SELECTIVE INCAPACITATION STRATEGIES BASED ON ESTIMATED CRIME RATES

Jan M. Chaiken and John E. Rolph

April 1978

P-6116

SELECTIVE INCAPACITATION STRATEGIES BASED ON ESTIMATED CRIME RATES

Jan M. Chaiken
John E. Rolph

The Rand Corporation
1700 Main Street
Santa Monica, California 90406

ABSTRACT

Methods for estimating the crime commission rates of criminal offenders are discussed in the context of a potential selective incapacitation strategy that would assign different sentence lengths according to whether the estimated crime rate is above or below a specified threshold. Any such strategy is subject to error because the true crime rate of an offender may differ from his estimated crime rate. For two strategies having the same cost, one of them is favored over the other if it has a higher expected number of crimes averted or if it has a lower probability of assigning long sentences to offenders with low crime rates. Both of these criteria are met by using a Bayes estimate of the crime rate rather than a maximum likelihood estimate. This is demonstrated by calculating the distribution of true crime rates for offenders whose estimates are above a threshold.

One rationale for incarcerating convicted criminals in prison is the fact that while in prison, criminals cannot commit crimes that affect "outside" society. Since individuals have different propensities for committing crimes, one can consider the possibility of sentencing policies that will give longer prison terms to those people with high crime commission propensities than to those with low propensities. Such a policy is called a selective incapacitation strategy because it is focused on reducing the amount of crime in society by physically preventing the offender from committing crimes, ignoring any other objectives of imprisonment, such as retribution and deterrence [1]. While we do not believe that selective incapacitation strategies would necessarily be desirable public policy, we will explore here how the effects of such strategies can be calculated. We shall show that even with "ideal" sources of information about offenders, there are inevitably inequities in selective incapacitation strategies, and their effect is less than might be anticipated.

The basic idea can be understood by examining an overly simplified model. We assume that N convicted offenders are to be sentenced to prison and that the i-th offender would have crime commission rate λ_i if he were free. (For the moment, λ_i is assumed known.) If the i-th offender is imprisoned for a length of time S_i, then the total incarceration cost is proportional to $Y = \Sigma S_i$, which is the total person-time spent in prison, and the expected number of crimes averted is $Z = \Sigma\lambda_i S_i$. By varying the sentence lengths S_1, S_2, ..., one affects both the "cost" Y and the incapacitation effect Z.

If the options for sentence lengths are specified, one can devise strategies that maximize the incapacitation effect at fixed "cost" Y_0. In this paper, we shall envision that there are two choices s_1 and s_2 for the sentence lengths (say, either 3 years or 5 years in prison), where s_1, the "ordinary" sentence, is less than s_2, the "enhanced" sentence. For convenience assume that the offenders are numbered in order of their λ_i, with λ_1 being the highest value. Let K be the largest integer such that $(N - K)s_1 + Ks_2 \leq Y_0$. Then the policy that maximizes the incapacitation effect while constraining cost to be no greater than Y_0 is as follows: Offenders 1, 2, ..., K receive the enhanced sentence of length s_2, while offenders K + 1, ..., N receive the ordinary sentence of length s_1. In other words, there is a cutoff $C = \lambda_K$ such that if $\lambda_i \geq C$ offender i receives the enhanced sentence, and if $\lambda_i < C$ he receives the ordinary sentence.

This optimal policy cannot be applied in practice, because we do not actually know the number of crimes averted by imprisoning offender i for length of time S_i. There are several reasons for this lack of knowledge:

1. The relationship between past crime commission rates and future crime commission rates is unknown, especially since conviction (even without incarceration) might potentially change an offender's crime commission rate.
2. Even if the offender's λ_i were an excellent guide to this future λ_i, we cannot determine his past λ_i exactly from information about the number of crimes he has committed; we can only estimate his past λ_i.

3. The above model ignores the possibility that the offender's criminal career might terminate naturally before time S_i has expired. In extreme cases, for example if S_i = 85 years, it is clearly erroneous to assume that the offender would have continued to commit crimes at rate λ_i for the entire 85 years if he were not imprisoned. However, our formula for the incapacitation effect Z assumes that 85 λ_i crimes are averted. Even if S_i is fairly small, it is possible that the offender would not have committed any more crimes.
4. Conversely, the model ignores the possibility that a prison term might extend the duration of a criminal career. For example, suppose that offender i is "predestined" to end his criminal career at age A_i if not incarcerated; but if he is incarcerated for time S_i his career will continue to age $A_i + S_i$. In this case, incarceration has not averted any crimes. It has simply caused the crimes to be committed at a later date. Thus, it is possible that incarceration has no incapacitation effect whatsoever, or at least a substantially smaller incapacitation effect than the one estimated in our formula for Z.

In this paper, we shall examine only the second of these difficulties, which is that an estimate of an offender's λ_i may not be equal to his true λ_i. As a consequence, some offenders with low crime commission rates can erroneously appear to qualify for an enhanced sentence, while some high-λ offenders escape the enhanced sentence; the incapacitation effect is then smaller than would have occurred if true values of λ_i were known. Throughout, we assume that it is somehow possible to determine the number of crimes committed by offender i in the past. Thus, we are examining the "best" possible results that can be obtained from a selective incapacitation policy, ignoring the difficult problem of estimating an offender's crime rate from information about his personal characteristics, previous history of arrests, etc.

1. RELATIONSHIP BETWEEN ESTIMATED AND TRUE CRIME COMMISSION RATES

First consider a very simple situation. Suppose all criminals commit crimes according to a Poisson process with the <u>same</u> parameter λ. Suppose further that for each offender one can determine the number of crimes committed during a street-time period of fixed length T, say two years. ("Street time" refers to periods when the offender is free to commit crimes, i.e., he is not incarcerated.) Let N_i be the (random) number of crimes committed by individual i during period T. Then the maximum likelihood estimate of λ_i, individual i's average crime rate, is

$$\hat{\lambda}_i = N_i/T. \tag{1}$$

These estimates will differ among offenders even though λ is the same for everyone. As an example, we might assume that λ = 5/year (as well as T = 2 years). Then the probability frequency function of $\hat{\lambda}_i$ is pictured in Figure 1. Supposing the number of convicted criminals observed is large, then the relative frequency of values of $\hat{\lambda}_i$ will approximate Figure 1.

Suppose that we now adopt a cutoff number C and attempt to incarcerate those offenders with values of $\lambda_i \geq C$ for a longer period than those who have $\lambda_i < C$. Since λ_i itself is unobservable, the value of

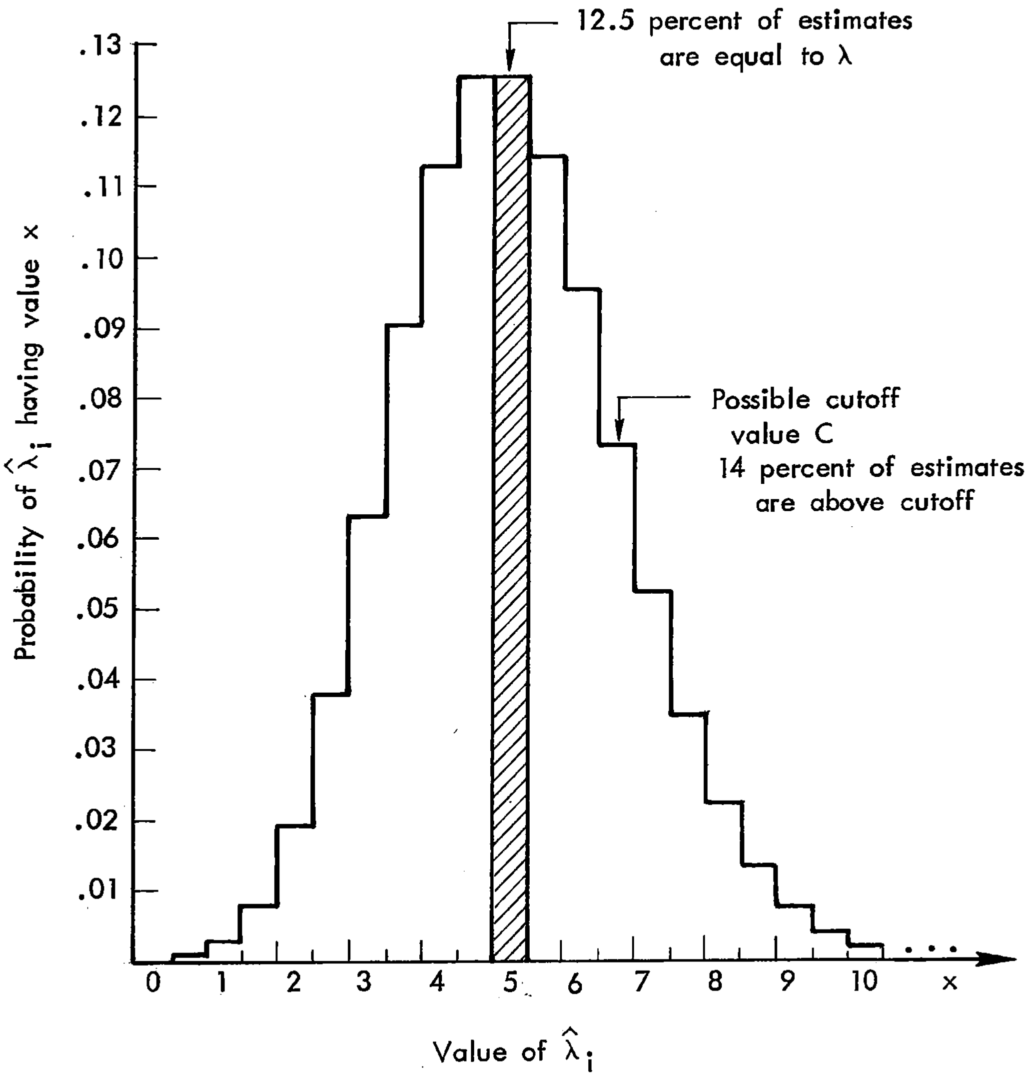

Fig. 1 — Frequency function for the observed value of the estimated crime rate. All offenders commit crimes according to a Poisson process with the same rate $\lambda = 5$/year, and their numbers of crimes are counted after $T = 2$ years.

the estimator, $\hat{\lambda}_i$, might be used instead. As Figure 1 shows, the probability distribution of $\hat{\lambda}_i$ is spread out even though all the λ_i are equal--the distribution of λ_i is concentrated at the one point $\lambda_i = 5$. Thus, whatever cutoff C is chosen, some criminals would be selected for enhanced sentences despite the fact that all of them have the same value λ_i. The incapacitation policy based on this cutoff, which appears to be selective, in fact has the same incapacitation effect as any other policy that gives enhanced sentences to the same proportion of offenders.

Figure 2 shows a more general situation, where the λ_i's are not all the same but instead have a probability distribution themselves. The distribution of the λ_i's results in a distribution of the $\hat{\lambda}_i$'s. Because each $\hat{\lambda}_i$ is equal to λ_i plus a random error, the group of offenders whose λ_i's are above any given cutoff C will be different from the group whose estimates $\hat{\lambda}_i$ are above C. In fact, as can be seen from Figure 2, the size of the group whose $\hat{\lambda}_i$ is above C is larger than the group of individuals whose true values of λ_i are above C (for C in the upper portion of the distribution).

If we choose the cutoff C = 8 crimes/year, then for the example shown in Figure 2 approximately 20 percent of offenders have their estimate $\hat{\lambda}_i \geq C$. Moreover, the average of the estimates $\hat{\lambda}_i$ for offenders with $\hat{\lambda}_i \geq C$ is 10.0 crimes per year. But it is erroneous to think that the average number of crimes prevented per year of incarceration of these offenders is 10.0. Rather, incapacitation effects must be calculated from the <u>true</u> crime commission rates of offenders, which will have a lower average.

In the sections that follow, we shall show how to estimate the distribution of true crime commission rates for offenders whose estimates

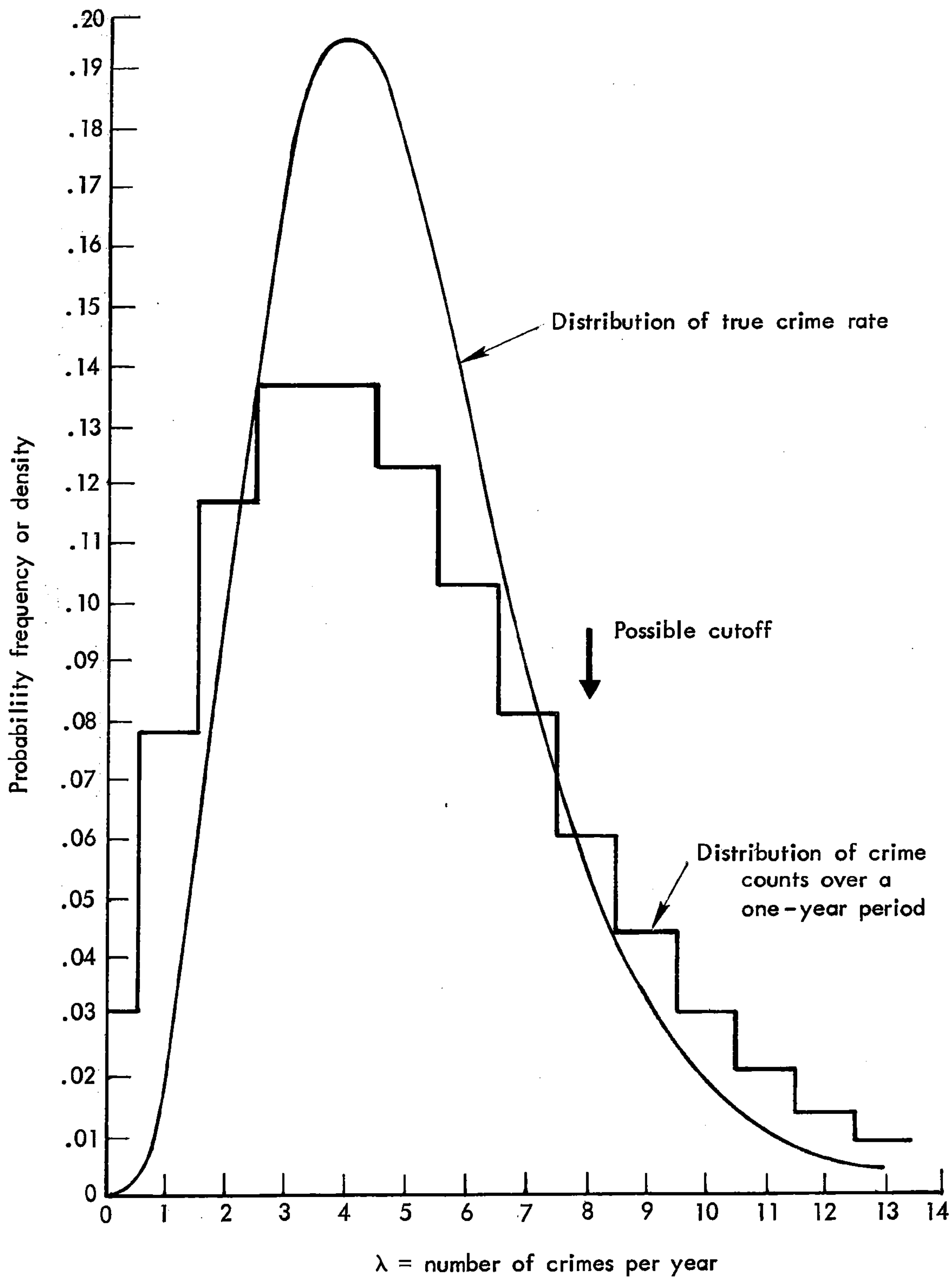

Fig. 2 — Comparison of true and estimated distributions of crime rates: The true crime rate is assumed to have a gamma distribution with mean and variance equal to 5. The estimate is the number of crimes committed during a one-year period.

$\hat{\lambda}_i$ lie above some cutoff. In so doing, we shall elaborate on the above model to take two considerations into account:

- that offenders are not eligible for incapacitation effects at arbitrary times in their careers, but only when they have just been convicted of a crime and are about to be incarcerated; and
- that the street-time period over which offenders' crime commission rates can be measured varies among offenders, since some of them may have begun committing crimes recently and others may have been incarcerated previously. When the length of street time T_i for offender i is allowed to vary with i, it turns out that one can achieve a greater incapacitation effect (at fixed cost) by using an estimate of λ_i other than N_i/T_i. Examples of such improved estimates will be given in the next section.

2. ESTIMATES OF CRIME COMMISSION RATES

We now suppose that offender i commits crimes according to a Poisson process with parameter Λ_i. After each crime the individual independently has probability Q_{1i} of being arrested for the crime and <u>if</u> arrested independently has probability Q_{2i} of being incarcerated for the crime. It is easy to show that his incarcerations occur according to a Poisson process with parameter $\Lambda_i Q_{1i} Q_{2i}$, while his crimes without incarceration occur according to an <u>independent</u> Poisson process with parameter $\Lambda_i(1 - Q_{1i}Q_{2i})$. For the purposes of an incapacitation strategy, we think of offenders as appearing as they are about to be incarcerated. The number of crimes committed by offender i is assumed to be measured over the period since he started committing crimes or since the end of his last incarceration, whichever is later. Thus T_i, the length of the measurement period for offender i, has an exponential distribution with parameter $\Lambda_i Q_{1i} Q_{2i}$. Let N_i be the number of crimes committed by offender i during T_i, excluding the last crime (which led to the incarceration). Then N_i is Poisson distributed with parameter $\lambda_i T_i$ where $\lambda_i = \Lambda_i(1 - Q_{1i}Q_{2i})$. Because the incarceration process and the crime-without-incarceration process are independent, we can condition on T_i in making probability calculations related to N_i.

Note that we have changed the notation slightly, and now the parameter λ_i that we will estimate is for the crime-without-incarceration process. This change was made for technical reasons, namely, to avoid the problem that offender i necessarily committed at least one crime during T_i--the crime that led to his incarceration. This circumstance causes $(N_i + 1)/T_i$ to be an upward-biased estimate of Λ_i. To obtain better estimates of Λ_i, one needs to make additional assumptions about the relationship of $Q_{1i}Q_{2i}$ to Λ_i, while λ_i can be estimated from the data N_i and T_i.

In fact, since the number N_i of crimes without incarceration has a Poisson distribution with parameter $\lambda_i T_i$, we have

$$P(N_i = n | T_i, \lambda_i) = \frac{(\lambda_i T_i)^n}{n!} e^{-\lambda_i T_i}; \qquad n = 0, 1, 2, \ldots . \tag{2}$$

The maximum likelihood estimator of λ_i from the data N_i and T_i is

$$\hat{\lambda}_i = N_i/T_i. \tag{3}$$

From Equation (2), it follows that given T_i and λ_i, $\hat{\lambda}_i$ has expectation $E(\hat{\lambda}_i) = \lambda_i$ and variance $\sigma^2(\hat{\lambda}_i) = \lambda_i/T_i$. In particular, $\hat{\lambda}_i$ is an unbiased estimate of λ_i.

To imitate the variation in λ_i in the real world, we assume that the λ_i's are sampled from a gamma distribution with parameters (α, β). Many empirical distributions can be fit fairly well with a gamma distribution, so this is not a particularly restrictive assumption. Moreover, data collected from self-reports of imprisoned felons are consistent with this assumption [5]. Then the probability density function of λ_i is

$$f_{\alpha,\beta}(\lambda_i) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda_i^{\alpha-1} e^{-\beta\lambda_i} \quad \text{for} \quad \lambda_i > 0 \tag{4}$$

where the gamma function Γ is defined by

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u}\, du, \qquad \alpha > 0.$$

The mean and variance of a gamma distribution are α/β and α/β^2, respectively. The λ_i's being drawn from a probability distribution corresponds to the offenders in our sample being drawn from a larger hypothetical criminal population.

With the assumption that the λ_i's have a distribution _a priori_, the relevant information for interence about λ_i given the data is the _a posteriori_ (or posterior) distribution of λ_i given N_i and T_i. A standard calculation [4, Chap. 9] shows this distribution to be gamma with parameters $(\alpha + N_i, \beta + T_i)$. Therefore, the mean and variance of the posterior distribution of λ_i given N_i and T_i are $(\alpha + N_i)/(\beta + T_i)$ and $(\alpha + N_i)/(\beta + T_i)^2$, respectively. The Bayes estimator of λ_i is defined to be the mean of this posterior distribution, or

$$\hat{\lambda}_i' = \frac{\alpha + N_i}{\beta + T_i}. \tag{5}$$

That is, the Bayes estimator $\hat{\lambda}_i'$ is the expected value of the true λ_i of offender i, given his data N_i and T_i. The Bayes estimator can be written as a weighted average of the prior mean α/β and the maximum likelihood estimator $\hat{\lambda}_i = N_i/T_i$. That is

$$\hat{\lambda}_i' = (1 - w) \frac{N_i}{T_i} + w \cdot \frac{\alpha}{\beta} \tag{6}$$

where

$$w = \frac{\beta}{\beta + T_i}.$$

Note that the longer individual i is observed (the larger T_i) the closer $\hat{\lambda}_i'$ is to the usual estimator $\hat{\lambda}_i = N_i/T_i$. Thus, a criminal whose behavior is observed for a very short time T_i will have his λ_i estimated as being close to the a priori mean of λ_i (namely, α/β), since $w_i = \beta/(\beta + T_i)$ is close to 1. If the parameters α and β are known a priori, perhaps from earlier studies, experience, etc., Equation (6) can be used as an estimator of λ_i. Otherwise the weight w in Equation (6) must itself be estimated from the data. The appendix gives the derivation for an estimate for $\hat{\alpha}$ and $\hat{\beta}$ that is used to get

$$\hat{w}_i = \frac{\hat{\beta}}{\hat{\beta} + T_i}. \tag{7}$$

We then estimate λ_i by

$$\hat{\lambda}_i'' = (1 - \hat{w}_i)\frac{N_i}{T_i} + \hat{w}_i \frac{\hat{\alpha}}{\hat{\beta}} \tag{8}$$

3. ALTERNATIVE INCAPACITATION STRATEGIES

We now envision that cutoff incapacitation strategies are to be based on the estimates of an offender's rate of crimes without incarceration. In one policy, an enhanced sentence would be given to offender i if $N_i/T_i \geq C$, while in the other he would receive an enhanced sentence if $\hat{\lambda}_i' \geq C'$ (or $\hat{\lambda}_i'' \geq C''$). To have equal-cost strategies, C' (or C'') must be chosen in relation to C so that the same proportion of offenders receive enhanced sentences.

To see the difference between the alternatives, consider two individuals with $T_1 = 1$ and $T_2 = 10$, respectively, and assume both have $N_i/T_i = 10$. Suppose $\alpha = 5$ and $\beta = 1$. Then

$$\hat{\lambda}_1' = (1 - w_1)N_1/T_1 + w_1 \cdot 5 = \frac{1}{2} \cdot 10 + \frac{1}{2} \cdot 5 = 7.5,$$

while

$$\hat{\lambda}_2' = (1 - w_2)N_2/T_2 + w_2 \cdot 5 = \frac{10}{11} \cdot 10 + \frac{1}{11} \cdot 5 = 9.5.$$

If we use the estimate $\hat{\lambda}_1 = N_i/T_i$, both individuals are either above or below the cutoff, while using $\hat{\lambda}_i'$ may result in offender 2 receiving an enhanced sentence and offender 1 receiving an ordinary sentence (say if $C' = 8$). Using the Bayes estimators $\hat{\lambda}_i'$ seems fairer, since criminals with high empirical crime rates $\hat{\lambda}_i$ do not receive enhanced sentences _unless_ they have a sufficiently large value of T_i--they have enough of a track record.

The Bayes strategy will also usually be preferable because it has a larger incapacitation effect. However, since the strategies are based on the crime-without-arrest process, we can only assert in general that the expected number of crimes without incarceration prevented by the Bayes strategy is larger than for the alternative. To see this, let s_1 and s_2 be the two sentence lengths ($s_1 < s_2$), and let $X = \{i:\hat{\lambda}_i \geq C\}$ and $X' = \{i:\hat{\lambda}'_i \geq C'\}$. The cutoffs C and C′ are chosen so that there are the same number of offenders in X as in X′. Then the effect of the strategy based on the estimate $\hat{\lambda}_i$ is measured by

$$Z = \sum_{i \notin X} \lambda_i s_1 + \sum_{i \varepsilon X} \lambda_i s_2$$

$$= (\Sigma\lambda_i)s_1 + \sum_{i \varepsilon X} \lambda_i (s_2 - s_1),$$

while the effect of the Bayes strategy is measured by

$$Z' = (\Sigma\lambda_i)s_1 + \sum_{i \varepsilon X'} \lambda_i (s_2 - s_1).$$

Hence

$$Z' - Z = (s_2 - s_1)\Big(\sum_{i \varepsilon X'-X} \lambda_i - \sum_{i \varepsilon X-X'} \lambda_i\Big).$$

Here $X' - X = \{i \varepsilon X' : i \notin X\} = \{i:\hat{\lambda}'_i \geq C' \text{ and } \hat{\lambda}_i < C\}$.

Since we do not know the true crime rate of offender i, but only his N_i and T_i, the <u>a posteriori</u> expected value of λ_i is $\hat{\lambda}'_i$, and the expected value of the difference is

$$E(Z - Z') = (s_2 - s_1)\Big(\sum_{i \varepsilon X'-X} \hat{\lambda}'_i - \sum_{i \varepsilon X-X'} \hat{\lambda}'_i\Big). \quad (9)$$

If M is the number of offenders in the set X′ − X (which necessarily equals the number in $X - X'$), the first sum in Equation (9) is greater than or equal to MC', while the second sum is less than MC', so the difference is positive. This shows that the expected effect for the Bayes strategy is larger than the expected effect for the other strategy. (The same argument shows that the Bayes strategy reduces the expected number of crimes without incarceration more than any equal-cost strategy that is based on knowing N_i and T_i for each offender.)

4. THE DISTRIBUTION OF TRUE CRIME RATES

A selective incapacitation strategy of the type we are discussing is specified by giving the value of the cutoff C or C' and the two sentence lengths s_1 and s_2. The operationally interesting effects of the strategy are then described by the distribution of the true Λ_i's for those offenders who are given the enhanced sentence. Typically, one cannot wait until all the offenders to be sentenced are in hand before assigning sentences to any of them, and therefore the cutoff must be selected in advance. So we shall show how to select the cutoff and estimate the distribution of true crime rates by making some assumption about the probability distribution of the T_i's.

In our model T_i has an exponential distribution with parameter $G_i = \Lambda_i Q_{1i} Q_{2i}$. For the calculation that follows we shall assume that $G_i = G$, the same for all offenders, and then later discuss other possibilities. Then the probability density of T_i is

$$g_{T_i}(t) = G\, e^{-Gt}. \qquad (10)$$

Under this assumption, an offender's Λ_i is simply related to λ_i by $\Lambda_i = \lambda_i + G$. Thus, if we estimate the distribution of λ_i for offenders above the cutoff, we automatically know the distribution of Λ_i. While the assumption is not to be taken too seriously, it does illustrate how the strategies behave when the values of T_i vary over a wide range.

Now define

$$V(C, \lambda) = P(\hat{\lambda}_i \geq C|\lambda)$$

and similarly

$$V'(C', \lambda) = P(\hat{\lambda}_i' \geq C'|\lambda).$$

Then we are interested in the probability density for the true value of λ_i, given that $\hat{\lambda}_i$ is above the cutoff, which is

$$f_{\alpha,\beta}(\lambda|\hat{\lambda} \geq C) = \frac{1}{K} f_{\alpha,\beta}(\lambda)V(C, \lambda), \tag{11}$$

where

$$K = \int f_{\alpha,\beta}(\lambda)V(C, \lambda)d\lambda. \tag{12}$$

The density $f_{\alpha,\beta}(\lambda|\hat{\lambda}' \geq C')$ is calculated similarly from $V'(C', \lambda)$.

The normalizing constant K is equal to the fraction of offenders who receive enhanced sentences under the policy with cutoff C, so its desired value is known if one specifies the desired cost of the selective incapacitation policy. Hence we need a formula relating C to K in order to select the value of C that yields the desired cost.

To evaluate Equation (11), we first evaluate $V(C, \lambda)$.

$$V(C, \lambda) = P(N_i/T_i \geq C|\lambda)$$

$$= \sum_{n=0}^{\infty} P(N_i = n,\ T_i \leq n/C|\lambda)$$

$$= \sum_{n=0}^{\infty} \int_0^{n/C} \frac{(\lambda t)^n}{n!} e^{-\lambda t}\ G\ e^{-Gt}\ dt$$

$$= \sum_{n=0}^{\infty} G \frac{\lambda^n}{(\lambda + G)^n} \int_0^{n/C} \frac{[(\lambda + G)t]^n}{n!} e^{-(\lambda+G)t}\ dt$$

$$= \sum_{n=0}^{\infty} G \frac{\lambda^n}{(\lambda + G)^{n+1}} \int_0^{\frac{n(\lambda+G)}{C}} \frac{y^n}{n!} e^{-y}\ dy$$

$$= \sum_{n=0}^{\infty} G \frac{\lambda^n}{(\lambda + G)^{n+1}} \Gamma_I(\frac{n(\lambda + G)}{C},\ n + 1) \tag{13}$$

where $\Gamma_I(x,\ n + 1)$ is the incomplete gamma function

$$\Gamma_I(x,\ n + 1) = \int_0^x \frac{y^n}{n!} e^{-y}\ dy.$$

Then Equation (11) becomes

$$f_{\alpha,\beta}(\lambda|\hat{\lambda} \geq C) = \frac{1}{K} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1}\ e^{-\beta\lambda} \sum_{n=0}^{\infty} G \frac{\lambda^n}{(\lambda + G)^{n+1}} \Gamma_I(\frac{n(\lambda + G)}{C},\ n + 1) \tag{14}$$

and the normalization constant K can be simplified as follows.

$$K = \int_0^\infty d\lambda \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1}\ e^{-\beta\lambda} \sum_{n=0}^{\infty} \int_0^{n/C} dt \frac{(\lambda t)^n}{n!} G\ e^{-(G+\lambda)t}\ dt$$

$$= G \frac{\beta^\alpha}{\Gamma(\alpha)} \sum_{n=0}^{\infty} \int_0^{n/C} dt\ e^{-Gt}\ t^n \int_0^\infty d\lambda \frac{\lambda^{\alpha+n-1}}{n!} e^{-\lambda(t+\beta)}$$

$$K = G \frac{\beta^{\alpha}}{\Gamma(\alpha)} \sum_{n=0}^{\infty} \int_{0}^{n/C} dt\, e^{-Gt} \frac{t^{n}}{(t+\beta)^{\alpha+n}} \int_{y=0}^{\infty} dy\, \frac{y^{\alpha+n-1}}{n!} e^{-y}$$

$$= G\, \beta^{\alpha} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\Gamma(n+1)} \int_{0}^{n/C} dt\, e^{-Gt} \frac{t^{n}}{(t+\beta)^{\alpha+n}}. \tag{15}$$

The result for $f_{\alpha,\beta}(\lambda|\hat{\lambda}' \geq C)$ is obtained by replacing n/C by $\frac{n+\alpha}{C'} - \beta$ in Equations (14) and (15).

Although these expressions appear lengthy, it is straightforward to evaluate them with a computer program. Figure 3 gives an example in which $\alpha = 5$, $\beta = 1$, $G = 0.2$, and $K = 0.2$. (That is, the underlying distribution of true λ is the same as in Figure 2, the average waiting time between incarcerations is 5 years, and 20 percent of offenders receive the enhanced sentence.) The cutoff C for $\hat{\lambda}_i$ corresponding to $K = 0.2$ was found by calculating Equation (15) for various values of C, with the result that $C = 7.05$. Similarly, the cutoff C' for $\hat{\lambda}_i'$ was found to be 6.40. Equation (14) and its analog for $f(\lambda|\hat{\lambda}' \geq C')$ were then evaluated to get the densities shown in the figure.

Two characteristics of incapacitation strategies based on estimated crime commission rates are illustrated in Figure 3:

1. Using either estimation procedure, some of the presumed high rate offenders have low values of their true λ. If it were possible to select the highest 20 percent of offenders according to their <u>true</u> crime commission rates, their average λ would be 8.44. But using estimated rates, the offenders selected for enhanced sentences have an average λ around 7.6. (This is the mean of the densities shown in Figure 3.)

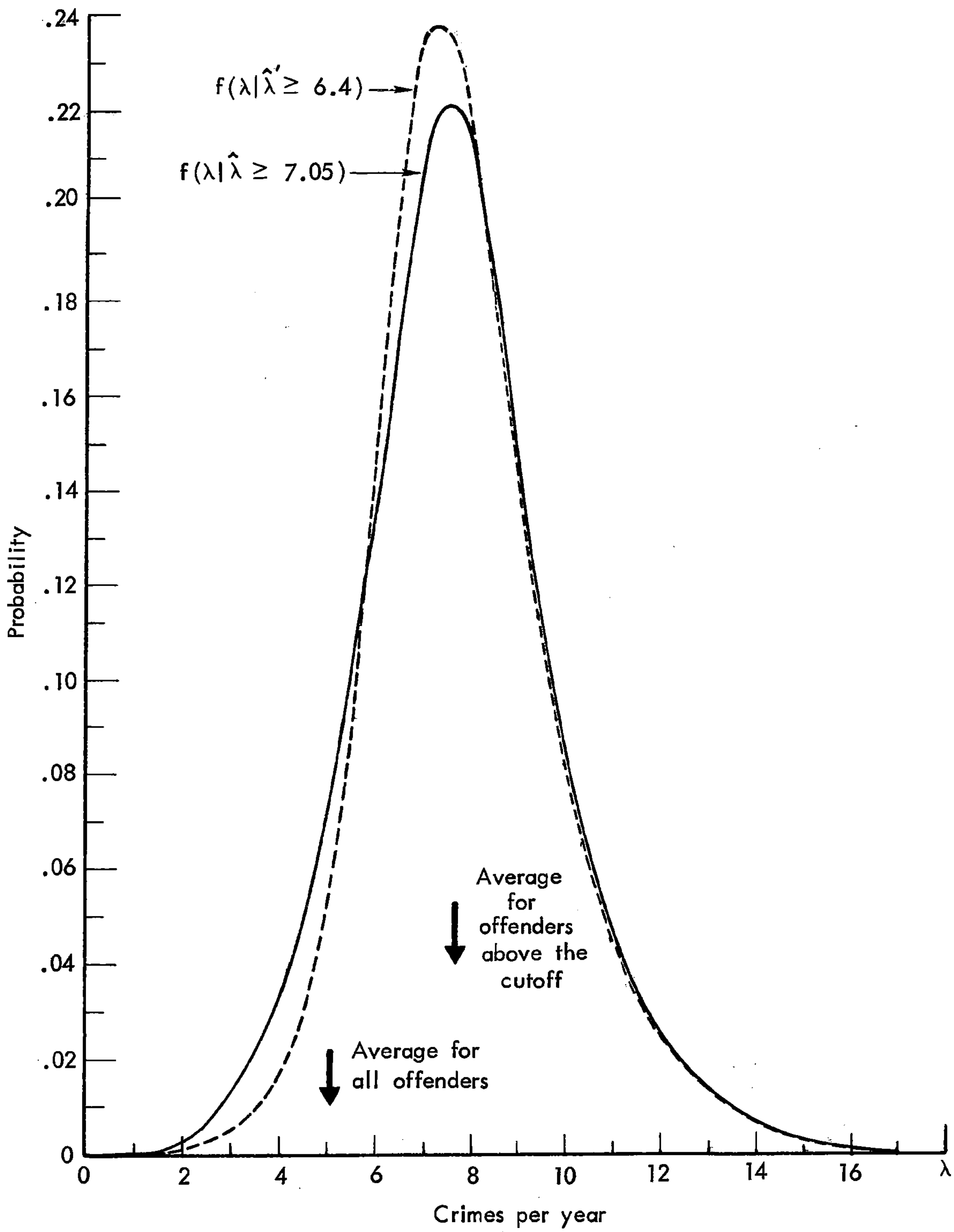

Fig. 3 — Two probability densities for the true λ of offenders, corresponding to two methods for estimating λ and setting a cutoff for the estimate. In each case, 20 percent of offenders have their estimated λ above the specified cutoff.

The average λ for those selected by the Bayes strategy is slightly higher than for those selected according to $\hat{\lambda} \geq C$, but the difference is not substantial in this example.

2. The Bayes strategy is slightly fairer, in that it is less likely to give enhanced sentences to offenders whose true λ is below the average for all offenders (5 crimes per year).

5. DISCUSSION

Although our model is overly simplified and the results have been shown for only one set of parameters, we believe that the above two conclusions about selective incapacitation strategies will hold up in more complicated and realistic models.

There are several directions in which our models can be generalized or the assumptions changed to make the models more realistic. First, the distribution of T_i, the time of observation of offender i, may be other than exponential. For example, T_i could be the accumulated street time since the beginning of the offender's career. Rather than assuming an exponential distribution for T_i, a gamma distribution would then be more appropriate. The formulas get only slightly more complicated in this case, so our approach can easily be adapted.

Second, the assumption that $G_i = \Lambda_i Q_{1i} Q_{2i}$ is a constant was made primarily for analytical convenience. That is, we assume that Q_{1i}, Q_{2i}, and Λ_i vary in such a way that G_i is the same for everybody and thus the value of T_i yields no information about $\Lambda_i = \lambda_i + G_i$. This assumption can be altered in a number of ways. One analytically attractive way is to let $q_i = Q_{1i} Q_{2i} \equiv q$ be the same for everyone and thus have Λ_i distributed according to $\Gamma(\alpha, (1 - q)\beta)$. The estimation problem becomes somewhat more difficult in this model, since T_i now contains information about λ_i. That is, one could develop an improved Bayes estimate of λ_i with this added information. Then the calculation of the true distribution of Λ for offenders above the cutoff would be similar to the one given here.

Third, we have assumed that N_i is the number of crimes without incarceration. In many situations, we will have only the number of crimes without arrests. To estimate the effect of an incapacitation strategy in this case, some assumption must be made about the relationship between Q_{1i} and the other parameters. Then the analysis goes through in a straightforward manner.

Finally, we have ignored information about individual i's crime commission rate that might be contained in variables like his age, race, and other background characteristics. Notwithstanding all of these simplifications in our models, two general conclusions remain: (1) care must be taken to account for the effects of estimating individual crime rates from data in assessing the benefits from incapacitation policies, and (2) taking into account the variation in true individual crime rates when estimating individual crime rates from data can lead to improved performance of the resulting incapacitation policy.

APPENDIX

We use a method-of-moments approach to devise empirical Bayes estimates of the λ_i's when λ has a gamma distribution with parameters α and β. This is similar to the method of Carter and Rolph [2,3] for the normal distribution. Let $X_i = N_i/T_i$. Suppose there are M individuals. For any $\gamma = (\gamma_1, \ldots, \gamma_M)$, with

$$\sum_1^M \gamma_i = 1, \text{ and } \gamma_i \geq 0,$$

and for any $\delta = (\delta_1, \ldots, \delta_M)$ with $\delta_i \geq 0$, define

$$\bar{X}(\gamma) = \sum_{i=1}^M \gamma_i X_i \tag{A.1}$$

and

$$S(\delta, \gamma) = \sum_{i=1}^M \delta_i (X_i - \bar{X}(\gamma))^2. \tag{A.2}$$

Now the marginal moments of X_i are

$$E(X_i) = E\{E(X_i|\lambda_i)\} = E(\lambda_i) = \alpha/\beta \tag{A.3}$$

$$\begin{aligned}
\operatorname{Var}(X_i) &= E\{\operatorname{Var}(X_i|\lambda_i)\} + E\{E(X_i) - E(X_i|\lambda_i)\}^2 \\
&= E\left(\frac{\lambda_i}{T_i}\right) + E\left(\frac{\alpha}{\beta} - \lambda_i\right)^2 \\
&= \frac{\alpha}{\beta T_i} + \frac{\alpha}{\beta^2} \\
&= \frac{\alpha}{\beta^2}\left(\frac{\beta + T_i}{T_i}\right) .
\end{aligned} \tag{A.4}$$

Now if $\delta_i(\beta) = T_i/(\beta + T_i)$ and $\gamma_i(\beta) = \delta_i(\beta)/\left(\sum_{i=1}^M \delta_i(\beta)\right)$ then

$$E(S(\delta(\beta), \gamma(\beta)) = (M - 1)\frac{\alpha}{\beta^2} \tag{A.5}$$

and

$$E(\bar{X}(\gamma)) = \frac{\alpha}{\beta} . \tag{A.6}$$

We use Equations (A.5) and (A.6) to devise an iterative scheme for estimating α and β. Define for any b

$$\delta_i(b) = \frac{T_i}{b + T_i} , \quad \gamma_i(b) = \frac{\delta_i(b)}{\sum_{i=1}^{M} \delta_i(b)}$$

with $\delta(b)$ and $\gamma(b)$ being the vectors of components. Define $\hat{\beta}$, $\hat{\alpha}$ to be the solutions to the equations:

$$\bar{X}(\gamma(\hat{\beta})) = \hat{\alpha}/\hat{\beta} \tag{A.7}$$

and

$$S(\delta(\hat{\beta}),\gamma(\hat{\beta})) = (M - 1) \frac{\hat{\alpha}}{\hat{\beta}^2} . \tag{A.8}$$

A solution $(\hat{\alpha}, \hat{\beta})$ of these equations satisfies

$$\hat{\beta} = \frac{(M-1)\bar{X}(\gamma(\hat{\beta}))}{S(\delta(\hat{\beta}),\gamma(\hat{\beta}))} . \tag{A.9}$$

We use Equation (A.9) to devise an iterative solution to Equation (A.7) and Equation (A.8) as follows. Initially, set $\gamma_i = \gamma_i(0) = 1/M$ for all i. From Equation (A.9) define

$$\hat{\beta}_1 = \frac{(M-1)\bar{X}(\gamma(0))}{S(\delta(0),\gamma(0))} .$$

Repeating, let

$$\hat{\beta}_2 = \frac{(M-1)\bar{X}(\gamma(\hat{\beta}_1))}{S(\delta(\hat{\beta}_1),\gamma(\hat{\beta}_1))}$$

and in general

$$\hat{\beta}_{i+1} = \frac{(M-1)\bar{X}(\gamma(\hat{\beta}_i))}{S(\delta(\beta_i),\gamma(\hat{\beta}_i))} .$$

When $\hat{\beta}_{i+1}$ and $\hat{\beta}_i$ are sufficiently close, let $\hat{\beta}$ be this common value. Then from Equation (A.7)

$$\hat{\alpha} = \hat{\beta}\, \bar{X}(\gamma(\hat{\beta})). \tag{A.10}$$

The estimates of Equation (A.10) are then used to get the empirical Bayes estimator as

$$\hat{\lambda}_i = (1 - \hat{w}_i)\,\frac{N_i}{T_i} + \hat{w}_i \bar{X}(\gamma(\hat{\beta}))$$

where

$$\hat{w}_i = \frac{\hat{\beta}}{T_i + \hat{\beta}}\,.$$

ACKNOWLEDGMENT

Computer programming for the illustrative examples in this paper was ably performed by Suzanne Polich.

This paper was prepared under Grant Number 77-NI-99-0053 from the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice. Points of view or opinions stated in this paper are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

REFERENCES

1. A. Blumstein, J. Cohen, and D. Nagin (Eds.), Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates, National Academy of Sciences, Washington, D.C., 1978.
2. G. M. Carter and J. E. Rolph, New York City Fire Alarm Prediction Models: I. Box-Reported Serious Fires, R-1214-NYC, The Rand Corporation, Santa Monica, 1973.
3. -----, "An Empirical Bayes Approach to Estimating Fire Alarm Probabilities," J. Amer. Statist. Assoc., 69, 880-885 (1974).
4. M. H. DeGroot, Optimal Statistical Decisions, McGraw Hill, New York, 1970.
5. H. Stambul, M. Peterson, and S. Polich, Doing Crime: A Survey of California Prison Inmates, R-2200-DOJ, The Rand Corporation, Santa Monica, forthcoming.

RAND/P-6116 SELECTIVE INCAPACITATION STRATEGIES BASED ON ESTIMATED CRIME RATES Chaiken & Rolph