An Investigation on the Rate of Crime in Sokoto State Using Principal Component Analysis

*U. Usman, M. Yakubu and A.Z. Bello
Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria

*Corresponding Author; E-mail: uusman07@gmail.com; ☎: +234(0)7033346191

| ABSTRACT: | Principal component analysis is a data analysis tool that is used to reduce the dimensionality of a large number of interrelated variables, while retaining as much of the information as possible. In this paper, PCA technique has been applied to know the number of principal components to be retained on the seven variables obtained from Criminal Investigation Department Sokoto State Police Headquarters Sokoto. Data analysis was carried out using NCSS and GESS 2007 Software. From the results, three principal components have been retained using the Scree plot and Loading plot indicating that correlation exist between crimes against persons and crime against properties. |
| Keyword: | Principal Components Analysis (PCA), Crime, Murder, Assault, Robbery, Theft, Store breaking, False, Grievous Harm and Wounding (GHW) |

INTRODUCTION
There is no universal definition of crime. This is as a result of changes in social, political, psychological and economic conditions. An act may be a crime in one society, but not in another (Danbazau, 2007). For example, prostitution, adultery and homosexuality between consenting adults have been wholly or partially removed from the criminal law in USA (Feldman, 1997) but are considered as crimes in Muslim communities such as Saudi Arabia. The constant changes in time also change the perception of society on crime. Today, it is becoming a crime to pollute the air and water. Pollution causes few problems and receives little attention in colonial days (WBE, 1974). Therefore, the perception of an “act” to be a crime varies with time and space. In addition, many scholars have defined crime in different views, mostly bordering on ethical and ideological orientation, on the definition of crime, a United Nations Research Institute (UN, 1995).

Perceptions of crime are not determined by any objective indicator of the degree of injury or damage but by cultural values and power relations (UN, 1995). In a strict legal definition, however, a crime is a violation of criminal law which in most societies can be defined broadly as any ‘act or omission forbidden law on pain of punishment’ (Carvell and Swinfen, 1970).

One of the fundamental techniques to combat criminal activities is the better understanding of the dynamics of crime. Crime is often thought of as a moral threat and injurious to the society. It afflicts the personality of individual and his property and lessens trust among members of the society (Louis et al., 1981). However, crime is an inescapable reality in human life, therefore no national characteristics, no political regime, no system of law, police or justice have rendered a country exempt from crime (Radzinowicz and King, 1977). The causes of crime are multiple and could be traced to bio-genetic factors, such as genetic mutation and heredity (Horton, 1939), psychological factors, such as personality disorders (Abramsom, 1944) and sociological factors, such as learning and environment (Sutherlands, 1939). The diverse differences in geographical areas in terms of population density, demographic characteristics, natural vegetation, location and socio-economic characteristics has rendered crime rate unevenly distributed globally. However, it has been observed that the entire world is experiencing high criminal rate. The report of international crime victim survey (ICVS) has confirmed the situation. The report which was conducted on six major world region including Africa, Asia, central and eastern Europe, Latin America, and western Europe for the 1989 – 1996 period as shown that more than half of the urban respondents reported being victim at least once regardless of what part of the world they inhabit (Ackermen and Murray, 2004).

Over the years the rate of crime in Nigeria has been on the increase and these crimes are being carried out with more perfect and sophistication. This has led to the formation of various vigilante groups, to combat crimes in some parts of the country (Fajemirokun et al., 2006).

In this paper we use Principal Component Analysis (PCA) in determining the numbers of principal components (PC) to be used in explaining the crime data in Sokoto State.
Classification of Crime
The classification of crime differs from one country to another. In the United States, the Federal Bureau of Investigation tabulates the annual crime data as Uniform Crime Reports (UCR). They classify violations of laws which derive from common law as part 1 (index) crimes in UCR data, further categorised as violent as property crimes. Part 1 violent crimes include murder and criminal homicide (voluntary manslaughter), forcible rape, aggravated assault, and robbery; while part 1 property crimes include burglary, arson, larceny/theft, and motor vehicle theft. All other crimes count as part II crimes (Wiki/Cr., 2009).

In Nigeria, the Police classification of crime also depends on what law prescribed. In Nigeria Police Abstract of Statistics (NPACS), offences are categorised into four main categories:

i. Offences against persons includes: manslaughter, murder and attempted murder, assault, rape, child stealing, grievous hurt and wounding, etc.
ii. Offences against property includes: armed robbery, house and store breakings, forgery, theft/thealing, etc.
iii. Offences against lawful authority include: forgery of current notes, gambling, breach of peace, bribery and corruption, etc.
iv. Offences against local act include: traffic offences, liquor offences, etc.

Causes of Crimes
Criminal behaviour cannot be explained by a single factor, because human behaviour is a complex interaction between genetic, environmental, social, psychological and cultural factor. Different types of crimes are being committed by different types of people, at different times, in different places, and under different circumstances (Danbazau, 2007). Here we discuss some of the causes of crime:

i. **Biogenetic factors**: Criminologists are with the opinion that criminal activity is due to the effect of biologically caused or inherited factors (Pratt and Cullen, 2000). According to Lombroso (1911), a criminal is born, not made; that criminals were the products of a genetic constitution unlike that found in the non-criminal population.

ii. **Social and environmental factor** (Sutherland, 1939): The environment is said to play significant role in determining criminal behaviour. Factors within the environment that mostly influence criminal behaviour include poverty, employment, corruption, urbanisation, family, moral decadence, poor education, technology, child abuse, drug trafficking and abuse, architectural or environmental design Oyebanji (1982) and Akpan (2002) have attribute the current crime problem. In Nigeria to urbanisation, industrialisation and lack of education. Kutigi (2008) has said that the factors of crime in Nigeria and poverty and ignorance which are at the same time the opinion of many Nigerians (Azaburke, 2007). In another dimension, according to Ayoola (2008), lack of integrity, transparency and accountability in the management of public funds, especially at all levels of government have been identified as the factors responsible for the endemic corruption that has eaten deep into the fabric of the Nigerian society over the years.

The Nigerian Police
The most important aspect of criminal justice system is the police. Criminal justice system can be defined as a procedure of processing the person accused of committing crime from arrest to the final disposal of the case (Danbazau, 2007). However, for the past three decades there have been serious dissatisfaction and public criticisms over the conduct of the police (Danbazau, 2007). Then, what are the causes of the police failure in preventing and controlling the crimes? So many factors can be attributed to the problem. There are the issue of inadequate manpower, equipment and professionalism (Danbazau, 2007), corruption (Al-Ghazali, 2004) and poor public perception on the Nigeria Police (Okeroko, 1993), which has consequently made the Nigerian Public unwilling to corporate with the police in crime prevention and control.

Statistics of Crimes in Nigeria
Nigeria has one of the highest crime rates in the world. Murder often accompanies minor burglaries. Rich Nigerians live in high – security compounds. Police in some states are empowered to “shoot on sight” violent criminals (Financial Times, 2009).

In the 1980s, serious crime grew to nearly epidemic proportions, particularly in Lagos and other urbanized areas characterised by rapid growth and change, by stark economic inequality and deprivation, by social disorganisation, and by inadequate government service and law enforcement capabilities (Nigeria, 1991).

Annual crime rates fluctuated at around 200 per 100,000 populations until the early 1960s and then
steadily increased to more than 300 per 100,000 by the mid-1970s. Available data from the 1980s indicated a continuing increase. Total reported crime rose from almost 211,000 in 1981 to between 330,000 and 355,000 during 1984 – 85. The British High Commission in Lagos cited more than 3000 cases of forgeries annually (Nigeria, 1991).

In the early 1990s, there was growing number of robberies from 1,937 in 1990 to 2,419 in 1996, and later the figure declined to 2,291 in 1999. Throughout the 1990s, assault and theft constituted the larger category of the crime. Generally, the crime data grow from 244,354 in 1991 to 289,156 in 1993 (Cleen, 1993) and continued to decline from 241,091 in 1994 to 167,492 in 1999 (Cleen, 2003). The number of crime slightly declined to 162,039 in 2006, a reduction of 8 percent from 2005 (Cleen, 2006).

Crime Analysis Using PCA
PCA is very useful in crime analysis because of its robustness in data reduction and in determining the overall criminality in a given geographical area. PCA is a data analysis tool that is usually used to reduce the dimensionality (number of variable) of a large number of interrelated variables while retaining as much of the information (variation) as possible. The computation of PCA reduced to an eigenvector – eigenvector problem. It is performed either on a correlation or a covariance matrix. If some group of measures constitutes the scores of the numerous variables, the researchers may wish to combine the score of the numerous variables into smaller number of super variables to form the group of the measures (Jolliffe, 2002).

This problem mostly happens in determining the relationship between socio-economic factors and crime incidences. PCA uses the correlation among the variables to develop a small set of components that empirically summarised the correlation among the variables.

In a study to examine the statistical relationship between crime and socio-economic status in Ottawa and Saskatoon, the PCA was employed to replace a set of variables with a smaller number of components, which are made up of inter-correlated variables representing as much of the original data set as possible (Exp, 2008).

Principal component analysis can also be used to determine the overall criminality. When the first eigenvector show approximately equal loadings on all variables then the first PC measures the overall crime rate. In Printcom (2003) for 1997 US crime data, the overall crime rate was determined from the first PC, and the same result was achieved by Hardle and Zdenek (2007) for the 1985 US crime data. The second PC which is interpreted as “type of crime component” has successively classified the seven crimes into violence and property crime.

MATERIALS AND METHODS
Data collection
The data used in this study was obtained from the Criminal Investigation Department, Sokoto State Police Headquarters, Sokoto from 2002-2009.

Principal Component Analysis
Principal Components Analysis, or PCA, is a data analysis tool that is usually used to reduce the dimensionality (number of variables) of a large number of interrelated variables, while retaining as much of the information (variation) as possible. PCA calculates an uncorrelated set of variables (factors or pc’s). These factors are ordered so that the first few retain most of the variation present in all of the original variables. Unlike its cousin Factor Analysis, PCA always yields the same solution from the same data (apart from arbitrary differences in the sign). The matrix of scores will be referred to as the matrix Y. The basic equation of PCA is, in matrix notation, given by:

\[ Y = W X \]  

\[ \text{(1)} \]

Where W is a matrix of coefficients that is determined by PCA.

The matrix of weights, W, is calculated from the variance-covariance matrix, S. This matrix is calculated using the formula:

\[ S_{ij} = \sum_{k=1}^{n} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \]

\[ \frac{1}{n-1} \]  

\[ \text{(2)} \]

The singular value decomposition of S provides the solution to the PCA problem. This may be defined as:

\[ USU^T = L \]

Where L is a diagonal matrix of the eigenvalues of S, and U is the matrix of eigenvectors of S. W is calculated from L and U, using the relationship:

\[ W = U L^{-\frac{1}{2}} \]

It is interesting to note that W is simply the eigenvector matrix U, scaled so that the variance of each factor, y_j, is one.

The correlation between an i th factor and the j th original variable may be computed using the formula:
\begin{equation}
    r_{ij} = \frac{u_{ji} \sqrt{c_l}}{s_{ji}}
\end{equation}

Here \( u_{ji} \) is an element of \( U \), \( l \) is a diagonal element of \( L \), and \( s_{ji} \) is a diagonal element of \( S \). The correlations are called the factor loadings and are provided in the Factor Loadings report.

When the correlation matrix, \( R \), is used instead of the covariance matrix, \( S \), the equation for \( Y \) must be modified. The new equation is:

\begin{equation}
    Y = WX + \epsilon
\end{equation}

Where \( D \) is a diagonal matrix made up of the diagonal elements of \( S \). In this case, the correlation formula may be simplified since the \( s_{jj} \) are equal to one.

Method for Determining Factors

Jolliffe (2002) suggests using a cutoff on the eigenvalue of 0.7 when correlation matrices are analyzed. Cattell (1966) documented the scree graph, which will be described later in this chapter. Studying this chart is probably the most popular method for determining the number of factors, but it is subjective, causing different people to analyze the same data with different results.

Principal Component Obtained From Covariance and Correlation Matrices

Consider the covariance matrix of a bivariate data

\begin{equation}
    \Sigma = \begin{pmatrix}
        \delta_{11} & \delta_{12} \\
        \delta_{21} & \delta_{22}
    \end{pmatrix}, \quad \text{Where } \delta_{12} = \delta_{21}
\end{equation}

and \( \delta_{22} > \delta_{11} \), assume \( \delta_{11} = 1 \) and \( \delta_{22} = 100 \) (a very large difference)

And the derived correlation matrix

\begin{equation}
    \rho = \begin{pmatrix}
        1 & r_{12} \\
        r_{21} & 1
    \end{pmatrix}
\end{equation}

Where \( r_{12} = r_{21} \). Because of its large variance, \( X_2 \) will completely dominate the first PC determined from \( \Sigma \). Moreover, this first PC explains a larger proportion of the total population variance as

\begin{equation}
    \frac{\lambda_1}{\lambda_1 + \lambda_2} = \psi_{x1}
\end{equation}

When the variables \( X_1 \) and \( X_2 \) are standardized, however, the resulting variables contributes equally to the PCs determine from \( \rho \) so that the first PC of the total population variance is explained as

\begin{equation}
    \frac{\lambda_1}{\lambda_1 + \lambda_2} = \psi_{z1}
\end{equation}

Thus \( \psi_{x1} \neq \psi_{z1} \), the proportion of variance accounted for by the component of \( \rho \) differs from the proportion for \( \Sigma \).

Since the entries of the covariance and correlation matrices are different, then the coefficient (eigenvectors) of the PC obtained from \( \rho \) differs from those obtained from \( \Sigma \), and therefore \( \alpha_j X \neq \alpha_j Z \); the PCs derived from \( \rho \) are different from those obtained from \( \Sigma \) (Rencher, 2002).

Procedure for Calculating PCs

For a random vector \( X = [X_1, X_2, ..., X_p] \) the corresponding standardized variables are \( Z = [Z_1, Z_2, ..., Z_p] \) so that \( \text{cov}(Z) = \rho \) (the correlation matrix of \( X \)). We denote the matrix of correlation between \( p \) variables by

\begin{equation}
    \rho = \begin{pmatrix}
        1 & r_{12} & \cdots & r_{1p} \\
        r_{21} & 1 & \cdots & r_{2p} \\
        \vdots & \vdots & \ddots & \vdots \\
        r_{p1} & r_{p2} & \cdots & 1
    \end{pmatrix}
\end{equation}

And the vector of the coefficient (weight or loadings) on the \( p \) variable for the \( j^{th} \) component by

\begin{equation}
    \alpha_j = \begin{pmatrix}
        \alpha_{j1} \\
        \alpha_{j2} \\
        \vdots \\
        \alpha_{jp}
    \end{pmatrix}
\end{equation}

The problem of determining the vector of \( \alpha_j \) which maximizes (1) the variance accounted for by the first component, (2) the variance accounted for by the second component, orthogonal to the first etc. the solution for \( \alpha_j \) can be solved by this equation

\begin{equation}
    (\rho - \lambda_j I)\alpha_j = 0
\end{equation}

In which \( I \) is the identity matrix, \( \lambda_j \)'s are the characteristic roots or eigenvalue of \( \rho \) and the \( \alpha_j \)'s are the associated eigenvectors.

RESULTS AND DISCUSSION

The major types of crime in the analysis are: crime against person which include: murder, grievous harm and wounding (GHW), assault and crime against
property which include armed robbery, theft and stealing, store breaking and false pretence and cheating.

The paper shows the comparison of the relative sizes of the mean in this data set. Theft and Stealing, GHW, false and murder have the highest rate. There is a high dispersion among the crimes which indicate an evenly distribution of crime in the state and make it necessary to classify them into category for reliable analysis. From Table 1 the communality predicted the variable retains factor kept as one.

Table 2 displays different levels of correlation between the crimes. There are very low correlation between crime against property and person which means that none of the variable can be used to predict (explain) another except in the case of store breaking. However, the correlation between crimes against person is at least moderate and can be fairly used to predict each other. The Gleason-Staelin redundancy measure, phi is 0.42 which is okay to some extent but care should be taken in using the phi when it is less than 0.5.

The eigenvalues are often used to determine how many factors are to be retained. One rule of thumb is to retain those factors whose eigenvalues are greater than one (Kresta, 1994). From Table 3 Considering the eigenvalues and cumulative percent, it will be reasonable to retain the first three PC’s that explain up to 89.40% of the total variability in the data set.
Figure 1 shows a rough bar plot of the eigenvalues. The scree plot shows that the first three factors are indeed the largest and we have the impression that these factors will adequately approximate this data and accept for over 89% of the variation.

From Table 4 the first PC combines the number of all the crimes with the classification that crime against person shows high rate of offence in the state, the second PC has high positive loading on assault of about 26% of the loading plot and there is a moderate correlation between the crimes which signifies low rate of robbery in the state.

Table 5 and Table 6 shows categorically that there is high rate of crime among persons with high moderate correlation and factor 1 is the average of all the variables while factor 3 is the contrast and factor 7 shows a very low correlation in the state.
Table 6: Bar Chart of Communalities

<table>
<thead>
<tr>
<th>Variables</th>
<th>Factor1</th>
<th>Factor2</th>
<th>Factor3</th>
<th>Factor4</th>
<th>Factor5</th>
<th>Factor6</th>
<th>Factor7</th>
<th>Communality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assault</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robbery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theft_Stealing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Store_Breaking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3: Factor Loading
Figure 3(a-d) shows there are the presence of 6 outliers in the plots and e shows positive and linear dependencies on the third factor. This means that there exist some interactions between the three Principal Components and there is also a significance difference between each crime so that attention should be given to crime against person in the state.

CONCLUSION
In this paper, we applied Principal Component Analysis to explore the number of principal components to be retained on the seven variables obtained from Criminal Investigation Department Sokoto State Police Headquarters Sokoto. The results of the statistical analysis proved the three components explain up to 89.40% (Table 3) of the total variability of the data set which gives evidence that Sokoto State have a low criminality record. The larger number of PC obtained could possibly be attributed to the problem of the Nigerian Police Data Collection.

We found that the highest and commonly committed crimes in Sokoto State among crime against person are Assault and GHW, while Store breaking is the highest committed crime against property.

There are low correlation in between crimes against person and property and therefore cannot be used to explain one another, however, at least moderate correlation exist in between crimes against person except murder with low rate. The second component has classified the crime into two categories with respect to the rate of occurrence using the loading plot (i) Assault, GHW and Store breaking (ii) Theft and Stealing, false pretence and Stealing, Murder and Robbery which shows a slight difference from the original classification.

REFERENCE


