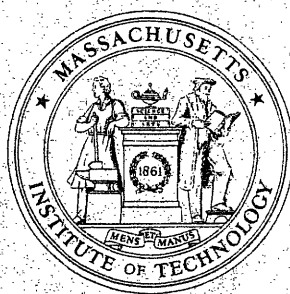


# OPERATIONS RESEARCH CENTER

working paper



**MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY**

BAYESIAN ANALYSIS OF CRIME RATE  
CHANGES IN BEFORE-AND-AFTER EXPERIMENTS

by

Thomas R. Willemain

OR 075-78

June 1978

Prepared under Grant Number 78NI-AX-0007 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 document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

## Table of Contents

	Page
Abstract. . . . .	i
Acknowledgements. . . . .	ii
1. Introduction. . . . .	1
2. Methodology for Difference in Crime Rates . . . . .	3
3. Application to Nashville Data . . . . .	7
4. Methodology for Ratio of Crime Rates. . . . .	12
References. . . . .	16

## Abstract

Bayesian analyses are developed for data consisting of counts of crimes before and after the introduction of an experimental crime control program. It is argued that Bayesian analysis is superior to conventional significance testing in that the entire probability distribution of the estimated change in crime rate can be displayed. Furthermore, the new Bayesian methods developed here are more appropriate than available Bayesian approaches to changes in time series because they make explicit use of the discreteness of the crime count data. The analysis assumes that crimes occur in the before and after periods according to homogeneous Poisson processes with possibly differing rates. This assumption is verified for the case of the Nashville, Tennessee experiment in saturation levels of police patrol. Application of the new Bayesian methods is illustrated by a re-analysis of the Nashville data.

### Acknowledgements

Professor Richard Larson offered suggestions for substantial improvements in the manuscript, which was ably typed by Kathleen Sumera. Any errors remain the responsibility of the author.

## 1. Introduction

A common form of program evaluation in the field of criminal justice involves a "before-and-after" comparison of crime rates in an area targeted by a new crime control program. This paper presents two Bayesian methodologies for analysing data which consist of two sequences of counts of random events; for instance, data on the daily number of crimes during an interval consisting of a "baseline" followed by a "trial" period. Section 2 of the paper develops the mathematical results when the performance measure of interest is the difference in the crime rates. Section 3 applies the method to a re-analysis of the Nashville experiment on saturation levels of police patrol [1]. Section 4 addresses the case in which the comparison is made in terms of the ratio of the crime rates.

The methodology is based on a mathematical model of crime occurrence which holds that the number of crimes in a given interval of time is a Poisson variable. The Poisson model is well-supported by the Nashville data. This analytic approach has certain advantages over both classical and Bayesian alternatives. The original analysis of the Nashville data by Schnelle et al. used student's t-test to examine the statistical significance of the difference in crime rates between baseline and trial periods. There are two drawbacks to their approach: first, the t-test assumes that daily crime counts vary continuously according to a Gaussian distribution, whereas in reality the counts are discrete and typically number only a few crimes per day, making the Gaussian a poor approximation; second, the statistical significance

of the difference in crime rates is less relevant to policy than the distribution of the magnitude of the change in crime rate - reporting only a significance result suppresses much of the information in the data and misses an opportunity to present the results in a more readily interpreted format. A Bayesian analysis would be more useful in presenting the distribution of the change in crime rates, but conventional Bayesian approaches to shifts in the level of a time series are based on the assumption that the variable of interest is continuous, with Gaussian increments from one time to another [2]. An approach which recognizes and exploits the discreteness of the data would be better matched to the problem. Such an approach is developed below; it assumes that crimes occur in the baseline and trial periods as Poisson processes with (possibly) different rates and determines the posterior distributions of the difference in crime rates (section 2) and the ratio of crime rates (section 4).

## 2. Methodology for Difference in Crime Rates

The typical data series consists of counts of crimes committed during a baseline reference period followed by a trial period during which the new crime control program is applied (see Figure 1). During the  $i^{\text{th}}$  interval of observation (assume that the basic interval is a day) a total of  $n_i$  crimes are committed. The sequence  $\{n_i\}$ , divided into baseline and trial counts, is the basic data input to the analysis; the output is an estimate of  $h(\Delta\lambda)$ , the distribution of the change in crime rate.

Let:

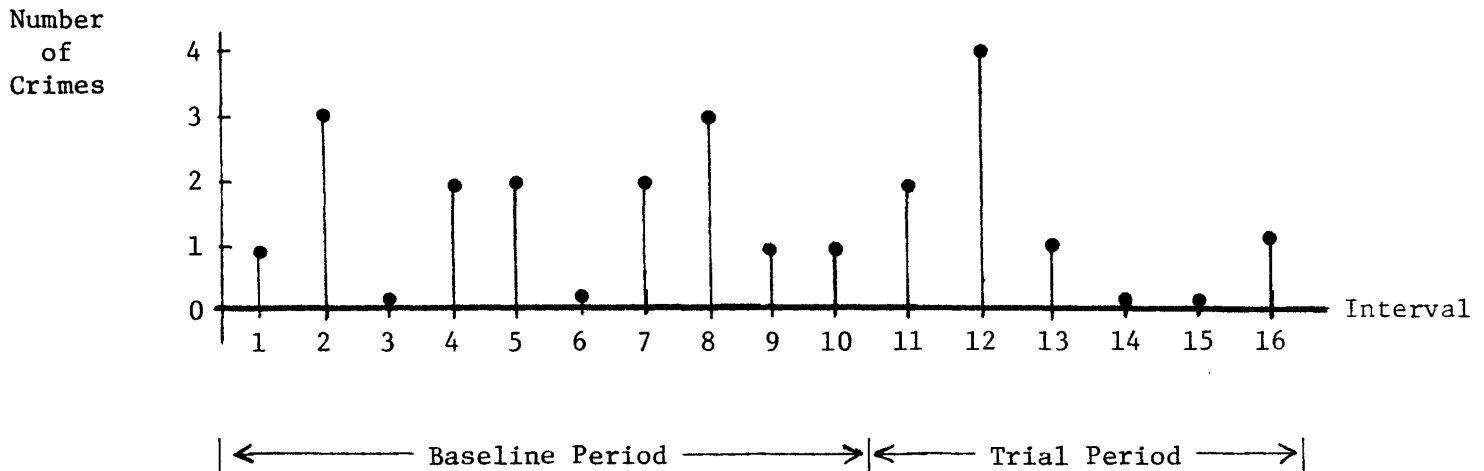
- $\lambda_b$  = crime rate during baseline period
- $\lambda_t$  = crime rate during trial period
- $\Delta\lambda = \lambda_t - \lambda_b$  = change in crime rate
- $g(\lambda_b)$  = posterior distribution of  $\lambda_b$
- $f(\Delta\lambda)$  = posterior distribution of  $\Delta\lambda$

In a Bayesian perspective, the crime rates  $\lambda_b$  and  $\Delta\lambda$  are treated as random variables whose different possible values are supported more or less strongly by the crime count data. The relative credibilities of these different values are expressed in the posterior distributions  $g(\lambda_b)$  and  $f(\Delta\lambda)$ .

Consider the distribution  $f(\Delta\lambda)$ . We begin by finding the distribution of the change in crime rate  $\Delta\lambda$  conditional on a particular value of baseline crime rate  $\lambda_b$ , then we uncondition by integrating over the posterior distribution of  $\lambda_b$ . For any given value of  $\lambda_b$ , the change in crime rate could be any number greater than or equal to  $-\lambda_b$ : this recognizes both that the trial period crime rate  $\lambda_t = \lambda_b + \Delta\lambda$  cannot be negative and that it might possibly be greater than  $\lambda_b$ . The latter situation might correspond to a trial program



Figure 1: Data and Variable Definitions



$I_b$  = interval of baseline data (here = 10)

$I_t$  = interval of trial data (here = 6)

$N_b$  = number of crimes during baseline interval (here = 15)

$N_t$  = number of crimes during trial interval (here = 8)

$i$  = index denoting  $i^{\text{th}}$  interval

$n_i$  = number of crimes in  $i^{\text{th}}$  interval (e.g.,  $n_7 = 2$ )

$\lambda_b$  = mean rate of occurrence of crimes during baseline period

$\lambda_t$  = mean rate of occurrence of crimes during trial period

$\Delta\lambda = \lambda_t - \lambda_b$  = change in crime rate

$g(\lambda_b)$  = posterior distribution of  $\lambda_b$

$f(\Delta\lambda)$  = posterior distribution of  $\Delta\lambda$

which, while reducing total crime, nevertheless increases total reported crime; it is reported crimes which constitute the data base for the program evaluation. We will make the conservative choice of assuming that the conditional prior distribution of the change in crime rate,  $f(\Delta\lambda | \lambda_b)$ , is non-zero and uniform over the range  $-\lambda_b \leq \Delta\lambda < \infty$ . Given this (improper) diffuse prior, the conditional posterior distribution of the change in crime rate is proportional to the conditional likelihood of observing the sequence of counts  $\{n_{I_b+1}, n_{I_b+2}, \dots, n_{I_b+I_t}\}$  during the trial period, generated by a Poisson process with rate  $\lambda_t = \lambda_b + \Delta\lambda$  [3]:

$$f(\Delta\lambda | \lambda_b) \propto \prod_{i=I_b+1}^{I_b+I_t} \exp[-(\lambda_b + \Delta\lambda)] (\lambda_b + \Delta\lambda)^{n_i} / n_i!$$

$$\propto \exp[-(\lambda_b + \Delta\lambda)I_t] (\lambda_b + \Delta\lambda)^{N_t} \quad \Delta\lambda \geq -\lambda_b. \quad (1)$$

Thus the unconditional posterior distribution of the change in crime rate,  $f(\Delta\lambda)$ , can be determined by integrating (1) over the possible values of baseline crime rate  $\lambda_b$ . Therefore the next task is to determine the posterior distribution of the baseline crime rate,  $g(\lambda_b)$ .

We begin with the choice of a prior distribution. We might either choose a diffuse prior or follow suggestions in the literature [3,4] and choose a logarithmically flat prior (i.e., a prior  $\propto \lambda_b^{-1}$ ). We opt for the latter, although with reasonably large counts of crimes  $N_b$  in the baseline period (as in the Nashville case) this choice makes little practical difference. Since the baseline data consist of the sequence of Poisson counts  $\{n_1, n_2, \dots, n_{I_b}\}$ , the posterior

distribution of  $\lambda_b$  becomes

$$g(\lambda_b) \propto \lambda_b^{-1} \prod_{i=1}^{I_b} \exp[-\lambda_b] \lambda_b^{n_i/n_i!}$$

$$\propto \exp[-\lambda_b I_b] \lambda_b^{N_b-1} \quad \lambda_b \geq 0 \quad (2)$$

We can now combine (1) and (2) to obtain the unconditional posterior distribution of the change in crime rate

$$f(\Delta\lambda) \propto \int_{\lambda_b=L(\Delta\lambda)}^{\infty} \exp[-(\lambda_b + \Delta\lambda)I_t] (\lambda_b + \Delta\lambda)^{N_t} \exp[-\lambda_b I_b] \lambda_b^{N_b-1} d\lambda_b \quad (3)$$

where

$$L(\Delta\lambda) = \max[0, -\Delta\lambda] \quad (4)$$

The integral (3) can be reduced to a simple finite sum when  $L(\Delta\lambda) = 0$  but becomes a finite sum of incomplete gamma functions when  $L(\Delta\lambda) = -\Delta\lambda$ . We will solve (3) by numerical integration. The special case  $\Delta\lambda = 0$  can be solved analytically using the fact that

$$\int_{x=0}^{\infty} x^n \exp(-ax) dx = \frac{n!}{a^{n+1}} \quad (5)$$

Thus

$$f(0) \propto \frac{(N_t + N_b - 1)!}{(I_b + I_t)^{N_t + N_b}} \quad ; \quad (6)$$

this value provides a useful check on the accuracy of the numerical integration.

### 3. Application to Nashville Data

The Nashville experiment set out to determine the impact on Part I crimes (robbery, larceny, burglary, motor-vehicle theft, forcible rape, aggravated assault and homicide) of saturating small areas with police patrol cars. Four zones with high crime rates were selected for study. The first two zones, referred to as "Day Patrol One" and "Day Patrol Two", each received in turn 10 days of saturation patrol during the 9AM - 5PM shift. Then the third and fourth zones - "Night Patrol One" and "Night Patrol Two" - each received 10 days of saturation patrol during the 7PM - 3AM shift. Thus the Nashville experiment consisted of four successive trials of 10 days each, two with daytime saturation and two with nighttime saturation. Crime counts were obtained in each of the four zones before, during, and after the 10 day trial periods. Briefly, the results obtained by Schnelle et al. [1] were that no statistically significant changes in crime rate were observed in the two daytime trials, but significant decreases were observed in both nighttime trials.

We will re-analyze the Nashville data using the Bayesian methodology developed above. The first step is to confirm that the data are in fact well described by a Poisson model. On the assumption that the impacts of saturation patrol do not persist after patrol returns to normal levels, we will not distinguish between crime data from before and after the 10 day trial periods of saturation patrol; rather, we combine these data to form the baselines. The four baseline periods range in duration from 78 to 116 days. Shown in Table 1 are the distributions of crimes per day in each of the four trials. Using the dispersion test [5], the hypothesis that the counts of crime arise

TABLE 1  
 Comparison of Nashville Baseline Data to Poisson Model  
 (source: Schnelle et al. [1])

# CRIMES	# DAYS WITH GIVEN # EVENTS (EXCLUDING SATURATION TRIAL DAYS)									
	"DAY PATROL 1"		"DAY PATROL 2"		"NIGHT PATROL 1"		"NIGHT PATROL 2"		"NIGHT PATROL 2"	
0	26	(26)	35	(36)	33	(36)	31	(26)	31	(28)
1	29	(29)	37	(35)	42	(42)	32	(36)	32	(36)
2	15	(16)	15	(17)	29	(25)	23	(24)	23	(24)
3	4	(6)	5	(5)	11	(10)	10	(11)	10	(10)
4	4	(2)	2	(1)	1	(3)	5	(4)	5	(3)
5	0	(0 <sup>+</sup> )	0	(0 <sup>+</sup> )	0	(1)	1	(1)	1	(1)
6	0		0		0	(0 <sup>+</sup> )	0	(0 <sup>+</sup> )	0	(0 <sup>+</sup> )
<u>7</u>	<u>0</u>		<u>0</u>		<u>0</u>		<u>1</u>		<u>0</u>	
Total days I <sub>b</sub>	78		94		116		103		102	
Total crimes N <sub>b</sub>	87		90		137		140		133	
$\lambda_b = \frac{N_b}{I_b}$	1.115		0.957		1.181		1.359		1.304	
X <sup>2</sup>	82.45		91.73		94.15		130.74		111.65	
d.f.	77		93		115		102		101	
Z	0.47		-0.06		-1.41		1.92		0.77	
Signif. (2 tail)	0.64		0.95		0.16		0.06		0.44	
Serial auto correlation	0.021		-0.013		-0.118		0.022			

Notes:

a)  $X^2 = \text{dispersion test } X^2 = \sum_{i=1}^{I_b} \frac{(n_i - \lambda_b)^2}{\lambda_b} \sim \text{chi square df} = I_b - 1$

for large I<sub>b</sub> refer  $\sqrt{2X^2} - \sqrt{2I_b - 3}$  to Gaussian (0,1)

(R.L. Plackett, Analysis of Categorical Data, pg. 10).

deleting the busy day from "Night 2" is conservative if want to establish a reduction in  $\lambda$ .

Counts in parenthesis are expected Poisson counts, rounded to nearest integer.

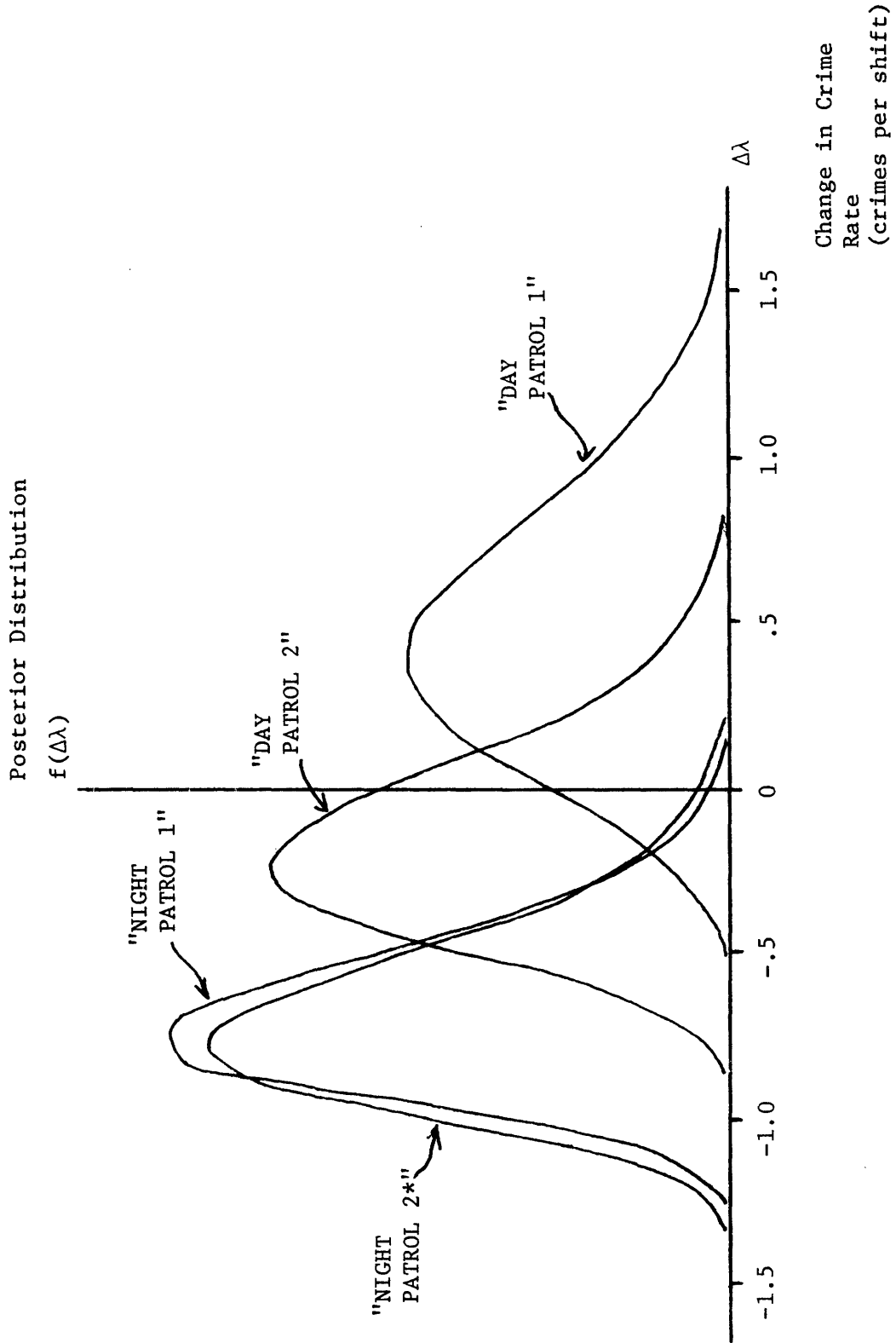
from a Poisson process cannot be rejected at the 0.05 level in any of the four trials. However, the second nighttime trial nearly achieves the 0.05 significance level and deserves special comment. That particular baseline is unusual in that it contains one night shift in which 7 crimes were committed - the highest total for any single shift in the entire study. This exceptional case occurred shortly after the saturation trial and may represent temporally-displaced crime, one very active criminal, or a just a typical fluctuation. If we set aside that datum (forming the fifth column in Table 1, labelled "Night Patrol 2\*"), we not only make a more conservative estimate of the impact of saturation patrol but also achieve a distribution very well described by the Poisson. Testing for serial correlation in each of the four series indicates that the Poisson assumption of independence between crime counts on successive days is also valid, since the serial correlation coefficient never exceeds 0.12 in absolute value. Thus the dispersion tests and tests for serial correlation both confirm the validity of modeling the count of crimes as a Poisson variate.

Given that the Poisson model is valid, we can use (3) to determine the shape of the distribution of the change in crime rate during saturation patrol. The results of the numerical integrations are displayed in Figure 2. As noted by Schnelle et al. [1], the nighttime saturation patrols produced a consistent and clear drop (about -0.8 crimes per shift), whereas the daytime patrols give neither clear nor a consistent indication of impact.

When interpreting these results for substantive purposes, one must be aware of two potential pitfalls. The first is a possible change in the crime reporting behavior of the public. It has been observed in other studies [6] that increasing police presence can lead to an increase in the fraction of crimes reported

Figure 2

Posterior Distribution of Change in Crime Rate  $\Delta\lambda = \lambda_t - \lambda_b$



to police. This problem may not be too serious in the Nashville study, since saturation trials lasted only ten days and since Schnelle et al. found that crimes were almost never reported directly to patrolling officers during either the baseline or trial periods. The second potential pitfall is spatio-temporal displacement of crime, which may be a more serious issue in the Nashville case. Since the trials were so brief, some crime may have merely been delayed, rather than permanently averted. Schnelle et al. were careful to check for spatial displacement of crime, but one wonders whether their test was sufficiently sensitive. Since the drop in crime rate amounts to roughly 1 crime per shift and since there are 33 patrol zones in Nashville, it would be quite possible for the one crime to be displaced and not detected. Even looking at only those zones contiguous to the experimental zones, a displacement of such a small number of crimes would still be hard to detect if the crimes were displaced to a different contiguous zone each night. While our main purpose in this paper is to develop the Bayesian methodology rather than to address the substantive question, and while Schnelle et al. took care to check for displacement, we should be aware of possible problems in interpretation of results such as those in Figure 2.



#### 4, Methodology for Ratio of Crime Rates

Section 2 of this paper developed a methodology for estimating the difference in the crime rates during the baseline and trial periods. In this section we outline an alternative which expresses the experimental impact as a proportionate reduction (or increase) in the crime rate. Our goal is to estimate the ratio of crime rates

$$R = \lambda_t / \lambda_b . \tag{7}$$

As before, we assume that the crime count in each period is a Poisson variate. We note that non-Bayesian methods are available for this problem [7] but again prefer the distributional form of results provided by the Bayesian approach. Values of R less than unity indicate successful crime-reduction programs; the relative credibility of various estimates of R will be readily grasped in the Bayesian framework.

Our approach again is to find the posterior distribution of R conditional on the value of  $\lambda_b$ ; then uncondition by integrating over the posterior distribution of  $\lambda_b$  given the baseline data. As before, we will take this posterior distribution of  $\lambda_b$  to be a gamma distribution

$$g(\lambda_b) \propto \exp(-\lambda_b I_b) \lambda_b^{N_b-1} \quad \lambda_b \geq 0. \tag{8}$$

Of particular interest is the choice of prior distribution for R conditional on the value of  $\lambda_b$ . If we expected that the experimental program would definitely reduce the crime rate, we might use a beta distribution for  $0 \leq R \leq 1.0$ . To allow for the possibility that  $R > 1.0$ , we can use a gamma prior  $\propto R^A \exp[-BR]$ . Note that this choice decouples the prior estimate of R

from the value of  $\lambda_b$ ; it is difficult to imagine a compelling, systematic way to link the value of  $R$  with the value of  $\lambda_b$ . If one prefers a diffuse prior, one can set  $A = B = 0$ . If one wants a roughly bell-shaped prior with mean  $\mu_R$  and standard deviation  $\sigma_R$ , one should choose the integer values closest to

$$A = (\mu_R/\sigma_R)^2 - 1 \quad (9)$$

$$B = \mu/(\sigma_R^2) . \quad (10)$$

With the gamma prior distribution for  $R$ , the conditional posterior distribution

$$f(R|\lambda_b) \propto \prod_{i=I_b+1}^{I_b+I_t} \frac{\exp[-R\lambda_b] (R\lambda_b)^{n_i}}{n_i!} R^A \exp[-BR] \quad (11)$$

$$\propto \exp(-R\lambda_b I_t) (R\lambda_b)^{N_t} R^A \exp[-BR] . \quad (12)$$

Unconditioning using (8) we find

$$f(R) \propto \int_{\lambda_b=0}^{\infty} R^{A+N_t} \exp[-BR] \lambda_b^{N_t} \exp[-RI_t \lambda_b] \lambda_b^{N_b-1} \exp[-\lambda_b I_b] d\lambda_b \quad (13)$$

$$\propto R^{A+N_t} \exp[-BR] \int_{\lambda_b=0}^{\infty} \lambda_b^{N_t+N_b-1} \exp[-(RI_t+I_b)\lambda_b] d\lambda_b \quad (14)$$

$$\propto R^{A+N_t} \exp[-BR] \frac{(N_t + N_b - 1)!}{(RI_t + I_b)^{N_t+N_b}}$$

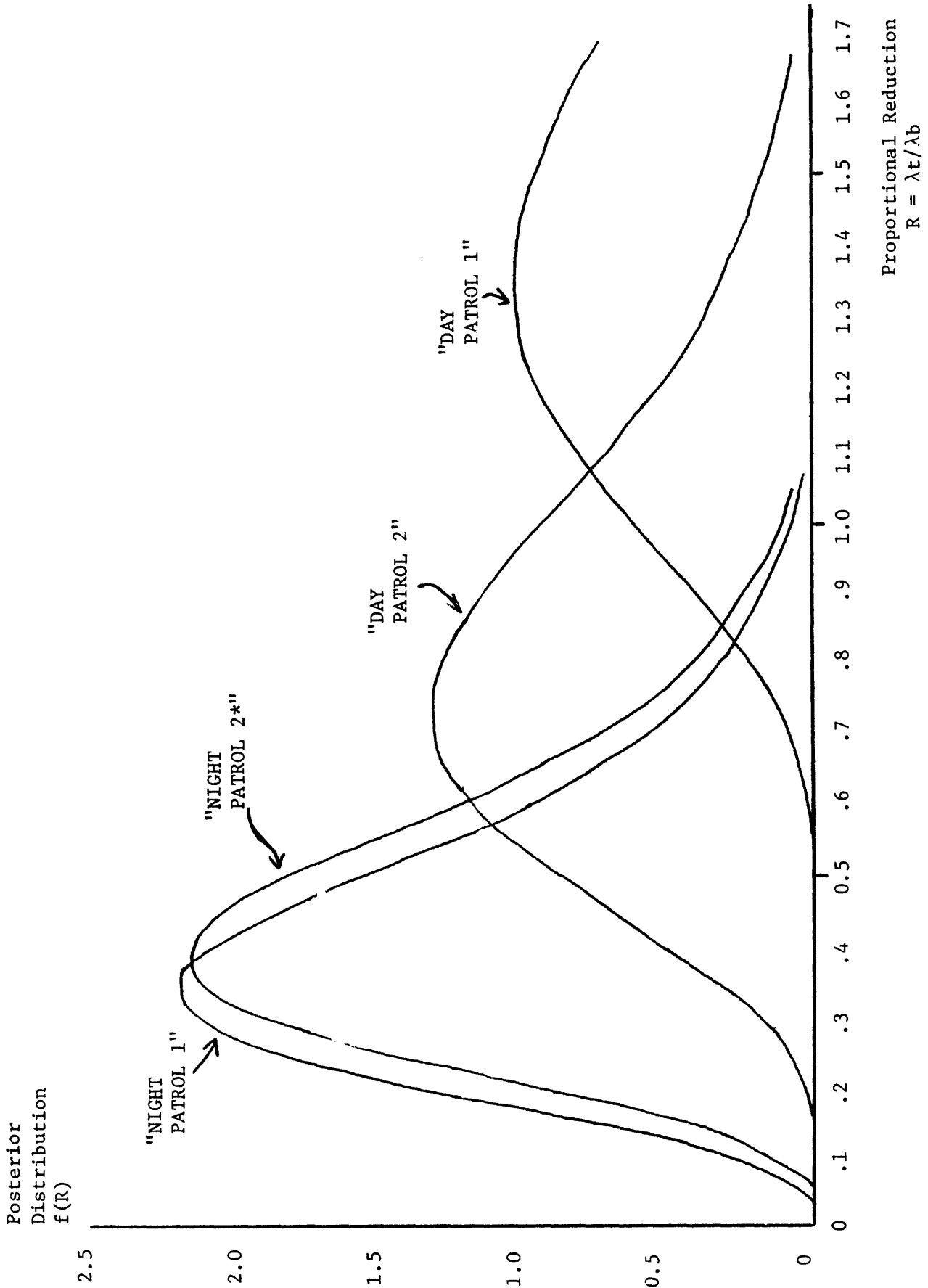
$$f(R) \propto R^{A+N_t} \exp[-BR] [RI_t + I_b]^{-(N_t+N_b)} . \quad (16)$$

In the case of a diffuse prior, (16) specializes to

$$f(R) \propto R^{N_t} [RI_t + I_b]^{-(N_t + N_b)} . \quad (17)$$

The Nashville data have been analyzed using (17); results are shown in Figure 3. As in Figure 2, the nighttime patrols were clearly successful, whereas the daytime results are weak and mixed.

Figure 3  
Posterior Distribution of Proportional Reduction in Crime Rate  $R = \lambda t / \lambda b$



References

1. J.F. Schnelle, R.E. Kirchner, J.D. Casey, P.H. Uselton, and M.P. McNees, "Patrol Evaluation Research: A Multiple Baseline Analysis of Saturation Police Patrolling During Day and Night Hours", Journal of Applied Behavior Analysis, Vol. 10, No. 1, 1977, pp. 33-40.
2. A.F.S. Lee and S.M. Heghinian, "A Shift in the Mean Level in a Sequence of Independent Normal Random Variables - A Bayesian Approach", Technometrics, Vol. 19, No. 4, November 1977, pp. 503-506.
3. S. Schmitt, Measuring Uncertainty: An Elementary Introduction to Bayesian Statistics (Addison-Wesley, Reading Mass., 1969).
4. C. Villegas, "On the Representation of Ignorance", Journal of the American Statistical Association, Vol. 72, No. 359, September 1977, pp. 651-654.
5. R.L. Plackett, The Analysis of Categorical Data (Griffin: London, 1974).
6. S. G. Chapman, Police Patrol Readings (Charles C Thomas, Springfield, Ill., 1972), pp. 342-352.
7. D.R. Cox and P.A.W. Lewis, The Statistical Analysis of Series of Events (Methuen: London, 1966).