Methods for Estimating Crime Rates of Individuals

Executive Summary

Jan M. Chaiken and John E. Rolph,
with Robert L. Houchens

Rand

Prepared under Grant No. 78-NI-AX-0129 from the National Institute of Justice, U.S. Department of Justice. Points of view or opinions stated in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Library of Congress Cataloging in Publication Data

Chaiken, Jan M.
Methods for estimating crime rates of individuals.

"Prepared for the National Institute of Justice."
"R-2730/1-NIJ."
Bibliography: p.
1. Crime forecasting--United States. 2. Crime forecasting. 3. Criminal behavior, Prediction of. I. Rolph, John E. II. Houchens, Robert L. III. Rand Corporation. IV. National Institute of Justice (U.S.) V. Title.
HV6791.C45 364.3 81-5132
ISBN 0-8330-0310-0 AACR2

Published by The Rand Corporation

R-2730/1-NIJ

Methods for Estimating Crime Rates of Individuals

Executive Summary

Jan M. Chaiken and John E. Rolph,
with Robert L. Houchens

March 1981

Prepared under a grant from the
National Institute of Justice, U.S. Department of Justice

Rand
SANTA MONICA, CA. 90406

PREFACE

This work was undertaken under grant number 78-NI-AX-0129 from the National Institute of Justice. The grant is one of several awarded for research in criminal justice evaluation methodologies. At The Rand Corporation, this project is part of the Criminal Justice Program and is closely related to several other projects concerned with policy-oriented research on criminal careers. One of those projects, the Second Inmate Survey, provided data which are used to illustrate the methods in this report.

This document is the Executive Summary of the report:

John E. Rolph, Jan M. Chaiken, and Robert L. Houchens, *Methods for Estimating Crime Rates of Individuals*, The Rand Corporation, R-2730-NIJ, March 1981.

ACKNOWLEDGMENTS

At the National Institute of Justice, this work was encouraged and supported by George Silberman and Richard Linster. At The Rand Corporation we are indebted to Suzanne Polich for calculating the crime rate data used for illustrative examples in our study and to Marcia Chaiken for providing data for examples of potential covariates of crime rates. Naihua Duan and Michael Maltz reviewed an earlier draft of this report with considerable dedication, and they made many helpful suggestions.

CONTENTS

I. INTRODUCTION

CONTEXT AND OBJECTIVES OF THE STUDY

In the course of processing an individual suspected of committing a crime or convicted of a particular crime, the criminal justice system must make many decisions based on beliefs or experiences that relate to the degree and seriousness of the offender's criminal behavior. These decisions affect the disposition of the individual's case and the length of time he or she will be incarcerated after conviction—if at all. Some examples of differential handling follow:

- *Law enforcement.* Some large police departments and sheriffs' departments operate "major offenders units" to keep track of certain known offenders and to arrest them when they commit a crime.[1] Because of the cost of tracking, only individuals the police anticipate will commit numerous serious crimes are assigned to such units.
- *Prosecution.* Some district attorneys operate "career criminal prosecution units," which bring special resources to bear in an attempt to assure conviction of certain individuals when they are arrested, and imprisonment when they are convicted (Bronx County, 1976; Dahmann and Lacy, 1977; INSLAW, 1977).
- *Sentencing.* In deciding whether a convicted person should be given probation or incarcerated, judges often consider whether the conviction crime appears to be an isolated incident or part of a pattern of substantial criminal activity. And in establishing the length of sentence to be imposed for particular crimes and combinations of crimes, especially when prior criminal record is taken into account, legislators may have in mind their own beliefs about the kinds of records that indicate a person is a high-rate offender.
- *Parole.* When parole boards decide whether a prisoner should be released, they consider the chances of "parole success," which essentially means the probability that the individual will commit a crime or violate a condition of parole during a specified future time period (Gottfredson et al., 1978; Hoffman and DeGostin, 1974).

Despite the facts that distinctions among individuals based at least in part on their crime commission rates are made within the criminal justice system, and that the people who make these distinctions may feel quite confident of the correctness of their decisions based on extensive personal experience, research shows that it is exceptionally difficult to predict accurately who *will be* a high-rate criminal offender, or even to determine from personal descriptors and criminal records who *has been* a high-rate offender during a specified period of time in the past. Nonetheless, research also shows that of the people who commit crimes, most commit only a relatively small number, while a few people commit crimes at substantially higher rates (Wolfgang et al., 1972; Peterson et al., 1980; Greene, 1977). In short, very high-rate offenders exist, but it is not easy to identify them.

The work described here develops methods that can be used to analyze crime commission rates and thereby shed light on the problem of distinguishing between low-rate and high-rate offenders. Analytically, the problem divides into two general categories of questions:

[1]Pate et al. (1976) describe the Perpetrator-Oriented Patrol in Kansas City (Missouri) and two predecessor projects in Miami (Florida) and Wilmington (Delaware). Additional police programs focusing on major offenders were instituted in Rochester (New York), Amarillo (Texas), Pueblo (Colorado), and Norfolk (Virginia).

1. Given the best possible information about an offender and his criminal behavior during a previous period, how can one estimate his crime commission rates during that period?
2. For a group of criminal offenders with specified characteristics, what can be said about their average rate of committing various crimes, the distribution of those crime commission rates, and the extent to which their rates differ from those of another group?

The second category of questions avoids the problem of identifying particular individuals as high-rate offenders. It focuses instead on aggregate behavior. Methods for answering questions of this type can be useful for determining whether rules already being used by police, prosecutors, parole boards, etc. actually distinguish the intended target group of high-rate offenders from others. Our methods can also help in devising better decision rules for selecting offenders to receive special attention or punishment, although other considerations certainly play a role, for example, the equity of treatment of similar persons in similar circumstances, the feasibility of obtaining the necessary information for the decision rule in a timely fashion, and "just deserts" as applied to the conviction crime (von Hirsch, 1976; Morris, 1974).

In addition, our methods can be used to analyze groups of offenders defined by characteristics presumably unrelated to their crime propensities, for example the city or state in which they reside. By permitting a determination of whether a city with relatively low per capita crime rates has (a) relatively fewer criminals than other cities or (b) lower crime commission rates among those who are offenders (or both), these methods can help in studying the effectiveness of various governmental anticrime activities and the deterrent effect of city- and state-level sanction policies.

RELATIONSHIP BETWEEN THIS STUDY AND POLICY RESEARCH

Other studies at The Rand Corporation are examining questions related to prediction of high-rate offenders, decision rules for selective handling of offenders, and deterrence (Greenwood, 1980). To carry out that research, self-report data were collected from incarcerated offenders about the crimes they committed during specified periods of time, and both self-report and official data were collected about their characteristics that presumably relate to criminal behavior (Peterson, Chaiken, and Ebener, forthcoming). The analysis in these research projects yields estimates of the crime commission rates of the surveyed offenders, relates these rates to their personal characteristics, and extrapolates the results to more general populations of offenders.

The methods described in this report can assist in endeavors like these, not only for Rand's inmate surveys but also for other research that uses any source of data concerning the criminal activity of individuals. Our illustrative examples, drawn from Rand inmate survey data, do not answer the major analytical questions related to that survey but rather show how our methods can be applied in practice.

METHODOLOGICAL ISSUES ADDRESSED

The basic model underlying our approach is that there are K types of crime of interest, and each criminal offender commits each of the crimes at a specified rate (possibly zero) when free to do so. An offender's annualized crime commission propensity for crime k is denoted λ_k; it represents the expected number of crimes he would commit per year of "street time"[2] if his current behavior persisted indefinitely. Due to statistical fluctuations and other reasons, the actual number of crimes of type k he commits in a particular year may differ from λ_k. For a group of offenders, who have been specifically selected for study, we assume information is available about the number of crimes of each type they committed during a particular period of time, called the "measurement period."

Our methods help accomplish the following goals:

- To describe the distribution of the observed crime rates for the selected offenders who provided information about their actual crime commissions.
- To estimate the crime commission propensities of any *one* of these individuals, taking into account the group's overall distribution of commission rates as well as the *individual's* reported crime commissions and other characteristics.
- To estimate the distribution of crime commission propensities for more general populations of offenders who differ from the selected offenders in known ways.

[2]"Street time" refers to the periods when an offender is free to commit crimes against the general public, as distinguished from periods of incarceration or hospitalization.

II. UNIVARIATE DISTRIBUTIONS OF OBSERVED CRIME COMMISSIONS

Although many offenders commit more than one type of crime, the distribution of commission rates for a particular crime type (i.e., the univariate distribution) is often of interest. For the individuals who were selected for study and provided data showing the number of crimes of that type they committed, estimating the univariate distribution appears to be a routine task. However, the following obstacles arise:

- *Skewed distribution.* A large proportion of offenders have very low λ_k for any given crime k. Yet the proportion that have very high λ_k is larger than would be anticipated from most commonly used smooth distributions that fit the data for low and moderate values of λ_k.
- *"Too many zeros."* Some of the individuals who did not commit *any* crimes of type k during the measurement period have $\lambda_k = 0$, while others have small (but nonzero) λ_k and/or a short measurement period. Hence the data do not permit a clear distinction between people with nonzero λ_k and those with zero λ_k. A parsimonious way to handle the data would be to assume that all values of λ_k are positive.[1] However, neither common sense nor typical data collected from criminals are consistent with such simplified assumptions.
- *Instrumentation error.* Although respondents to a survey may have been asked to state the number of crimes of type k they committed during a specified period, their answers may be ambiguous or imprecise due to the format of the survey instrument, incomplete responses, or respondent error.

EXPLORATORY ANALYSIS

In working with data having these characteristics, the problem of imprecise responses must be handled first. Mathematically, most imprecise responses can be treated as censored observations, whose values are known only to the extent that they fall in an interval, possibly unbounded on the right or the left. Many methods are available for handling censored observations. One method we used is a nonparametric maximum likelihood estimation of the empirical distribution reflecting both censored and uncensored observations.[2] Then, if one wishes to test the fit of parametric distributional forms to the data, the estimated nonparametric distribution can serve as a comparison.

To illustrate the method, we worked with a subset of data from Rand's second inmate survey. These were preliminary data giving, for each of 440 respondents, a low and a high estimate of the *length* of his measurement period (in years) and low and high estimates (from self-reports) of the *counts* of crimes of each of the following types he committed during the measurement period:

---

[1]Under this assumption, people who do not commit any crimes of type k are envisioned to have very low values of λ_k, rather than $\lambda_k = 0$.

[2]See Turnbull (1976).

1. Burglary
2. Robbery of businesses
3. Robbery of persons
4. Theft (other than auto theft)
5. Auto theft
6. Forgery and credit cards
7. Fraud
8. Dealing drugs

Various transformations were applied to the annualized reported crime rates to see whether the data might (approximately) follow a lognormal, gamma, or Pareto distribution. When individuals with zero counts were treated as if they had low (but nonzero) values of λ_k, the lognormal and gamma forms were decisively rejected, and the Pareto form was accepted for four of the eight crime types. The values of the parameters of the Pareto distributions that fit the data for these four crime types implied that the means and variances of the distributions are infinite. Hence, the Pareto parametric form, while statistically acceptable as a description of four of the crime types including their zero counts, was not a practically satisfying description of the data. This result demonstrated that the problem of "too many zeros" cannot be satisfactorily handled by the expedient of considering zero values of λ the same as very small positive values.

When, on the other hand, all respondents with zero counts were treated as if they had $\lambda_k = 0$ (i.e., they were considered irrelevant for estimating the distribution of crime commission rates and were excluded as missing), only one crime type (dealing drugs) fit one of the smooth functional forms we tried, namely, a gamma distribution. We then tried intermediate formulations, in which some but not all of the zero counts were considered to represent nonzero λ_k's, and found that the commission rates for four additional crime types approximately fit a gamma distribution.[3]

In summary, the exploratory analysis, while not revealing any single method or parametric form that sufficed for all crime types, did yield a reasonably satisfactory parametric description of observed crime commission rates for six of the eight crime types examined. Figure 1 gives an example of the empirical distributions and the parametric distributions that fit the data for one crime type: robbery of persons.

Each person in the sample reported a number Y of robberies committed and a length T of his measurement period. The bar graph in Figure 1 shows the distribution of the observed annualized crime commission rate, Y/T. Only the truncated distributions are shown; that is, individuals who reported zero robberies of persons are excluded from the graphs even though they were included when fitting the Pareto distribution. The ordinate shows the percentage of respondents in each interval of width 2, i.e., $Y/T \in (0, 2), [2, 4), [4, 6)$, etc., and the percentage with $Y/T \geq 20$.

Figure 1 shows that the Pareto distribution overestimates the frequency of small, but nonzero, crime commission rates (under 4), but it gives an excellent fit in the tail of the distribution (over 14). By contrast, the gamma distribution is closer to the empirical distribution at the low end, but it overestimates the frequency of moderate values and underestimates in the tail.

[3]These included three of the crime types previously fit to a Pareto distribution. Two crime types (burglary and theft) did not fit any of the distributional forms studied.

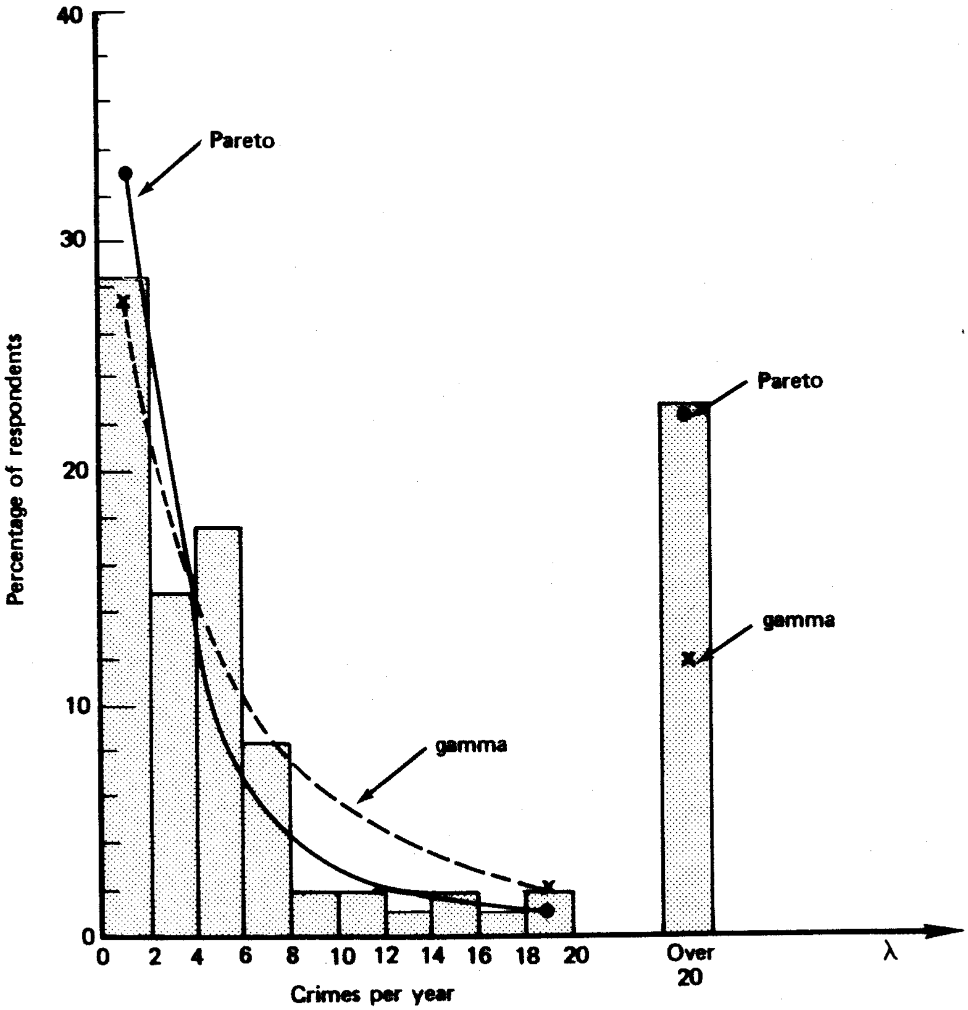

Fig. 1—Distribution of crime rate for robbery of persons
(respondents who rob persons)

GAMMA-POISSON MODEL

When the observed crime commission rates appear to fit a gamma distribution approximately, a reasonable hypothesis is that the distribution of *counts* of crimes committed, conditioned on the length of the measurement period, is negative binomial.[4] The negative binomial distribution can arise in many different ways, one of which is as a gamma mixture of Poisson distributions. In the mixed gamma-Poisson model, the underlying crime commission propensity λ_k is assumed to have a gamma distribution, and each offender is assumed to commit crime k according to a Poisson process with his value of λ_k during the measurement period. Under these assumptions, all offenders whose measurement period has length T will have their counts of crime k drawn from a negative binomial distribution (Johnson and Kotz, 1969, Ch. 5). The functional forms of these distributions and the relationships among their parameters are shown in Display 1.

Whether or not one accepts the mixed gamma-Poisson model, it is possible to test the hypothesis that, conditioned on the length of the measurement period, the counts of crimes have a negative binomial distribution. If all offenders who provided data have measurement periods of the same length, the negative binomial model is convenient to fit and test. In the case of the data we received from Rand's Second Inmate Survey, respondents had different measurement periods, so we fit a negative binomial distribution to the data for a subset of offenders who had approximately the same length measurement period.

To handle the problem of "too many zeros," we fit the *nonzero* counts to a *truncated* negative binomial distribution (i.e., the distribution conditioned on positive counts—see Display 1). This fitting process yields estimates $\hat{a}$ and $\hat{P}$ of the parameters of the distribution, which can then be used to estimate how many of the zero counts belong to the negative binomial distribution. Denoting by m the number of respondents with nonzero counts, the total number of respondents with nonzero λ can be estimated as the integer part of $m/(1 - \hat{P}^{\hat{a}})$. The correction is nontrivial. For example, in a subset containing 78 respondents who claimed zero drug deals during the measurement period, we estimated 19 of them had nonzero λ.

If one accepts the gamma-Poisson model, then estimates of the parameters a and P yield estimates of the shape and scale parameters α and β of the underlying gamma distribution. These can be used to estimate the mixture of negative binomial distributions (corresponding to the different lengths T of the measurement periods). Not every crime type in our illustrative data base could be fit well by the truncated negative binomial. However, the method can also be applied to other truncated distributional forms, such as the truncated Waring distribution discussed below.

ADEQUACY OF THE POISSON ASSUMPTION

The appropriateness of the mixed gamma-Poisson model depends not only on the gamma form of the underlying crime rate distribution but also on the assumption that each offender's crime commissions occur according to a Poisson process. We examined the latter assumption also, using data from Rand's Second Inmate Survey.

The results suggest that the Poisson assumption may be correct for most types of crimes but not for all of them. For example, the activities of robbing people, stealing cars, dealing in

[4]The negative binomial distribution, which is described in the text that follows, is potentially correct in theory for the *counts*. The rates (ratios of counts to measurement times) cannot in principle be described exactly as drawn from a gamma distribution.

Display 1

1. *Gamma Distribution*

$$\text{density } f(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda) \qquad \text{for } \lambda > 0.$$

α = shape parameter
β = scale parameter
mean = α/β
variance = α/β^2

2. *Poisson Distribution*

$$\text{Prob}(Y = y|\lambda, T) = \frac{(\lambda T)^y}{y!} \exp(-\lambda T) \qquad y = 0, 1, 2, \ldots$$

λ = commission rate (crimes per year)
T = length of measurement period (years)
mean = λT variance = λT

3. *Negative Binomial Distribution*

$$\text{Prob}(Y = y) = \binom{a + y - 1}{y} (1 - P)^y P^a \qquad y = 0, 1, 2, \ldots$$

$$\text{mean} = a \frac{1 - P}{P}$$

$$\text{variance} = a \frac{1 - P}{P^2}$$

In the compound Gamma-Poisson Model:

$a = \alpha$
$P = \beta/(T + \beta)$
mean = $\alpha T/\beta$
variance = $\alpha T(T + \beta)/\beta^2$

4. *Truncated Negative Binomial Distribution*

$$\text{Prob}(Y = y|y > 0) = \binom{a + y - 1}{y} \frac{(1 - P)^y P^a}{1 - P^a} \qquad y = 1, 2, \ldots$$

drugs, or doing forgery or fraud appear to be spread out over the offender's time on the street in a way that is consistent with a Poisson process. (The temporal disaggregation of the data was inadequate for a formal test of the Poisson process.) However, the robbing of businesses appears in some instances to be concentrated in short time periods within the overall street time. Including such "spurting" behavior in models of crime commissions does not present substantial technical difficulties in some contexts, while it does in others. In the remainder of this report we continue to assume a Poisson process for crime commissions.

III. ESTIMATES OF INDIVIDUAL CRIME PROPENSITIES

If offender i commits Y_i crimes during a measurement period of length T_i, the usual (maximum likelihood) estimate of his crime commission propensity $\lambda(i)$ is

$$\hat{\lambda}(i) = \frac{Y_i}{T_i} .$$

However, the distribution of these estimates, $\hat{\lambda}$, is more spread out than the distribution of the underlying true λ. Intuition suggests that a better estimate of $\lambda(i)$ can be obtained by shrinking $\hat{\lambda}(i)$ toward the value of λ that is expected for offenders "similar" to offender i. In case the distribution of λ is gamma (α, β) and offender i commits crimes according to a Poisson process at rate $\lambda(i)$ during a period of length T_i, an improved estimate of $\lambda(i)$ is the Bayes estimate

$$(1 - w_i) \frac{Y_i}{T_i} + w_i \frac{\alpha}{\beta} ,$$

where $w_i = \beta/(\beta + T_i)$. Here α/β is the mean of the underlying gamma distribution, so the Bayes estimate is a weighted average of the ordinary estimate and the a priori mean.

More generally, Hudson and Tsui (1980) and Hudson (1980) have developed shrinkage estimators for Poisson data that are improvements over the usual estimator, whatever the form of the underlying distribution. Shrinkage estimators show greatest improvement when applied separately to homogeneous groups of data (Carter and Rolph, 1974; Efron and Morris, 1973). Thus, it is desirable to divide the offenders into subgroups according to characteristics other than the number of crimes they committed in the measurement period. Homogeneity is obtained if the members of each subgroup are similar in terms of the relationship between their personal characteristics (or prior behavior) and their crime commission propensities.

In applying the method to data from Rand's Second Inmate Survey, we developed regression estimates of crime commission propensities in the measurement period from covariates such as age, criminal behavior as a juvenile, and use of specified drugs. For illustrative purposes only, three subgroups were considered for each crime type, although elaborations using more subgroups are possible and would be desirable for analysis of substantive issues related to the inmate survey.

One motive underlying our division into subgroups was to separate individuals with $\lambda_k = 0$ from those with $\lambda_k > 0$. However, one cannot actually define subgroups in this way because

- The data do not tell us exactly who has $\lambda_k = 0$.
- The Hudson/Tsui method does not permit defining subgroups in terms of the crime commission data. That is, the offenders with $Y_i = 0$ cannot be used to define one subgroup since the remaining subgroups would then have no zero counts. But the Hudson/Tsui theory assumes that the count of crimes committed by offender i during T_i has a Poisson distribution, not a truncated Poisson distribution. Consequently, the data for any subgroup should, with high probability, include some zero counts.

We felt a reasonable approach was to divide the respondents into "previous doers" of crime k and "previous nondoers," according to whether they did or did not commit crime k in the four years prior to the measurement period. The estimates $\hat{\lambda}(i)$ for previous nondoers were shrunk toward zero.[1] The estimator $\hat{\lambda}(i)$ for previous doers was pulled toward a regression estimator of $\lambda(i)$, based on covariates such as age, use of specified illegal drugs, extent of crime as a juvenile, and employment history. (If $Y_i = 0$ for a previous doer and his regression estimate is positive, his Hudson/Tsui estimate is also positive.)

To illustrate possible additional elaborations of the method, we also separated out those previous doers whose data appeared (from validity analyses not reported here) to be of possibly low quality. Such respondents were excluded when fitting regressions but were included when using the regression estimator to obtain the final shrinkage estimators. The result of the shrinkage process was that most respondents' estimates were practically unchanged, although those with small original estimates sometimes experienced substantial percentage change.

[1]In particular, if $Y_i = 0$ for a previous nondoer, his Hudson/Tsui estimate is zero; otherwise it is positive.

IV. MULTIVARIATE MODELING

Our efforts to describe a multivariate distribution capturing not only the univariate distributions for each of the K crime types but also the covariance among them were frustrated by several factors. Two of the factors have already been noted in the discussion of univariate distributions:

- Each offender could have $\lambda_k = 0$ or $\lambda_k > 0$ for $k = 1, 2, ..., K$, thereby necessitating $2^K - 1$ parameters just to describe the patterns of "doing" vs. "nondoing."[1]
- No single family of distributional forms was found to fit all the univariate distributions simultaneously, so we could not hypothesize a suitable multivariate generalization.

We made a modest effort to explore the multivariate distribution that might apply to three Rand Survey crime types whose univariate distributions suggested an underlying gamma distribution of propensities. This is the model: An offender's value of λ_k is determined by the sum of two independent draws from gamma distributions, the first gamma distribution having parameters (α_0, β) and the second having parameters (α_k, β). Note that the first distribution is common to all crime types, and the scale parameter β is the same in all the gamma distributions. For three crime types, this results in a five-parameter description of the multivariate distribution.

In this model, the counts of crimes committed by offenders during the measurement period have a form, derived in the text, that we call a generalized multivariate negative binomial (GMNB) distribution. A generalized least-squares method was developed for estimating the parameters of the GMNB distribution.

Since every offender is assumed to have nonzero λ_k for each k, we applied the method for estimating parameters only to respondents who were "previous doers" of at least one of the three crimes. This admittedly inadequate approach to the problem of "too many zeros" led to a fit GMNB distribution that was unsatisfactory, and we attribute this to the zero problem. This finding suggests that a better method of fitting the multivariate distribution might be a truncated GMNB distribution in the same spirit as we used for the univariate case. However, no simple solution to the "zero problem" exists. It is not sensible to include in the model only those offenders who have nonzero counts for *every* crime of interest.

Exploration of the literature on distributions similar to the GMNB revealed a promising avenue for future efforts to model the multivariate distribution of crime commission rates. A generalized Waring distribution discussed by Irwin (1968, 1975) appears to meet the requirements we have identified here for the univariate distribution. (It includes the negative binomial as a special case.) A multivariate version of the generalized Waring has been developed by Sibuya (1980), who includes a treatment of truncating zero counts. The parameters can be estimated by maximum likelihood methods, but we did not carry out the extensive programming necessary to test this distribution against our sample data from Rand's Second Inmate Survey.

[1]In practice, somewhat fewer than $2^K - 1$ distinct combinations actually occur, but it remains true that many parameters are needed.

V. EXTRAPOLATIONS TO MORE GENERAL POPULATIONS OF OFFENDERS

For various practical reasons, most self-report surveys of adult criminal behavior are restricted to offenders who are incarcerated or otherwise under the control of the criminal justice system. For policy purposes, however, their crime commission propensities are not very interesting.

Incarcerated offenders—and especially imprisoned offenders—are frequently seen as an extraordinarily atypical group, the "losers" among criminals. For this reason, most people who are interested in individual crime commission rates do not want to know the distribution of rates for prisoners but rather for some other specific group. For example, prosecutors may be interested in the crime commission propensities of *arrestees* or the characteristics of high-crime-propensity arrestees. Researchers concerned with deterrence or the sociodemographic factors related to crime may be more interested in a general population of active offenders than in arrestees. By contrast, judges may want to know the characteristics of convicted persons with high-crime-commission propensities.[1]

To extrapolate data collected from one group of offenders into estimates of distributions for some other target population of interest, one merely needs to know the sampling probabilities for members of the study group in relation to the target population. This approach fails only if some offenders in the target population have zero probability of being in the sample.

One method for estimating the sampling probabilities applies to a situation where the target population includes all active offenders in a particular jurisdiction (whether on the street or not) and the sample is a random subset of a cohort of offenders *arrested* in that jurisdiction in a given time period.[2] For specificity, suppose the sample is a cohort of arrestees for a given year. The target population can be defined as consisting of everyone who has a positive crime commission propensity λ for one or more of the crimes for which an arrest can be made.

In the model, offender i in the target population is assumed to commit crime k according to a Poisson process at rate $\lambda_k(i)$ when free ("on the street"), and he has probability $Q_k(i)$ of being arrested for crime k, given that he has committed crime k. For estimation purposes, it is preferable to postulate a different arrest probability for each crime type, rather than an overall arrest probability for offender i, which would reflect the mix of crime types he commits as well as his chances of arrest for each crime. Nonetheless, his probability of being a member of the arrest cohort is estimated from his arrest *rate* (arrests per year), namely, $\sum_k \lambda_k(i)\, Q_k(i)$.

Assuming that we know the number of crime commissions of each type that led to an arrest for the sampled offenders during the measurement period (as we do in Rand's Second Inmate Survey), we can estimate the arrest rate for each such offender by shrinkage methods analogous to those described above for estimating the crime commission propensities $\lambda_k(i)$. We can shrink the offender's data toward values calculated from his values of $\lambda_k(i)$ and a priori estimates of $Q_k(i)$. Examples of suitable a priori estimates are:

---

[1]Convicted persons, as a group, have characteristics differing from those of prisoners, because many of them do not go to prison.

[2]The same method applies if the sample is drawn from a cohort convicted in the period, or incarcerated in the period, or imprisoned in the period.

- The average arrest probabilities Q_k for the jurisdiction, estimated from data external to the survey.
- The average arrest probabilities Q_k for the sampled offenders, estimated from their data.

Although it would also be possible to shrink toward a rate estimated by taking into account the *characteristics* of each offender, this approach does not appear attractive thus far, since we have not yet found independent covariates that are significantly related to arrest probabilities.

Under the above assumptions about crime commissions and arrest probabilities, together with some independence assumptions, we can deduce the stochastic process for arrests. For example, when the crime commissions for each crime type are Poisson, the overall arrest process (all crimes together) is also Poisson. From this information we can deduce the probability that an offender who is on the street at the beginning of the year will be arrested at least once during the year (and hence will belong to the one-year arrest cohort). A slight correction is then introduced to account for offenders who are incarcerated at the start of the year and may be released and arrested again during the year.

This method produces an estimate of the probability that an offender with the same parameters as sampled offender i will be found in a one-year arrest cohort. Hence, sampled offender i can be considered to represent a larger group of offenders whose size is inversely proportional to his sampling probability. Using these sampling weights, the characteristics of the target population of offenders can be produced.

Details of these probability calculations are given in our report (Rolph, Chaiken, and Houchens, 1981), as well as the analogous results when different assumptions are made about the sampled offenders. For example, the sample could consist of all offenders from a given jurisdiction who are found in jail or prison on a specified date.

In general, offenders who are in custody of the criminal justice system overrepresent offenders who have high crime commission rates and those who commit crimes with high probabilities of arrest, conviction, and incarceration (e.g., homicide and kidnapping). Consequently, the distributions of crime commission rates for target populations typically differ quite substantially from the distributions estimated for the study group.

REFERENCES

Bronx County (New York) District Attorney's Office, "The Major Offense Bureau: An Exemplary Project," for LEAA, U.S. Department of Justice, 1976.

Carter, G. M., and J. E. Rolph, "Empirical Bayes Methods Applied to Estimating Fire Alarm Probabilities," *Journal of American Statistical Association,* Vol. 69, December 1974, pp. 880-885.

Dahmann, J., and J. L. Lacy, *Criminal Prosecution in Four Jurisdictions: Departures from Routine Processing in the Career Criminal Program,* METREK-MITRE, *Mitre Technical Report,* MTR-7550, 1977.

Efron, B., and C. Morris, "Stein's Estimation Rule and Its Competitors—an Empirical Bayes Approach," *Journal of American Statistical Association,* Vol. 68, 1973, pp. 117-130.

Gottfredson, D. M., et al., *Classification for Parole Decision Policy,* National Institute of Law Enforcement and Criminal Justice, Washington, D.C., 1978.

Greene, M. A., "The Incapacitative Effect of Imprisonment Policies on Crime," Ph.D. dissertation, Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1977.

Greenwood, P. W., *Rand Research on Criminal Careers: An Update on Progress to Date,* The Rand Corporation, N-1572-DOJ, October 1980.

Hoffman, P., and L. K. DeGostin, *Parole Decision-Making: Structuring Discretion,* U.S. Board of Parole Research Unit, Report 5, 1974. Reprinted in L. Radzinowicz and M. E. Wolfgang, *Crime and Justice,* Vol. II, Basic Books, New York, 2d ed., 1977.

Hudson, H. M., "Adaptive Estimators for Simultaneous Estimation of Poisson Means," submitted for publication, 1980.

Hudson, H. M., and K. Tsui, "Simultaneous Poisson Estimators for A Priori Hypotheses About Means," *Journal of the American Statistical Association,* Vol. 76, 1981, pp. 182-187.

Institute for Law and Social Research, "Curbing the Repeat Offender: A Strategy for Prosecutors," PROMIS Research Project, Publication 3, 1977.

Irwin, J. O., "The Generalized Waring Distribution Applied to Accident Theory," *Journal of Royal Statistical Society,* Series A, Vol. 131, 1968, pp. 205-225.

Irwin, J. O., "The Generalized Waring Distribution," *Journal of Royal Statistical Society,* Series A, Vol. 138, 1975, Part I, pp. 18-31; Part II, pp. 205-225; Part III, pp. 374-384.

Johnson, N. L., and S. Kotz, *Discrete Distributions,* Wiley, New York, 1972.

Morris, N., *The Future of Imprisonment,* University of Chicago Press, Chicago, 1974.

Pate, T., R. A. Bowers, and R. Parks, *Three Approaches to Criminal Apprehension in Kansas City: An Evaluation Report,* Police Foundation, Washington, D.C., 1976.

Peterson, M. A., and H. Braiker, with S. Polich, *Doing Crime: A Survey of California Prison Inmates,* The Rand Corporation, R-2200-DOJ, 1980.

Peterson, M., J. Chaiken, and P. Ebener, *Survey of Prison and Jail Inmates: Background and Methods,* The Rand Corporation, N-1635-NIJ (forthcoming).

Rolph, J., J. Chaiken, and R. Houchens, *Methods for Estimating Crime Rates of Individuals,* The Rand Corporation, R-2730-NIJ, 1981.

Sibuya, M., "Multivariate Digamma Distribution," *Annals of Institute of Statistics and Mathematics,* Vol. 32, Part A, 1980, pp. 25-36.

Turnbull, B. W., "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data," *Journal of the Royal Statistical Society,* Series B, Vol. 38, No. 3, 1976, pp. 290-295.

von Hirsch, A., *Doing Justice: The Choice of Punishment,* Hill and Wang, New York, 1976.

Wolfgang, M. E., R. M. Figlio, and T. Sellin, *Delinquency in a Birth Cohort,* University of Chicago Press, Chicago, 1972.

RAND/R-2730/1-NIJ