Available online at www.elixirpublishers.com (Elixir International Journal)

Statistics

Elixir Statistics 61 (2013) 16808-16812

ISSN: 2229-712X

On the suitability of probability density functions in modelling Nigeria crime data: a pearson system approach

Osowole, O. I[1,*], Aghamie S.O[2] and Anderson, N[3]

[1]Department of Statistics, University of Ibadan, Nigeria.

[2]Department of Mathematics, Agbor, Delta State, Nigeria

[3]Department of Statistics, Abia State Polytechnic, Aba, Nigeria.

ARTICLE INFO

Article history:

Received: 6 June 2013;

Received in revised form: 24 July 2013;

Accepted: 1 August 2013;

Keywords

Pearson system of distribution approach,

Four parameter gamma density function,

Crime prevention and control,

Mental stress,

Morality of society.

ABSTRACT

This paper attempted to fit an appropriate probability density function to available crime data in Nigeria by using the Pearson System of Distribution Approach. Rising crime rates have necessitated the need to model adequately crime occurrences to enhance effective management and control. Aggregated dataset of crime rates from 1994 to 2003 were used to estimate the parameters of the selected density functions from the Pearson system and for the goodness of fit test. Results showed that the Four Parameter Gamma density function was the most suitable for the data under study. This study has revealed the potentials of probability density functions in crime prevention and control.

Introduction

Crime started in the primitive days as a simple and less organised issue, and ended today as very complex and organised. Therefore, the existence of crime and its problems have spanned the history of mankind (Gulumbe *et. al.* 2012). Generally crime is treated as the unexpected behaviour of an individual which goes against the law. There are many reasons due to which an individual produces this behaviour. Sometimes crime is committed by a person because of mental stress and sometimes crime is committed without any reason because some people are habitual to do so. There is no exact definition of crime; it depends on different time period and different regions. According to the Cruzen, "A crime is an act of human conduct harmful to others which the state is bound to prevent. It renders the deviant person liable to punishment as a result of proceeding initiated by the state organs assigned to ascertain the nature, the extent and the legal consequences of the person's wrongness (Aoulak, M. A. (1999)). Schaefer (1989) referred to crime as a violation of criminal law which its formal penalties are applied by some governmental authority. Crime according to Dan Bazau (1994) is something which offends the morality of society, or that violates the divine law.

Research on the geographic distribution and the determinants of urban crime has long been an important area of interest for criminologists, sociologists, and geographers. Criminologists and sociologists believe that crime results from social stress and conflicts and the rates of crime in urban neighbourhood are highly affected by the demographic and socio-economic contexts (Haifeng Zhang and Michael P. Peterson (2007) and Reith, M. (1996)).

The growth in urban crime rate in Nigeria is one of the major social problems facing the country in recent time. The dominance of crime in developing countries increases the volatility of the issue, for it pyramids one fear upon others. The concentration of violent crimes in major urban cities worldwide is therefore heralded as an indicator of the breakdown of urban systems. In many urban cities of Nigeria today, criminal activities and violence are assuming dangerous tendencies as they threaten lives and property, the national sense of well-being and coherence, peace, social order and security, thus, reducing the citizens" quality of life (Agbola, (2000); Ahmed (2010); and Ahmed, Y. A.(2012)). The fear of armed robbery keeps Nigerians sleepless at night and they tend to live one day at a time with the fear of whether they will see the light of tomorrow. They are especially afraid of armed-robbers, paid assassins, political thugs and other criminals who assess life as being worthless. Nigerians find it difficult to put their trust on police protection because Nigeria is under policed with an average of one policeman to 5000 Nigerians, compared to that of one policeman to 400 persons in the developed world. Nigerian police are, at times, in collusion with the men of the underworld to unleash terror on their fellow countrymen (Agbola, (1997) and Ahmed, Y.A. (2012)).

No systematic study on the suitability of probability density functions in modelling crime incidence in Nigeria has been undertaken in the past. Most past studies on crime rates in Nigeria have centred more on the social and economic aspects of crime and poor management of urban centres. Authors who have contributed in this regard include Oyebanji (1982); Omisakin (1998); Alemika (2003); Agbola (2002); Omotor (2009) and Ahmed (2010). This study is therefore an attempt to determine

Tele:
E-mail addresses: dosowole@yahoo.com

an appropriate probability density function from the Pearson system of distributions that will fit well crime data in Nigeria following the work of Osowole and Bamiduro (2012). It is expected that the determined probability density function will be useful in future forecast of crime occurrence and control.

2.0 The Pearson System of Distributions

Several well known distributions like Gaussian, Gamma, Beta and Student's t-distributions belong to the Pearson family. The system was introduced by Pearson, K.(1985) who worked out a set of four-parameter probability density functions as solutions to the differential equation

$$\frac{f'(x)}{f(x)}=\frac{P(x)}{Q(x)}=\frac{x-a}{b_0+b_1x+b_2x^2} \quad \ldots\ldots(1)$$

where f is a density function and a,b_0,b_1 and b_2 are the parameters of the distribution. What makes the Pearson's four-parameter system particularly appealing is the direct correspondence between the parameters and the central moments $(\mu_1,\mu_2,\mu_3\ and\,\mu_4)$ of the distribution (Stuart, A. And Ord. J.(1994)). The parameters are defined as

$$\left.\begin{aligned} b_1=a&=-\frac{\mu_3(\mu_4+3\mu_2^2)}{A}\\ b_0&=-\frac{\mu_2(4\mu_2\mu_4-3\mu_3^2)}{A}\\ b_2&=-\frac{(2\mu_2\mu_4-3\mu_3^2-6\mu_2^3)}{A}\end{aligned}\right\} \quad \ldots\ldots(2)$$

The scaling parameter A is obtained from

$$A=10\mu_4\mu_2-18\mu_2^3-12\mu_3^2 \quad \ldots\ldots(3)$$

When the theoretical central moments are replaced by their sample estimates, the above equations define the moment estimators for the Pearson parameters a,b_0,b_1 and b_2.As alternatives to the basic four-parameter systems, various extensions have been proposed with the use of higher-order polynomials or restrictions on the parameters. Typical extension modifies (1) by setting $P(x) = a_O + a_1x$ so that

$$\frac{f'(x)}{f(x)}=\frac{P(x)}{Q(x)}=\frac{a_0+a_1x}{b_0+b_1x+b_2x^2} \quad \ldots\ldots(4)$$

This parameterization characterizes the same distributions but has the advantage that a_1 can be zero and the values of the parameters are bound when the fourth cumulant exists (Karvanen, J. (2002)). Several attempts to parameterize the model using cubic curves have been made already by Pearson and others, but these systems proved too cumbersome for general use. Instead the simpler scheme with linear numerator and quadratic denominator are more acceptable.

2.1 Classification and Selection of Distributions in the Pearson System

There are different ways to classify the distributions generated by the roots of the polynomials in (1) and (4). Pearson himself organized the solution to his equation in a system of twelve classes identified by a number.

Table 1: Pearson Distributions

The table provides a classification of the Pearson Distributions, f(x) satisfying the differential equation $(\frac{1}{f})\frac{df}{dx}=\frac{P(x)}{Q(x)}=\frac{(a_0+a_1x)}{(b_0+b_1x+b_2x^2)}$ **. The signs and values for selection criteria,** $D=b_0b_2-b_1^2$ **and** $\lambda=\frac{b_1^2}{b_0b_2}$ **, are given in columns three and four.**

	Restrictions	D	λ	Support	Density
P(x) = a_0, Q(x) = 1					
1.	**$a_0 < 0$**	**0**	**0/0**	$\mathbf{R}^+$	$\gamma e^{-\gamma x},\gamma>0$
P(x) = a_0, $Q(x)=b_2x(x+\alpha)$					
2(a).	**α > 0**	**< 0**	∞	**[- α , 0]**	$\frac{m+1}{\alpha^{m+1}}(x+\alpha)^m, m<-1$
2(b).	**α > 0**	**< 0**	∞	**[- α , 0]**	$\frac{m+1}{\alpha^{m+1}}(x+\alpha)^m, -1<m<0$
P(x) = a_0, $Q(x)=b_0+2b_1x+x^2=(x-\alpha)(x-\beta),\alpha<\beta$					
3(a).	$a_0\neq 0$, $0<\alpha<\beta$	**< 0**	**>1**	**[β, ∞]**	$\frac{(\beta-\alpha)^{-(m+n+1)}}{B(-m-n-1,n+1)}(x-\alpha)^m(x-\beta)^n$ $m>-1,n>-1,m\neq 0,n\neq 0,m=-n$
3(b).	$a_0\neq 0$, $\alpha<\beta<0$	**< 0**	**>1**	**[-∞,α]**	$\frac{(\beta-\alpha)^{-(m+n+1)}}{B(-m-n-1,m+1)}(x-\alpha)^m(x-\beta)^n$ $m>-1,n>-1,m\neq 0,n\neq 0,m=-n$
4.	$a_0\neq 0$, $\alpha<0<\beta$	**< 0**	**< 0**	**[α, β]**	$\frac{\alpha^{2m}\beta^{2n}}{(\alpha+\beta)^{m+n+1}B(m+1,n+1)}(x-\alpha)^m(x-\beta)^n$ $m>-1,n>-1,m\neq 0,n\neq 0,m=-n$
$P(x)=a_0+a_1x, Q(x)=1$					
5.	$a_1\neq 0$	**0**	**0/0**	**R**	$\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Table 2: Pearson Distributions (Continued)

	Restrictions	D	λ	Support	Density
$P(x)=a_0+a_1x, Q(x)=x-\alpha$					
6.	$a_1 \neq 0$	< 0	∞	[α,∞]	$\frac{k^{m+1}}{\Gamma(m+1)}(x-\alpha)^{-m}e^{-k(x-\alpha)}, k>0$
$P(x)=a_0+a_1x, Q(x)=b_0+2b_1x+x^2=(x-\alpha)(x-\beta), \alpha\neq\beta$					
7(a).	$a_1 \neq 0$ $0<\alpha<\beta$	< 0	>1	[β, ∞]	$\frac{(\beta-\alpha)^{-(m+n+1)}}{B(-m-n-1,n+1)}(x-\alpha)^m(x-\beta)^n$ $m>-1, n>-1, m\neq 0, n\neq 0, m\neq -n$
7(b).	$a_1 \neq 0$ $\alpha<\beta<0$	< 0	>1	[-∞,α]	$\frac{(\beta-\alpha)^{-(m+n+1)}}{B(-m-n-1,m+1)}(x-\alpha)^m(x-\beta)^n$ $m>-1, n>-1, m\neq 0, n\neq 0, m\neq -n$
8.	$a_1 \neq 0$ $\alpha<0<\beta$	< 0	< 0	[α, β]	$\frac{\alpha^{2m}\beta^{2n}}{(\alpha+\beta)^{m+n+1}B(m+1,n+1)}(x-\alpha)^m(x-\beta)^n$ $m>-1, n>-1, m\neq 0, n\neq 0, m\neq -n$
	$P(x)=a_0+a_1x, Q(x)=b_0+2b_1x+x^2=(x-\alpha)(x-\beta), \alpha=\beta$				
9.	$a_1>0$ $\alpha=\beta$	0	1	[α, ∞]	$\frac{\gamma^{m-1}}{\Gamma(m-1)}(x-\alpha)^{-m}e^{-\gamma/x}, \gamma>0, m>1$
	$P(x)=a_0+a_1x, Q(x)=b_0+2b_1x+x^2$, complex roots				
	Restrictions	D	λ	Support	Density
10.	$a_0=0, a_1<0$ $b_1=0, b_0=\beta^2$ $\beta\neq 0$	>0	0	**R**	$\frac{\alpha^{2m-1}}{B(m-\frac{1}{2},\frac{1}{2})}(x^2+\beta^2)^{-m}, m>\frac{1}{2}$
11.	$a_0\neq 0, a_1<0$ $b_1\neq\frac{a_0}{a_1}$	>0	0> <1	**R**	$c(b_0+2b_1x+x^2)^{-m}e^{-v\arctan\left((x+b_1)/\beta\right)}$ $m>\frac{1}{2}, \beta=\sqrt{b_0-b_1^2}$

The numbering criterion has no systematic basis and it has varied depending on the source. An alternative approach suggested by Andreev, A. *et. al.*(2005) for distribution selection based on two statistics that are functions of the four Pearson parameters will be adopted. The scheme is presented in Tables 1 and 2 where *D* and λ denote the selection criteria. D and λ are defined as

$$\left.\begin{array}{l} D=b_0b_2-b_1^2 \\ \lambda=\frac{b_1^2}{b_0b_2} \end{array}\right\} \qquad \text{……………………………………(5)}$$

The advantage of this approach in statistical modelling in the Pearson framework is its simplicity. Implementation is done in accordance with the following steps:

(1) Estimate moments from data.

(2) Calculate the Pearson parameters a, b_0, b_1 and b_2 using (2) and (3).

(3) Use the estimates of the parameters to compute the selection criteria D and λ as given in (5)

(4) Select an appropriate distribution from Tables (1) and (2) based on the signs of the values of the selection criteria

3.0 Results and Discussion

The methods presented are applied to the disaggregated data on different crimes committed in Nigeria between 1994 and 2003. The data were compiled by the Research Department of CLEEN Foundation and downloaded from the website of the foundation. The data consisted of offences against persons: manslaughter, murder and attempted murder, assault, rape, child stealing, grievous hurt and wounding; offences against property: armed robbery, house and store breakings, forgery, and theft/stealing; offences against lawful authority include: forgery of current notes, gambling, breach of public peace, bribery and corruption. The data were aggregated and transformed appropriately to conform to the adopted methodology of the study. The total number all cases of crime considered was 260.

The estimates of the selection criteria for the selection of probability distributions from the Pearson system were obtained as shown in Table 3. Based on the values and signs of these criteria, the beta and gamma distributions (rows (3a) and (6) in Table 3) were the possible candidates to be selected for the crime data based on the classifications in Tables 1 and 2. However a closer look at the values of the selection criteria indicated the selection of the three parameter gamma distribution over the four parameter beta distribution. The estimates of the parameters of these distributions (Table 4) were obtained using the Easy Fit 5.5 statistical software specially designed for this purpose. Easy fit allows three different specifications of the gamma distribution. The three forms were considered despite the fact that the Pearson table earlier suggested the three parameter case. In order to further validate the choice made from the Pearson table of distributions, a Kolmogorov Smirnov goodness of fit test was also conducted on the data. The test showed that both the 4-parameter and 2-parameter gamma distributions fitted well the data under consideration at the 1% level of significance. The 3-parameter gamma and 4-parameter beta distributions did not perform well under this setting. However the 4-parameter gamma distribution performed better than its 2-parameter counterpart because it had a higher p-value. This is further collaborated by the cumulative distribution function graphs in Figures 1 and 2. From the graphs it can be seen that the 4-parameter performed better than the 2-parameter.

Table 3: Estimates of Selection Criteria ($D = b_0 b_2 - b_1^2$ and $\lambda = \frac{b_1^2}{b_0 b_2}$) for Crime Data

	Estimate
A	-4.01303 X10^{23}
b_0	-820.7813
b_1= a	-473.148
b_2	-0.25
D	-2.223663554x10^5
λ	1090.985

Table 4: Parameter Estimates of the Gamma and Beta Distributions

Generalized Gamma (4P)	**Generalized Gamma (3P)**	**Gamma (2P)**	**Beta (4P)**
k =0.83397	α =0.74792	α =0.7408	α_1 =0.9437
α =0.9437	β=64.006	β=66.702	α_2=45.668
β=45.668	γ= 2.0		a= 2.0
γ= 2.0			b =0.9437

P=number of parameters

Table 5: Kolmogorov Smirnov Goodness of Fit Test (α =0.01)

	Generalized Gamma (4P)	**Generalized Gamma (3P)**	**Gamma (2P)**	**Beta (4P)**
Test Statistic	0.08615	0.10889	0.09261	0.10622
p-value	0.03967	0.00386	0.02163	0.00522
Remark	Sig.	NSig.	Sig.	NSig.

Sig. = significant, NSig.= not significant, P=number of parameters

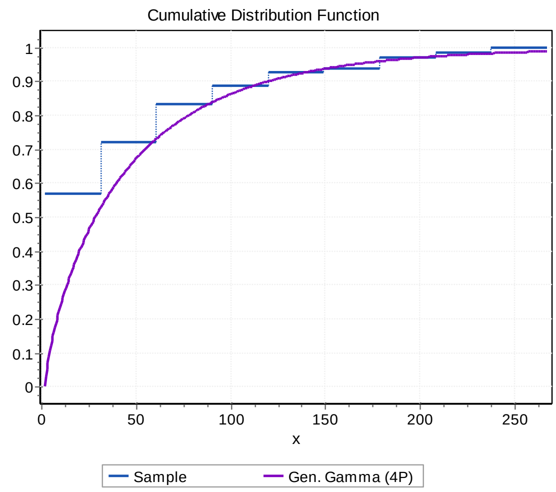

Figure 1: Cumulative Distribution Graph for Gamma (4P)

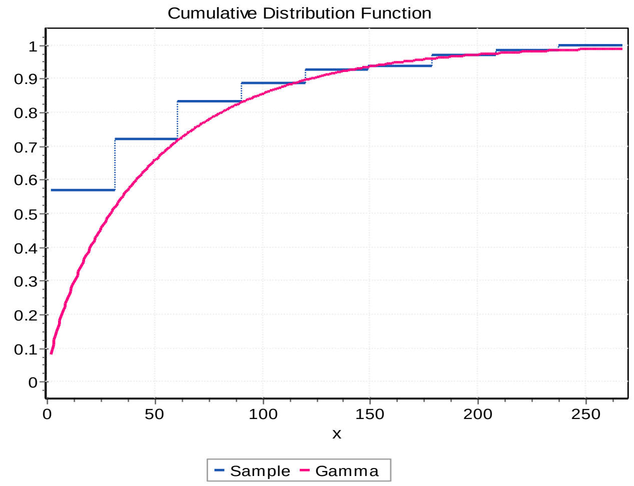

Figure 2: Cumulative Distribution unction Graph for Gamma (2P)

4.0. Conclusion

This study has derived the most appropriate probability distribution function for the modelling of crime occurrences in Nigeria. Two specifications of the gamma distributions were found suitable. The 4-parameter case gave a higher p-value at 1% level of significance from the Kolmogorov Smirnov goodness of fit test and was therefore considered to be more appropriate. It is hoped that the findings of this study will help all stakeholders involved in crime prevention and control.

References

1. Agbola, T. (1997): 'The architecture of fear" Urban Design and Construction, Response to Urban Violence in Lagos, Nigeria. IFRA/African Book Builders.

2. ________ (2000): 'Urban Violence, Urban Insecurity and the Challenge of Good Urban Governance; The evolving Disturbing Scenario from Abuja, Nigeria' In the Proceedings of the 33rd Annual Conference of NITP.Ilorin.30th October-1stNovember.

3. __________ (2002); A Study of Urban Violence and Insecurity in Abuja, Journal of the Nigerian Institute of Town Planners (JNITP) 1(XV): 59-78.

4. Alemika, E. (2003): 'Nigeria recorded 10,345 robbery cases in four years' The Punch, News Paper, Monday 7th July, 2003.

5. Andreev, A., Kanto, A., and Malo, P.(2005): 'Simple Approach for Distribution Selection in the Pearson System' Helsinki School of Economics Working Papers: W-388.

6. Ahmed, Y.A. (2010): 'Trend and Pattern of Urban Crime in South-western Nigeria' Unpublished Ph.D. Thesis, University of Ilorin, Nigeria

7. Ahmed, Y.A.(2012): 'The Pattern and Distribution of Crime Incidence in an Urban Environment: A Case Study of Osun State South Western Nigeria' International Journal of Humanities and Social Science Vol.2.No5, 2012

8. Aoulak, M. A. (1999): 'Prison administration in Pakistan' Lahore : SnS Publishers.

9. Dan Bazau, A.B.(1994): 'Law and Criminality in Nigeria: An Analytical Discourse' University Press Plc, Ibadan, Oyo State, Nigeria

10. Gulumbe, S.U, Dikko, H.G., and Yusuf, B. (2012): 'Analysis of Crime Data using Principal Component Analysis: A Case Study of Katsina State' CBN Journal of Applied Statistics Vol. 3 No.2.

11. Haifeng Zhang and Michael P. Peterson (2007): 'A Spatial Analysis of Neighbourhood Crime in Omaha, Nebraska using

Alternative Measures of Crime Rates' Internet Journal of Criminology(www.internetjournalofcriminology.com)
12. Karvanen, J.2002. Adaptive Methods for Score Function Modelling in Blind Source Separation. Unpublished Ph.D. Thesis, Helsinki University of Technology
13. Omisakin, I.S. (1998): 'Crime Trends and Prevention Strategies in Nigeria' A Study of Old Oyo State, Monograph Series No.9, NISER Ibadan.
14. Omotor, D. G. (2009): 'Socio-Economic Determinants of Crime in Nigeria' Pakistan Journal of Social Sciences 6(2):54-59.
15. Osowole, O.I and Bamiduro, T.A.(2012)): 'On the Probability Density Functions of Forster-Greer-Thorbecke (FGT) Poverty Indices' West African Journal of Industrial and Academic Research, WAJIAR 4(1)Pp 9-15.ISSN 2276-9129.www.wajiaredu.com
16. Oyebanji, J. O. (1982): 'Economic Development and the Geography of Crime: An Empirical Analysis' Geo Journal, (NITP) 6 (5): 453-458.
17. Pearson, K. (1895): 'Memoir on Skew Variation in Homogeneous Material' Philosophical Transactions of the Royal Society. A186: 323-414.
18. Reith, M. (1996): 'The Relationship between Unemployment in Society and the Popularity of Crime Drama on TV' Journal of Broadcasting & Electronic Media, 40: 258-264.
19. Schaefer, R.T. (1989): 'Sociology' Third Edition Western Illinois University in Collaboration with Robert P. Lamm McGraw, Hill, Inc.
20. Stuart, A. and Ord. J.(1994): 'Kendall's Advanced Theory of Statistics Vol. I: Distribution Theory' London: Edward Arnold.