Kragujevac J. Math. 22 (2000) 21–31.

STATISTICAL MODELS FOR ANALYSING CRIMINAL CAREERS

M. O. Ojo

Department of Mathematics, Obafemi Awolowo University, Ife-Ife, Nigeria

(Received June 17, 1999)

ABSTRACT. The main purpose of this paper is to present and discuss the analytical tools that will be necessary to carry forward research in criminal careers. In an earlier model that was used to analyses some crime data (Ojo [4]), the rate of offending or arrest γ, say, of individual criminals, was taken as a constant, which can be regarded as over simplification. This model and other similar basic models in this context are extended in this paper, first, to reflect variations in γ across the explanatory variables, x (such as age, drug use, previous convictions, etc.), then to reflect variations in γ over time within a criminal career and finally, to take into account both changes in γ over time and variations across x.

1. Introduction

The high rate of crime is a common problem all over the world. A lot of research has been done and there continues to be more efforts in either finding ways of controlling crime or developing methods of analysing crime data. There are several variables that can explain variations in criminology. Such variables include age, drug use, stability in job, Psychiatric History and many others. Many other time dependent factors also play important roles in criminology. For example the rate of committing crime in Nigeria was on the increase before the civil war that broke out in 1967. During this civil war, attention was diverted from crime as able – boided men were drafted into the army and the crime rate automatically dropped. At the end of the war many became jobless but shortly after was an era of oil boom. Contacts were awarded here and there and many contractors and business men became very rich and consequently crime rate started to increase again.

Apart from high rate of crime, the rate of offending or arrest or conviction varies from criminal to criminal and from one type of offenese to another.

Because of these reasons and some others, there has ever been a need to carry out research in criminology. Comprehensive data are needed for effective research in criminology. Such data must include the behavior of each individual criminal over a sufficiently long period of time. These types of data can be obtained from either prison, police records or psychiatric hospitals in some countries.

Unfortunately, these data are often not made available to researchers in Nigeria. Indeed, more often than not these information are not kept, and where they are kept, they are quite scanty. This has drastically slowed down the pace of statistical research in quantitative criminology in Nigeria.

However, statistical models can still be developed and made ready to be used when are available.

Thus, the main purpose of this paper is to present and discuss analytical tools needed to carry forward research in criminal careers.

An important parameter in analysing criminal careers is the rate at which offenders commit offense or simply the rate of offending or arrest or conviction which is denoted by γ in this paper. Most of the earlier models assume that this rate is constant across criminals (for example see Cohen [2]). Subsequent research has shown that there is a considerable heterogeneity in the values of γ both between individual offenders and for the same offender at different points during his criminal career. Research such as the survey of inmates conducted by Rand (Paterson and Braiker [5], Chaiken and Chaiken [1]) indicates that γ varies considerably across offenders and its distribution is Skewed.

It is usually assumed that the rate γ starts from the maximum and decreases with time to zero, marking the end of the career. However, it is strongly felt in this paper that the rate γ of offending may not necessarily start at the maximum particularly if a criminal starts committing crime at a tender age. These criminals usually start committing crime by pick-pocketing, shop-lifting or even raping. As they grow in age, they continue to gain more experience and skill in their chosen career. They continue in this way until they reach the peak of their career after which the rate of committing crime starts to decrease (probably because of age, inability to perform, lack of interest or Soberness) until the career eventually terminates. The models for the rate γ of offending that will meet the above requirements must satisfy the following basic condition. If $\gamma(t)$ denotes the rate at which an offense is committed at time t, then (i) $\gamma(0) = 0$ and (ii) $\gamma(t) = 0$ for sufficiently large value of t.

In addition to the fact that the distribution of γ is Skewed, models that satisfy the above conditions are presented in sections 4 and 5.

The number of crimes committed in a given interval of time and the times at which these crimes are committed may be considered as random processes and this concept is discussed in section 2.

In section 3, we mention a few number of models that can be used for the purpose of prediction in the context of criminology. Model fitting is discussed in section 6.

The attendant problem of getting relevant and reliable data from sources like prisons or police records, in Nigeria, has made a full illustration impossible in this paper. However, a few number of models presented in section 5 are fitted to the crime data obtained from a local prison.

2. Model as a Markov Process

We assume that within a period of time or within an individual's criminal career, the number of crimes committed and the exact times at which these crime are committed are determined by a chance process.

Now let $X(t)$ be the number of crimes committed at time $t > 0$.

Thus $X(t, t{+}h) - X(t)$ is the number of crimes committed in the interval $(t, t{+}h)$.

The process can be regarded as a Markov process and is specified by the following basic assumptions.

(i) Pr (One crime in the interval $(t, t+h)$)$= \lambda h$.

$$ \tag{2.1} $$

(ii) Pr (Number of crimes > 1 in the interval $(t, t+h)) = o(h)$ where λ is the crime rate and h is small and positive.

We are interested in finding the probability

$$\begin{aligned} &P(X(t) = j | X(0) = 0) \\ &\qquad\qquad\qquad = P_{0j}(t) \end{aligned} \tag{2.2}$$

Obviously,

$$P_{0j}(X(t)) = 0 \quad \text{if} \quad 0 > j.$$

Let us consider two contiguous intervals $(0, t)$ and $(t, t + h)$.

If at last one crime is committed during the interval $(0, t+h)$, we must have one of the following three mutually exclusive events:

(i) No crime during $(t, t + h)$ and j crimes during $(0, t)$

(ii) 1 crime during $(t, t + h)$ and $j - 1$ crimes during $(0, t)$.

$$ \tag{2.3} $$

(iii) > 1 crimes during $(t, t + h)$ and $< j - 1$ crimes during $(0, t)$.

Combining (2.1) and (2.3) we get the following difference equation

$$P_{0j}(t + h) = (1 - \lambda h)P_{0j}(t) + \lambda h P_{0,j-1}(t) + o(h).$$

That is

$$P_{0j}(t + h) - P_{0j}(t) = -\lambda h P_{0j}(t) + \lambda h P_{0,j-1}(t) + o(h).$$

Dividing through by h and taking limit as $h \to 0$, we get

$$P'_{0j}(t) = -\lambda P_{0j}(t) + \lambda P_{0,j-1}(t) \ \ldots\ . \tag{2.4}$$

By using the initial conditions

$$P_{00}(0) = 1 \quad \text{and} \quad P_{0j}(0) = 0 \quad \text{if} \quad j \neq 0$$

it can be swown that the general solution of (2.4) is given by

$$P_{0j}(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \ j = 0, 1, 2, \ldots$$

which is a Poisson distribution.

3. Models varying across x, the explanation Variable alone

In a criminal, after being discharged from prison, does not come back to crime n years after discharge, we consider it to be a "success" in this context. Statistical models can be developed to predict the probability of success of such a criminal.

Model 3.1

Let $x_1, x_2, \dots$ be independent or explanatory variables (such as age drug use, previous convictions etc.). A model that quickly comes to mind is the linear model

$$P_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \dots$$

where P_j is the probability of success of the jth criminal. The model parameters can easily be estimated by the method of least squares. However, the model can predict the value of P_j outside the range $(0,1)$ and thus may be considered inadequate for this purpose.

Model 3.2

A model that can also easily predict P_j and is free of the above limitation is the logistic model given as

$$P_j = [1 + \exp -(\beta_0 + \beta_1 x_{1j} + \dots)]^{-1}.$$

The logistic model has been successfully used in analysing bioassay and quanta response data.

4. Models varying with time, t, alone

Here we present time dependent models for the rate $\gamma(t)$ that satisfies the conditions stated in section 1 above.

Model 4.1 $\gamma(t) = t(p_0)^{\alpha t},\ t \geq 0,\ \alpha \geq 0$

where p_0 is the proportion of "successful" criminals in a given sample of criminals and α is the model parameter. The model takes off from the origin, reaches the maximum at time $t = (-\alpha \log P_0)^{-1}$ and finally approaches zero for sufficiently large t corresponding to the end of a career.

Model 4.2: Another appropriate model for the rate γ is given as

$$\gamma(t) = \frac{te^{-\lambda t}}{1 + e^{\lambda}},\ t \geq 0,\ \lambda > 0;$$

where λ is a suitable decay coefficient. Here also the rate starts at the origin, reaches the maximum at time $t = 1/\lambda$ and decays to zero.

More flexible models are given and discussed as follows.

Model 4.3:

$$\gamma(t) = \frac{\beta - \alpha}{\Gamma\alpha} t^{\alpha-1} e^{-t/\beta},\ t \geq 0,$$

$\alpha, \beta > 0$ where α and β are the model parameters (α being the shape parameter and β the decay coefficient).

The model has an initial rate rate of zero, attains a maximum at time $t = \beta(\alpha - 1)$ and eventual tends to zero.

5. Models varying with time t and with x

A more general model for the rate γ can be obtained by combining the dependence of the rate function on both time and variate x.

We consider a few of these types of models.

Model 5.1:

$$\log \gamma(t, x) = \log t + \theta t \log P_0,$$

where P_0 is the proposition of "successful" cases in a sample of criminals and $\theta = \beta_0 + \beta_1 x_1 + \dots$. The model starts at the origin, attains the maximum at time $t = (-\theta \log P_0)^{-1}$ and tends to zero for large t.

Model 5.2

Another general model is proposed as

$$\log \gamma(t, x) = \log t - \theta t - \log(1 + e^{\theta}),$$

where θ is as given in model 5.1.

Model 5.3:

This general model is presented as

$$\log \gamma(t, x) = \log \lambda - \alpha \log t - \theta t$$

where θ is also as given in model 5.1. Similar to the earlier ones, models 5.2 and 5.3 start at the origin, attain maximum values at times t equals $1/\theta$ and α/θ respectively and decay to zero.

Remark. It is interesting to note that by using model (3.2) in model (5.1) or (5.2), the rate $\gamma(t, x)$ of offending decreases as P_j, the probability of success, increases and vice versa. This is to be expected since a criminal with high probability of success will have a tendency of committing fewer number of crimes.

6. Model Fitting

One of the best methods of estimating the parameters of a statistical model is the method of maximum likelihood. It has many desirable properties. Suppose there are N subjects, the ith subject is observed over a period of time from a to b and that during this period, he commits n_i offenses at times $t_{i1}, t_{i2}, \dots, t_{im}$. Then the probability of what has been observed for this subject is proportional to

$$\gamma(t_{i1}, x_i)\gamma(t_{i2}, x_i) \dots \gamma(t_{in_i}, x_i) \exp\left(-\int_a^b \gamma(t, x_i)\, dt\right)$$

(See Cox and Isham [3] for the relevant Mathematical background)

The logarithm of the likelihood function for the complete data is

$$\sum_{i=1}^{N} \sum_{i=1}^{n_i} \log\left(\gamma(t_{ij}, x_i) - \int_{a_i}^{b_i} \gamma(t, x_i)\, dt\right).$$

The integral has to be evaluated for each particular form of $\gamma(t, x)$ and then the complete log likelihood function is coded into machine readable form so that a numerical maximization package can call on its value for any trail set of values of the model parameters.

7. Illustration

Because a comprehensive information on individual criminals was not available at the time of compiling these results, a full illustration cannot be given in this paper. However, some information were collected on a number of criminals in a local prison and models (5.1) and (5.2) were fitted to the data.

Information were obtained on ten explanatory variables on 76 criminals as follows:

$$
\begin{aligned}
x_1 &= \text{age} \\
x_2 &= \text{Institutional life under 15 years} \\
x_3 &= \text{Educational attainment} \\
x_4 &= \text{longest period in any one job} \\
x_5 &= \text{Marital Status} \\
x_6 &= \text{Number of convictions} \\
x_7 &= \text{Recidivism} \\
x_8 &= \text{Psychiatric admission} \\
x_9 &= \text{Approved School} \\
x_{10} &= \text{Number of Juvenile Convictions}
\end{aligned}
$$

All hese variables have been coded to take numerical values. The complete data on 76 cases are displayed in table (7.1). The variable y, which takes values 1 and 0, denotes "outcome" which is classified as "success" or "failure".

A criminal is judged to have attained a success if he does not come back to crime two years after discharge. The data consists of 26 successes and 50 failures. Since there was no information on the times the offenses were committed, the time t in the model is considered fixed. In this situation, we obtain the maximum likelihood estimates of the β's in model (3.2) and use these estimates to fit models (5.1) and (5.2). The maximum likelihood estimates of the β_j's are

$$
\begin{aligned}
b_0 &= -2.5734 \\
b_1 &= 0.1232 \\
b_2 &= 0.7709 \\
b_3 &= 0.7095 \\
b_4 &= 0.4776 \\
b_5 &= 0.2152 \\
b_6 &= 0.3198 \\
b_7 &= -1.6989 \\
b_8 &= 0.6744 \\
b_9 &= 0.6144 \\
b_{10} &= -8.6802
\end{aligned}
$$

The value of P_0 in model (5.1) is 26/76. For a particular criminal, the offending rate functions, $\gamma(t, x)$ are graphed in figures (7.1) and (7.2).

Table 7.1: Criminal career data

References

[1] J. M. Chaiken and M. R. Chaiken, *Varieties of Criminal behavior*, Rand. Corp. santa Monica, Calif (1982).

[2] J. Cohen, *The incapacitative effect of inprisonment: A critical review of literature*, In Blumstein, A., Cohen, J. and Nagin, D. (eds.) Deference and incapacitation. National academy Press, Washington, D. C. (1978).

[3] D. R. Cox and V. Isham, *Point Processes*, Chapman and Hall, London (1980).

[4] M. O. Ojo, *Analysis of some prison data*, Journal of applied Statistics, Vol. 16, no. 6, pp. 377–383 (1989).

[5] M. A. Paterson and H. B. Braiker, *Doing Crime. A survey of California prison inmates*, Rand. Corp. santa Monica, Calif (1980).