causes. Traffic Engineering and Control. 10:364-371.