Regression Analysis of Road Traffic Accidents and Population Growth in Ghana

Boakye Agyemang¹, Dr G. K. Abledu², Reuben Semevoh³

ARTICLE INFO
Available Online October 2013
Key words: regression model; road traffic accidents; population; statistical evidence; Ghana.

ABSTRACT
Vehicular accidents in Ghana have become one of the growing concerns to most Ghanaians in recent times. This is as a result of the tremendous negative effects of accidents on human lives, properties, and the environment as well as and the economy at large. Many researchers have come out with the causes, effects and recommendations to vehicular accidents, which include drunk driving, machine failure and over speeding just to mention a few. Yet every year the Road Safety Commission, Ghana Statistical Service and other statutory organizations report an increase in vehicular accidents in Ghana.

The systematic yearly increase in the number of accidents therefore calls for statistical analysis of this global and of course national canker. This study attempts to show statistical evidence of relationship between road traffic accidents and population growth in Ghana in order to ascertain additional information in contributing to previous researches that have emerged in dealing with this menace. Time series data on yearly road traffic accidents and population values for Ghana covering the period 1990 to 2012 was used. The results from the analysis shows three key findings: a systematic visible pattern of growth in both road traffic accidents and population over the period; evidence of statistical relationship between road traffic accidents and population growth in Ghana as given by the correlation coefficient (r) of 0.854, with a corresponding coefficient of determination (r-square) of 72.9% indicating that for the period under study based on the available data, population is able to account for 72.9% of the changes in accidents in Ghana; and finally a regression model developed for the purposes of estimating and forecasting on the basis of the analysis, specifically based on test of hypothesis and model validation.

1 Introduction

Vehicular accidents in Ghana have become one of the worrying and growing concerns to most Ghanaians in recent times. This is as a result of the tremendous negative effects of road traffic accidents on human lives, properties and the environment. As a result of the negative impact of road traffic accidents in Ghana, the National Road Safety Commission (NRSC) was therefore established by an Act of Parliament (Act 567) in 1999 with the mission of promoting best road safety practices for all categories of road users through the vision of making Ghana a country with the safest road transportation system in Africa.

Many researchers have come out with the causes, effects and recommendations to vehicular accidents, which include drunk driving, machine failure and over speeding just to mention a few. Yet every year there seems to be an un-ending situation with regard to road accidents in Ghana. Some important global facts pertaining to road traffic accidents are as follows: at least 1.3 million people are killed every year through road crashes; some 20-50 million suffer various forms of disability; 90% of these road casualties are in low and middle income countries; at current rates, it is estimated that road traffic crashes will be the 3rd leading cause of death worldwide by the year 2020 if rigorous actions are not taken. Also, 1,800 lives are lost in Ghana annually through road crashes with 14,000 injuries from an average of 11,000 road traffic crashes; road traffic crashes cost the nation 1.6% of GDP which translates to US$ 288 MILLION (in 2009): 42% of

¹ Koforidua Polytechnic, Applied Mathematics Department
² Koforidua Polytechnic, Applied Mathematics Department
³ Koforidua Polytechnic, Applied Mathematics Department
these crash victims are pedestrians; 60% of all crash victims are people within the productive age group (NRSC, 2010). The above statistics are worrying and also implies the indirect destruction of the nation’s human assets, as well as economic underdevelopment and retrogression rather than economic development and progression.

According to a World Health Organization (WHO) & World Bank (1999) report on “The Global Burden of Disease”, deaths from non-communicable diseases are expected to climb from 28.1 million a year in 1990 to 49.7 million by 2020 (an increase in absolute numbers of 77%). Road traffic crashes will contribute significantly to this rise. According to the report, road traffic injuries are expected to move from ninth place to take third place in the rank order of disease burden by the year 2020.

Table 1: Projected Rank Order of Disease Burden from 1990 to 2020

<table>
<thead>
<tr>
<th>1990</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Respiratory Infections</td>
<td>1</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>2</td>
</tr>
<tr>
<td>Perinatal</td>
<td>3</td>
</tr>
<tr>
<td>Unipopular Major Depression</td>
<td>4</td>
</tr>
<tr>
<td>Ischaemic heart</td>
<td>5</td>
</tr>
<tr>
<td>Cerebrovascular</td>
<td>6</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>7</td>
</tr>
<tr>
<td>Measles</td>
<td>8</td>
</tr>
<tr>
<td>Road Traffic Crashes</td>
<td>9</td>
</tr>
<tr>
<td>Congenital Anomalies</td>
<td>10</td>
</tr>
</tbody>
</table>

Source: www.grsroadsafety.org

In assessing the magnitude of the problem of road traffic crashes, according to WHO, 1.2 million people die through road traffic crashes annually. On the average, in the industrialized countries, and also in many developing countries, one out of every ten hospital beds is occupied by a road traffic crash accident victim (NRSC 2010).

The 1999 WHO publication on "Injury: A Leading Cause of the Global Burden of Disease," reports that road traffic crashes are the major cause of severe injuries in most countries and the leading injury-related cause of death among people aged 15-44 years. Globally, the WHO reports that 38,848,625 people were injured through motor vehicle crashes in 1998. Out of the 5.8 million people who died of injuries, 1,170,694 (20%) died as a direct result of injuries sustained in motor vehicle crashes. Figure 2 below summarizes the traffic statistics in the WHO report.

The above facts reveal unacceptable levels of road traffic accidents and casualties and therefore have both global and national social and economic burden, especially in Ghana as 1.6% of our GDP. There is therefore the need for interventions and strategies to deal with the menace especially by reducing it by 50% by the end of 2020 as recommended in the United Nations (UN) Global Plan for the Decade of Action for Road Safety 2011-2020.

Unfortunately, in Ghana there is an information gap in respect of population and road traffic crashes as enough research have not been conducted in this field, hence the need to analyze the accidents data statistically in relation to the population in order to bridge this information gap in Ghana.

1.1 Objectives
The objectives of this study are:

- To identify the population and road traffic trends in Ghana;
- To ascertain whether there exists a relationship between road traffic accidents and population in Ghana;
- To establish the nature of the relationship between road traffic accidents and population in Ghana;
- To derive a model for the relationship between road traffic accidents and population in Ghana;
- To forecast the road traffic accidents based on population in Ghana;
1.2 Hypothesis of the study
Null Hypothesis (Ho): There is no significant relationship between road traffic accidents and population growth in Ghana.
Alternative Hypothesis (H₁): There is a significant relationship between road traffic accidents and population growth in Ghana.

1.3 Significance of the study
The study would serve as a guide to stakeholders within the road transport system in making informed and intelligent policy decisions with regard to the management of road traffic accidents in Ghana. Also, the information that would be unraveled from this study when put into effective and better utilization in the management of road traffic accidents would inure to the economic benefit of saving cost to the nation at large.

Finally, it would provide information that would help bridge the information gap that exists in our part of the world and add to the body of literature for researchers who may require such information.

2 Methodology
A linear regression model was used to analyze the dynamics of changes, variations and interruptions in road traffic accidents and population growth of Ghana through time-series data as described below:

2.1 Data source
The study relied heavily on secondary data obtained from the National Road Safety Commission of Ghana. The data specifically comprised of time series data on yearly road traffic accidents and corresponding population values for Ghana covering the period 1990 to 2012. Details of the data are presented in appendix I.

2.2 Method of data analyses
Data analyses were performed using Minitab. Exploratory data analyses were conducted to investigate patterns, relationships and checking of assumptions underlying the use of regression, specifically using scatter and residual plots of the data. This was followed by inferential analysis and subsequently interpretation of results as well as drawing valid conclusions. It is important to state that for all analyses P-value < 0.05 was considered to be statistically significant.

2.3 Model development and assumptions
The following assumptions as applied to fitting regression models were verified.
1. There must be a linear relationship between X and Y, in this case road traffic accidents and population. Of course non-linear relationships are very common, and can be easily addressed.
2. The error terms should average out to zero, in the long run.
3. The variance of the error terms should be constant across observations. That is also known as homoskedasticity.
4. The error terms are not autocorrelated, any one residual is not correlated with any other residual.

3 Results
3.1 Exploratory data analysis
Figure 1 is the scatter plot showing the relationship between number of road traffic accidents and population in Ghana. It can be seen from the scatter plot that there exists a systematic visible pattern in the data resulting in the existence of a relationship between the two variables, but no clear outlier. As the population grew from year to year, it grew along with the number or the rate road traffic accidents indicating a positive relationship between these variables. However, the magnitude of the relationship is given by correlation coefficient in the subsequent analysis.
Figure 1: Scatter plot of road accidents and population.

Table 2: Summary of Regression Coefficients

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>P</th>
<th>R-sq</th>
<th>R-Sq(adj)</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-263</td>
<td>1499</td>
<td>0.863</td>
<td>72.9%</td>
<td>71.3%</td>
<td></td>
</tr>
<tr>
<td>Pop (X1)</td>
<td>0.00053198</td>
<td>0.00007654</td>
<td>0.000</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Table 2 above reports the regression coefficients and other vital statistics, which are interpreted as follows:

The regression model, which establishes the relationship between total yearly accidents and population is thus given as $Y = -263 + 0.000532X$, where $Y$ is the total number of accidents in a year and $X$ being the total yearly population.

The value of -263 is interpreted in absolute terms to be the total number of yearly accidents when the total population is set to zero and all other factors are held constant, whilst the coefficient of $X$ of 0.000532 is the rate or magnitude of change in the number of accidents as a result of a change in the population. Its positive sign is an indication of the fact that there is a positive association between road traffic accidents and the population as already established by the scatter plot in the preliminary analysis above.

Again, the p-values indicate that the constant term is insignificant since it is greater than the chosen alpha level of 0.05, whilst that of the predictor variable ($X$) of 0.000 shows a highly significant even though it quite smaller, with approximately 0 standard error.

Furthermore, the coefficient of determination (R-sq) of 72.9% indicates that for the period under study based on the available data, population is able to account for 72.9% of the changes in accidents in the country with only 27.1% not being explained by population but rather by other variables which were not included in the study.

The adjusted coefficient of determination (R-Sq(adj)) of 71.3 indicates that the model specifically has an explanatory power of 71.3% specifically adjusted for the degrees of freedom with regard to the explanatory variable. It implies that population as an explanatory variable did not improve the model more than expected by chance.
Finally, variance inflation factor (VIF) of 1.00 implies the absence of correlation in the values of the predictor variable.

**Table 3: Analysis of Variance**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>51006785</td>
<td>51006785</td>
<td>48.30</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Error</td>
<td>18</td>
<td>19007715</td>
<td>1055984</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>70014500</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The analysis of variance table is used here to test the hypothesis for the study:

**Null Hypothesis (Ho):** There is no significant relationship between road traffic accidents and population growth in Ghana.

**Alternative Hypothesis (H1):** There is a significant relationship between road traffic accidents and population growth in Ghana.

The p-value of 0.000 as reported in the analysis of variance table indicates that the test is significant at 0.05. Therefore, we fail to accept the null hypothesis and conclude that indeed there is a significant relationship between road traffic accidents and the population in Ghana.

**3.2 Model Diagnostics**

The model diagnostics is done using graphical displays of the residuals as shown in figure 2 below.

![Residual Plots for Accidents(Y)](image)

- **Normal Probability Plot**
  - The probability plot of the residuals indicates that the individual probabilities of the residuals are normally distributed since the points are near or closer to one another with the plotted points forming an approximately straight line.

- **Versus Fits**
  - The plot of the residuals verses the fitted values clearly shows that the residuals appear to be randomly scattered about zero, thereby satisfying the assumption of constant variance.

- **Versus Order**
  - The histogram plot of the residuals also appears to be normal with only a slight deviation from normality which is allowed and so therefore satisfying the normality assumption.

From the Figure 2 above, it is clear that all the four (4) diagnostic residual plots do not show any anomalies for the fitted regression model. Clearly, the probability plot of the residuals indicates that the individual probabilities of the residuals are normally distributed since the points are near or closer to one another with the plotted points forming an approximately straight line.

The plot of the residuals versus the fitted values as well as time plot of the residuals clearly shows that the residuals appear to be randomly scattered about zero, thereby satisfying the assumption of constant variance. Finally, the histogram plot of the residuals also appears to be normal with only a slight deviation from normality which is allowed and so therefore satisfying the normality assumption.
4.0 Conclusion

Based on the analyses and interpretations of the appropriate data, findings from the results of the study are discussed as follows:

Firstly, the study revealed that there exists a strong positive correlation between road traffic accidents and population in Ghana, and that specifically, population growth accounts for 72.9% of the changes in the number of accidents in the country. This implies that as the population grows, the number of accidents will also increase.

Secondly, from the analysis the relationship was found to be statistically significant at 5% level of significance, with the exact regression model found as $Y = -263 + 0.000532X$. The test of hypothesis using analysis of variance concluded that the overall model was significant, but the constant was not significant. Therefore, the model can be used to estimate and forecast road traffic accidents in Ghana based on population.

Finally, the study also revealed that the model actually fits the data very well as it is clearly portrayed in the analysis of the residual plots in which the assumptions were satisfied.

REFERENCES


APPENDIX I:

**ROAD TRAFFIC AND POPULATION DATA FOR GHANA**

<table>
<thead>
<tr>
<th>Year</th>
<th>Accidents(Y)</th>
<th>Pop (X1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>8370</td>
<td>14821000</td>
</tr>
<tr>
<td>1992</td>
<td>6922</td>
<td>15222000</td>
</tr>
<tr>
<td>1993</td>
<td>6467</td>
<td>15634000</td>
</tr>
<tr>
<td>1994</td>
<td>6584</td>
<td>16056000</td>
</tr>
<tr>
<td>1995</td>
<td>8313</td>
<td>16491000</td>
</tr>
<tr>
<td>1996</td>
<td>8488</td>
<td>16937000</td>
</tr>
<tr>
<td>1997</td>
<td>9918</td>
<td>17295000</td>
</tr>
<tr>
<td>1998</td>
<td>10996</td>
<td>17665000</td>
</tr>
<tr>
<td>1999</td>
<td>8763</td>
<td>18349000</td>
</tr>
<tr>
<td>2000</td>
<td>11087</td>
<td>18845000</td>
</tr>
<tr>
<td>2001</td>
<td>11293</td>
<td>19328000</td>
</tr>
<tr>
<td>2002</td>
<td>10715</td>
<td>19811000</td>
</tr>
<tr>
<td>2003</td>
<td>10542</td>
<td>20506000</td>
</tr>
<tr>
<td>2004</td>
<td>12175</td>
<td>21093000</td>
</tr>
<tr>
<td>2005</td>
<td>11320</td>
<td>21693000</td>
</tr>
<tr>
<td>2006</td>
<td>11668</td>
<td>22294000</td>
</tr>
<tr>
<td>2007</td>
<td>12038</td>
<td>22911000</td>
</tr>
<tr>
<td>2008</td>
<td>11214</td>
<td>23544000</td>
</tr>
<tr>
<td>2009</td>
<td>12299</td>
<td>24196000</td>
</tr>
<tr>
<td>2010</td>
<td>11506</td>
<td>24233000</td>
</tr>
</tbody>
</table>

*Source: NRSC, 2010*