

Past Intensity of a Terminated Poisson Process

S. M. Pollock
The University of Michigan

R. L. Farrell
Vector Research, Incorporated

Technical Report 83-14

Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109

July, 1983

Past Intensity of a Terminated Poisson Process

S. M. Pollock
The University of Michigan

R. L. Farrell
Vector Research, Incorporated

INTRODUCTION

One rule for the termination of a stochastic point process is to stop the process when the number of events within a window of length w exceeds a threshold number k . Such a behavior was hypothesized in [1] in an attempt to develop a model that might represent the decision-making of judges when facing a juvenile delinquent with a record of "police contact" events extending into the past.

An interesting question arises when one wishes to look at the past behavior of a collection of processes terminated in this way. This was examined in [1] in an applied (criminal justice) setting, but unfortunately the mathematical analysis given there is flawed in not having taken into account the true conditional nature of the process before it is terminated, given that it indeed does become terminated at some time t . In this paper we present an analysis that rectifies this flaw.

There are a broad variety of applied situations where an analyst is faced with a stream of data from a controlled process and wishes to analyze the parameters of the uncontrolled underlying process. Any such analyses, whether in the criminal justice problems that originated the current work or in other quality control problems, require an understanding of the relation between the unobservable parameters of the

uncontrolled process and the observable behavior of the controlled process. In many instances apparent facts about the underlying process, derived from observation of the controlled process, may actually be related to the control mechanism and not the underlying process.

FUNDAMENTAL PROBLEM

Consider a point process commencing at time 0 -- that is, with no points in $(-\infty, 0]$ -- and an associated stopping rule that terminates the process at the first time t that there are k events within the interval $[t-w, t]$. Let the stopping time be the random variable T , with density function $f(t)$. For example, if the process is Poisson with rate λ , then when $k = 1$, $f(t)$ is trivially $\lambda e^{-\lambda t}$. However, if $k \geq 2$, even in this simple case $f(t)$ becomes an extremely complicated function (the case of $k = 2$ is discussed in detail below).

We wish to find the rate of a stopped process y time units into the "past", e.g., previous to being stopped, but of course after the process started. The situation is illustrated in Figure 1a, which schematically shows a number of realizations of stopped processes, and Figure 1b which shows these same realizations "right justified" at the stopping time. The rate of interest is seen to be proportional to the conditional probability of an event in an interval around the time y in the past, given that the process was in existence at time y in the past. That is, in looking back at time y we consider only evidence from realizations which were "alive" at that time; other realizations are disregarded.

This rate can be computed by first considering a rate conditioned on a particular value of the stopping time. In particular, we can first compute this conditional rate, defined to be

Figure 1a. Process realizations.

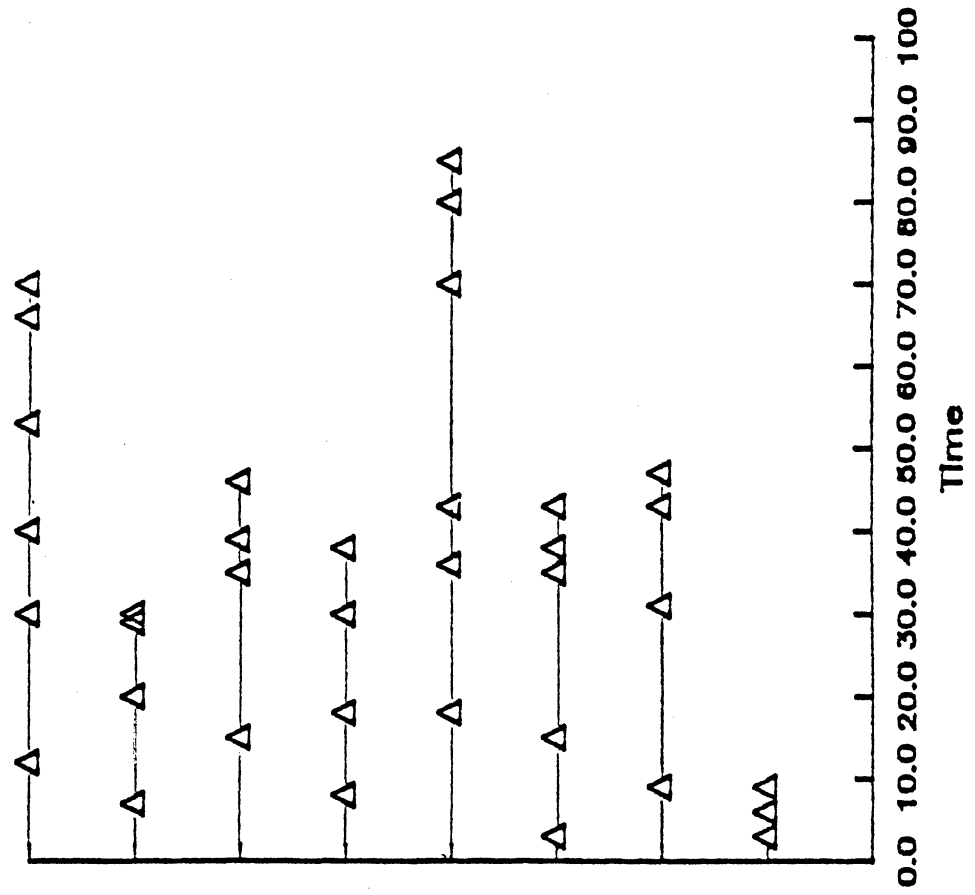
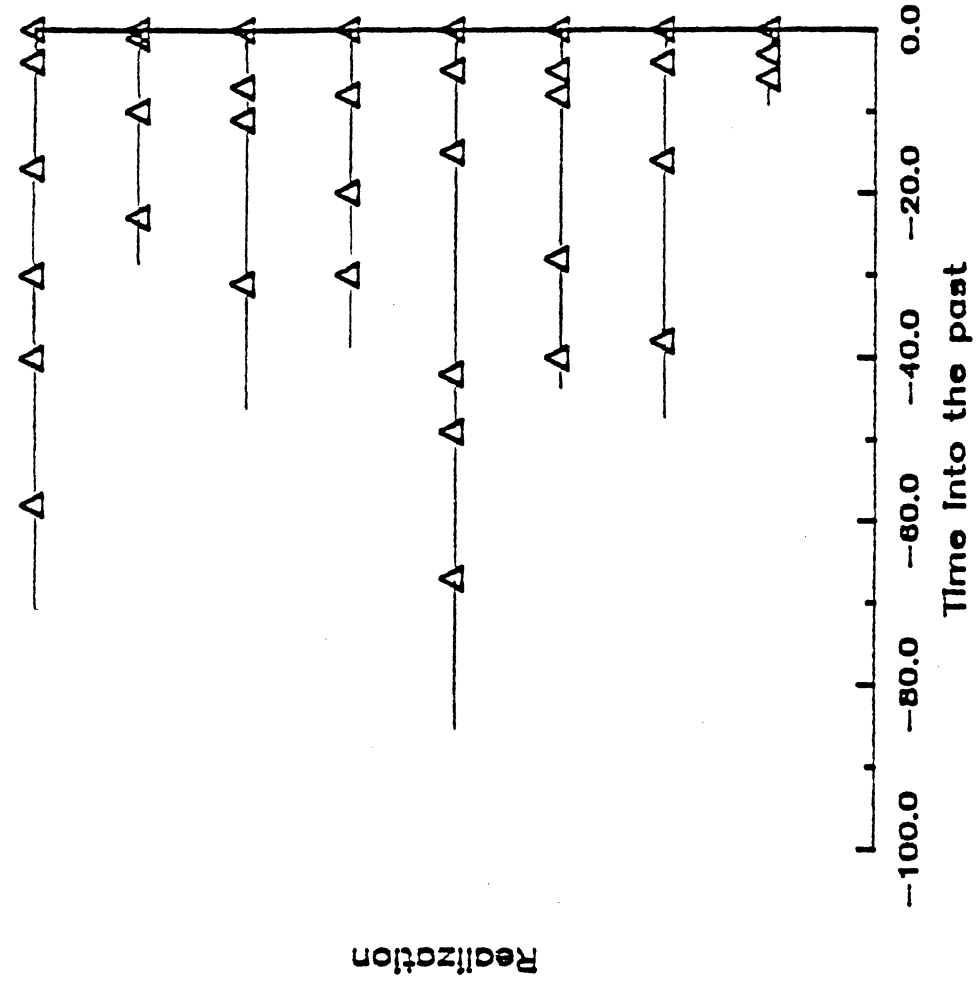


Figure 1b. Looking back from termination.



$$\rho(y|t) \equiv \text{prob. dens.}\{\text{non-stopping event at } t-y \mid \text{process stops at } t\} \quad (1)$$

(Throughout this paper, we assume that the relevant densities exist, that functions are smooth, that limits taken exist, etc., without giving mathematical details or describing discontinuous analogs of the cases described.)

Once this has been computed, this allows us to find the unconditional rate:

$$r(y) = \int_0^{\infty} \rho(y|t) h(y,t) dt \quad (2)$$

where

$h(y,t) \equiv$ p.d.f. for the stopping time T given $T > y$

$$= \begin{cases} 0 & 0 \leq t \leq y \\ \frac{f(t)}{\bar{F}(y)} & y < t \leq \infty \end{cases} \quad (3)$$

with $\bar{F}(y) \equiv \int_y^{\infty} f(t) dt$. The conditioning event $\{T > y\}$ is important here, since it guarantees that the time y into the past is smaller than the stopping time T .

One expects this rate $r(y)$ to be high for $y < w$, reflecting the stopping condition. One might expect it to be monotone non-increasing: as will be shown, it is not. One might also expect the "controlled" rate $r(y)$ for $y > w$ to be less than the "uncontrolled" rate λ and less than $(k-1)/w$ due to the stopping condition. The first of these assertions is true for the cases analyzed; the second is not and, in fact, λ can exceed k/w . Thus, the controlled process, viewed from its termination, has several non-intuitive properties.

SOLUTION FOR A POISSON PROCESS AND $k = 2$

In the case where $k = 2$ and the point process is a renewal process, Equation (1) may be usefully re-written by using Bayes' rule. In particular, by defining $g(y) \equiv \text{prob. dens.}\{\text{process stops at } t \mid \text{non-stopping event at } t-y\}$ and $e(t-y) \equiv \text{prob. dens.}\{\text{non-stopping event at } t-y\}$ then equation (1) becomes

$$p(y|t) = \frac{g(y)e(t-y)}{f(t)}$$

Combining this with equations (2) and (3) gives, after some manipulation,

$$r(y) = \frac{g(y) \int_0^{\infty} e(x) dx}{\bar{F}(y)} \quad (4)$$

(This formula does not hold for $k > 2$, in which case g is a function of y and $t-y$.)

Evaluation of equation (4) is made extremely difficult by the nature of the stopping process and its effects on $g(\cdot)$ and $f(\cdot)$. Even in the case of a stationary Poisson process with rate λ an analytic solution is readily obtainable only for $k = 1$ and $k = 2$. In the former case, of course $r(y) = 0$ for all $y > 0$ since the process stops at the first event, and thus the rate of events earlier than this is zero.

We can also analyze the case where $k = 2$. First, to find $g(y)$, we note that this is the p.d.f. for the time Y from a (non-stopping) event to termination of the process. Simple renewal arguments lead to the generator equations:

$$g(t) = \begin{cases} \lambda e^{-\lambda t} & 0 \leq t \leq w \\ \int_w^t \lambda e^{-\lambda x} g(t-x) dx & w < t \end{cases} \quad (5)$$

The upper expression on the right hand side represents a termination at t due to the next event's arrival time being less than w ; the lower expression convolutes the exponential next-event time (larger than w) with $g(\cdot)$. Successive solutions of equation (5) for intervals $(w, 2w]$, $(2w, 3w]$, ... leads eventually to

$$g(y) = \lambda e^{-\lambda y} \left[\sum_{n=0}^N \frac{\lambda^n (y-nw)^n}{n!} - \sum_{n=0}^{N-1} \frac{\lambda^n (y-(n+1)w)^n}{n!} \right] \quad (6)$$

where $N \equiv \left[\frac{y}{w} \right]$ = greatest integer part of $\frac{y}{w}$, and $\sum_{n=0}^{-1} \equiv 0$.

By similar arguments it can be shown that

$$\bar{F}(y) = e^{-\lambda y} \sum_{n=0}^{N+1} \frac{\lambda^n [y - (n-1)w]^n}{n!} \quad (7)$$

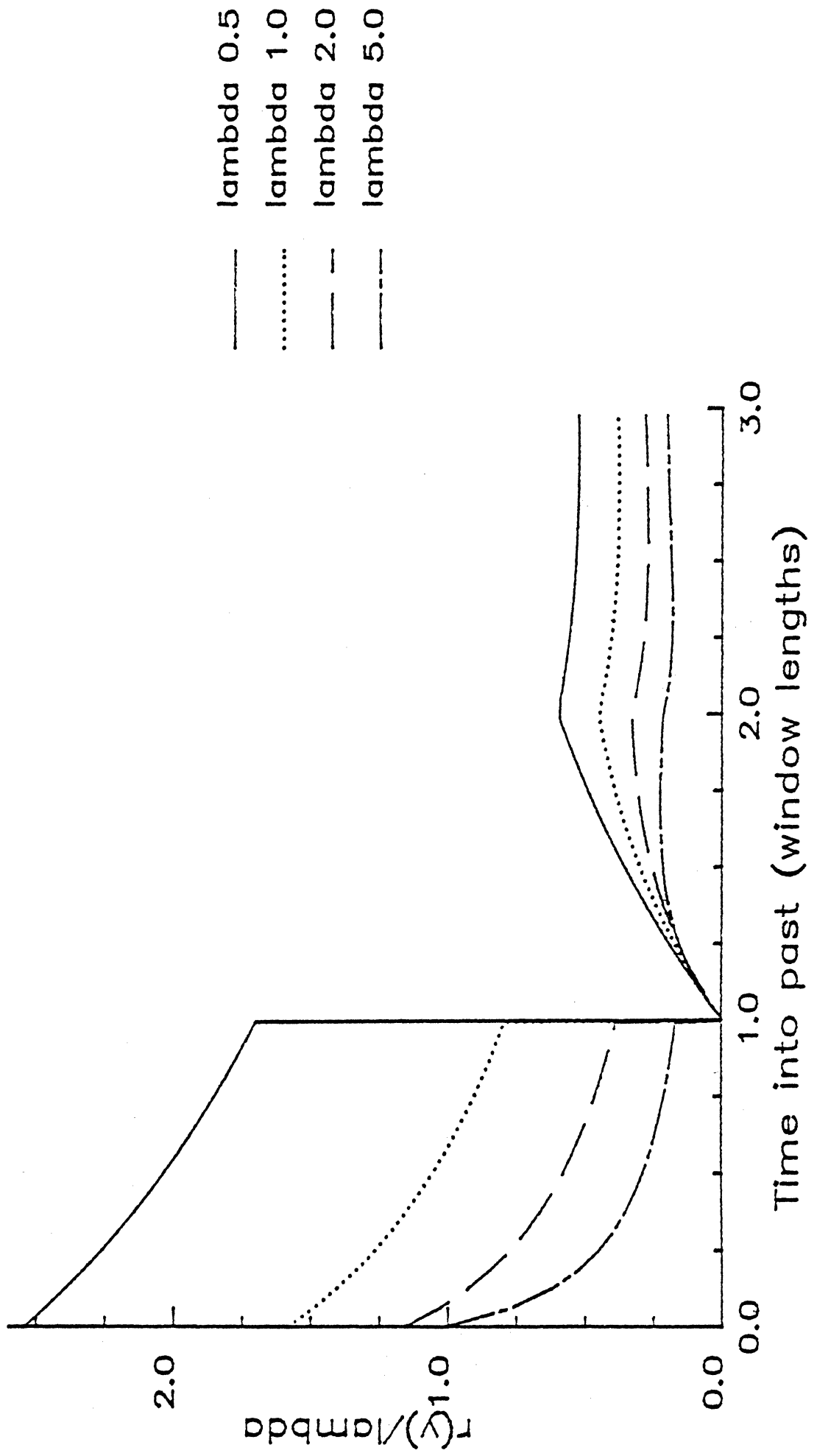
By the additivity of expectations, $\int_0^{\infty} e(x) dx$ is the expected number of non-stopping events. The number of such events is distributed geometrically with parameter $e^{-\lambda w}$. Therefore

$$\int_0^{\infty} e(x) dx = \frac{1}{1 - e^{-\lambda w}} \quad (8)$$

Figure 2 shows $r(y)/\lambda$, as calculated from equation (4), using the results of equation (6), (7) and (8). In the figure $w = 1$, so that time is in "window units", and $r(y)/\lambda$ is shown for $\lambda = .5, 1, 2$ and 5 . As can be seen, there are a number of interesting features to this retrospective look at the terminated process:

- (a) For all λ , the rate of occurrences into the past falls off until $t = 1$ where it drops to zero.

Figure 2. Ratio of $r(y)$ to λ



- (b) Values of $r > 1/w$ can occur even for $y > w$.
- (c) For times greater than 1 into the past, although the rate climbs in some time periods, it is never greater than λ , and even for $\lambda < 1$ can be considerably less than λ .
- (d) As $y \rightarrow \infty$, $r(y)/\lambda$ approaches a non-zero value dependent on λ and w .

To analyze the limiting behavior of $r(y)$ for large y , consider the defective renewal equations

$$g(y) = b(y) + \int_0^y g(y-u) dL(u)$$

$$F(y) = A(y) + \int_0^y F(y-u) dL(u)$$

where

$$A(y) = \begin{cases} 1 - e^{-\lambda y}(1 + \lambda y) & y < w \\ 1 - e^{-\lambda w}(1 + w) + w(e^{-\lambda w} - e^{-\lambda x}) & y > w \end{cases}$$

$$b(y) = \begin{cases} \lambda e^{-\lambda y} & 0 < y < w \\ 0 & y > w \end{cases}$$

$$L(y) = \begin{cases} 0 & y < w \\ e^{-\lambda w} - e^{-\lambda y} & y > w \end{cases}$$

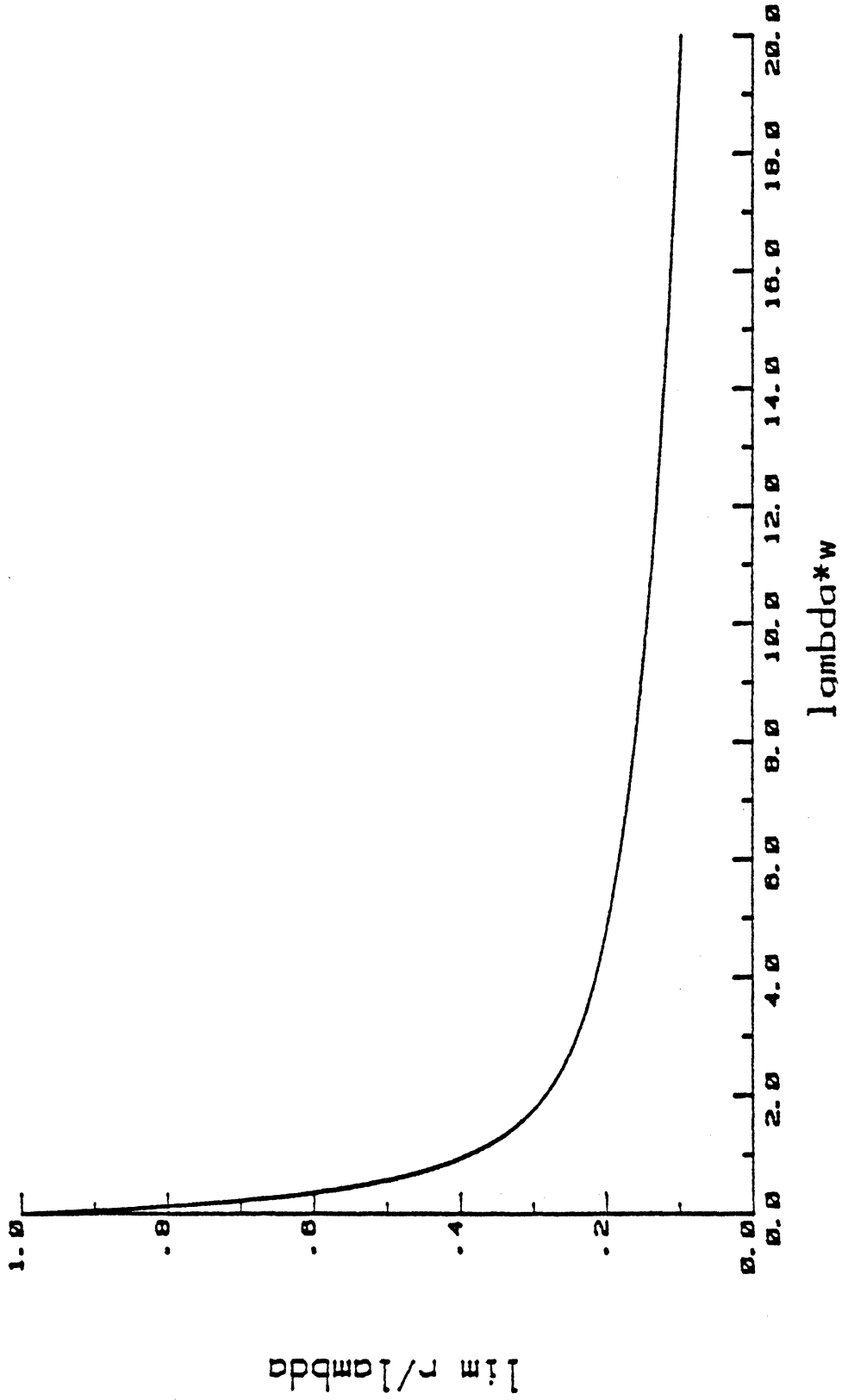
Using methods for the asymptotic analysis of renewal equations with defective distributions (note that $L(\infty) = e^{-\lambda w} = 1 - P_0 < 1$) given, for instance, in [2] and [3], one can eventually obtain

$$\lim_{y \rightarrow \infty} r(y) = \lim_{y \rightarrow \infty} \frac{g(y)/P_0}{1 - F(y)} = \frac{\lambda \eta (1 - \eta)}{1 - e^{-\lambda w}}$$

where η is the (unique) solution to $\eta = e^{-\lambda w \eta}$. Figure 3 shows

$\lim_{y \rightarrow \infty} r(y)/\lambda$ as a function of λw .

Figure 3. Graph of $\lim r/\lambda$.



DISCRETE TIME APPROACH FOR RENEWAL PROCESSES WITH $k > 2$

Analysis of even Poisson processes with $k > 2$ has thus far proven to be intractable. A discrete-time approximation for general renewal cases with arbitrary k however, is possible -- one that yields messy but computable results. A sketch of this method follows.

Re-define the stochastic process to be a discrete-time renewal process. Similarly re-define the stopping rule to be: the process terminates if there are k successes (events) within the past μ trials. Thus $\mu \approx w/\tau$. We wish to find the probability $r(j)$ that the j th trial before termination was a success, given that the process was "alive" then. Thus, $j \approx y/\tau$.

We can now identify an induced Markov Process (IMP) generated by the terminating process. The state space of IMP consists of a possibly countable number of states in three classes:

- (1) $1+\mu$ states which can be entered at most once each, representing initial histories of 0, 1, ..., μ failures (the first of these is the initial state);
- (2) a finite or countable number of states representing histories of $\mu+1$, ..., $\mu+i$, ... failures since the last success (or process initiation) -- the number of these states is either the maximum possible length of an interarrival interval, or the point at which the tail of the interarrival distribution becomes geometric; and
- (3) $2^\mu-1$ states representing all possible histories of length μ with one or more successes.

Some class (3) states are absorbing (and some -- those with more than k successes -- are impossible to reach from the initial state). These may

be lumped together into a single state, leaving $J = \sum_{i=1}^{K-1} \binom{\mu}{i}$ non-absorbing states in class (3) and a single absorbing state. The process after this lumping is called the Reduced IMP (RIMP); it has $K+J+1$ states where $K+J$, the number of class (1) and class (2) states, may be infinite.

The transition matrix P for the RIMP consists of all zeroes except for at most two elements in each row: one representing a success on the next trial, the other a failure. The transition probabilities are easily determined from the interarrival distribution of the renewal process.

Assume the absorbing state is numbered 0 and the initial state is numbered 1. Define the events:

$$S(n) \equiv \{\text{The } n\text{th trial is a success}\}$$

$$A_i(n) \equiv \{\text{The RIMP is in state } i \text{ after the } n\text{th trial}\}$$

$$T(n) \equiv \{\text{The RIMP terminates (first reaches its absorbing state) on the } n\text{th trial}\}.$$

Furthermore, let

$$\lambda^*(i) \equiv \text{the state of the RIMP that results from state } i \text{ when a success occurs.}$$

We wish to calculate the success probability on the j th trial preceding termination, given the process was alive at that time. This is, in analogy to equations (2) and (3)

$$\bar{r}(j) = \frac{\sum_{n=j+1}^{\infty} \bar{p}(j|n) \text{ pr.}\{T(n)\}}{\sum_{n=j+1}^{\infty} \text{ pr.}\{T(n)\}} \quad (9)$$

where

$$\begin{aligned} \rho(j|n) &= \text{pr.}\{\text{success on } j\text{th trial preceding termination |} \\ &\quad \text{termination on the } n\text{th trial}\} \\ &= \text{pr.}\{S(n-j) \mid T(n)\} \end{aligned}$$

which can be written

$$= \sum_{i=1}^{J+K} \text{pr.}\{S(n-j) \cap A_i(n-j-1) \mid T(n)\} \quad (10)$$

The terms in the summation in equation (10) can be re-written, by using Bayes' rule, as.

$$\begin{aligned} \text{pr.}\{S(n-j) \cap A_i(n-j-1) \mid T(n)\} &= \\ \left\{ \text{pr.}\{T(n) \mid S(n-j) \cap A_i(n-j-1)\} \times \right. & \quad (11a) \end{aligned}$$

$$\text{pr.}\{S(n-j) \mid A_i(n-j-1)\} \times \quad (11b)$$

$$\left. \text{pr.}\{A_i(n-j-1)\} \right\} \frac{1}{\text{pr.}\{T(n)\}} \quad (11c)$$

The conditioning event in term (11a) is the equivalent of $A_{\ell^*(i)}(n-j)$, since the RIMP is, after the $(n-j+1)$ st transition, in the state resulting from a success in state i .

The probability in term (11b) is, by definition, $P_{i\ell^*(i)}$. Combining equations (9), (10) and (11), then, gives

$$\bar{r}(j) = \frac{P_{i\ell^*(i)} \sum_{n=j+1}^{\infty} \sum_{i=1}^{J+K} \text{pr.}\{T(n) \mid A_{\ell^*(i)}(n-j)\} \text{pr.}\{A_i(n-j-1)\}}{\sum_{n=j+1}^{\infty} \text{pr.}\{T(n)\}} \quad (12)$$

The terms can be readily (if laboriously) computed from elements of powers of the matrix P:

$$\begin{aligned} \text{pr. } \{T(n) \mid A_{\ell^*(i)}(n-j)\} \\ = \sum_{h=1}^{J+k} [p^{j-1}]_{\ell^*(i),n} P_{h,0} \end{aligned}$$

$$\begin{aligned} \text{pr. } \{A_i(n-j-1)\} \\ = [p^{n-j-1}]_{1,i} \end{aligned}$$

$$\begin{aligned} \sum_{n=j+1}^{\infty} \text{pr. } \{T(n)\} \\ = 1 - [p^j]_{1,0} \end{aligned}$$

These formulae offer a computational approach to terminating processes with $k > 2$.

ACKNOWLEDGEMENT

We should point out that L. Tierney brought to our attention the flaws in the development of $r(t)$ in [1]. Jeff Alden was particularly helpful in pointing out an elegant transform approach that allows a verification of equations (6) and (7). M. Maltz has developed a simulation of the process that confirms our results. Part of this research was supported by ~~National Institute of Justice~~ ^{National Institute} of Justice Grant #81-IJ-CX 0064.

REFERENCES

- 1 Maltz, M.D. and Pollock, S. M. "Artificial Inflation of a Delinquency Rate by a Selection Artifact", Operations Research 28, pp. 547-559 (1980).
- 2 Feller, W., An Introduction to Probability Theory and Its Applications, Volume II, Wiley (1966).
- 3 Karlin, S. and Taylor, H. M., A First Course in Stochastic Processes, Second Edition, Academic Press (1975).