1.0

2.8 2.5

3.2 2.2

3.6

1.1 2.0

1.8

1.25 1.4 1.6

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963 A

AN ESTIMATION PROBLEM WITH POISSON PROCESSES

By

Alan E. Gelfand

TECHNICAL REPORT NO. 317

April 20, 1982

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

An Estimation Problem with Poisson Processes

By

Alan E. Gelfand

1. Introduction

This paper considers the following problem. Suppose we observe a collection of independent Poisson processes, $X_i(t)$, $i = 1, 2, \ldots, n$ such that

$$X_i(t) \sim P_0(\lambda_i(t)$$

with

$$\lambda_i(t) = \delta_i \rho(t) \quad.$$

Let

$$p(t) = \int_0^t \rho(u) du \quad.$$

If all the processes are observed up to a fixed time point T, how may we estimate $p(t)$? If we assume ρ integrable on $[0, \infty)$ and without loss of generality set $p(\infty) = 1$ then each $X_i(t)$ converges in distribution to $X_i(\infty)$ where

$$X_i(\infty) \sim P_0(\delta_i) \quad.$$

In section 2 we formalize the mathematical setting for this problem. In section 3 we describe two estimation approaches and illustrate these approaches with two choices of ρ. Application of this sort of estimation problem arises when there is interest in the conditional distribution of $X(t)$ given $X(\infty)$ that is, in the proportion of

1

occurrences of a process which may be expected by time t. In section 4 a successful application of the methodology to the active life of a judicial opinion is described.

2. The Set Up

Suppose

$$X(t) \sim P_0(\lambda(t))$$

with

$$\lambda(t) = \delta\rho(t) \quad .$$

Suppose further that

(i) ρ is integrable with

$$p(t) = \int_0^t \rho(u) \, du$$

and that $\lim_{t \to \infty} p(t) \equiv p(\infty) = 1$.

(ii) $\rho(t)$ strictly decreases in t , $\rho(0) < \infty$

(iii) $\rho'(t), \rho''(t)$ exist, $t > 0$.

Two particular examples of interest in the sequel are

$$\rho_1(t;\alpha) = \alpha e^{-\alpha t} \quad , \quad \alpha > 0 \quad \text{with}$$

$$p_1(t;\alpha) = 1 - e^{-\alpha t}$$

and

$$\rho_2(t;\alpha) = \frac{\alpha}{(\alpha+t)^2} \quad , \quad \alpha > 0 \quad \text{with}$$

$$p_2(t;\alpha) = t/(\alpha+t) \quad .$$

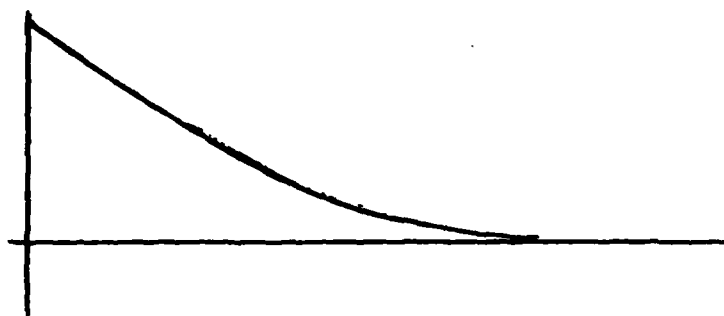In general ρ and p tend to look as in Figures 1 and 2.
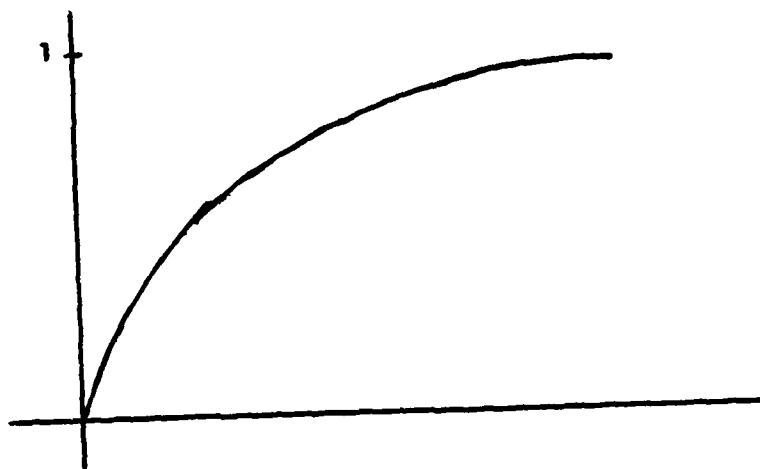

Fig 1:   A typical   ρ(t)

Fig 2:   A typical   p(t)

In view of our assumptions

$$X(t) \xrightarrow{d} X(\infty) \quad \text{where} \quad X(\infty) \sim P_0(\delta)$$

with interpretation

3

$X(t) = \#$ of occurrences of the process by time t .

$X(\infty) = \#$ of occurrences of the process at the end of time.

Why is $p(t)$ of interest? The conditional distribution of $X(t)$ given $X(\infty)$ is binomial i.e.

$$X(t)\,|\,X(\infty) \sim B_1(X(\infty),\ p(t))$$

and therefore simple computation shows that

$$E\left\{\frac{X(t)}{X(\infty)} \mid X(\infty) > 0\right\} = p(t)$$

$$\text{var}\left\{\frac{X(t)}{X(\infty)} \mid X(\infty) > 0\right\} = c(\delta)\ p(t)(1-p(t)) \ .$$

where

$$c(\delta) = (1-e^{-\delta})^{-1} \sum_{y=1}^{\infty} \frac{\delta^y e^{-y}}{y\ y!} \ .$$

Hence $p(t)$ measures the expected proportion of occurrences of the process by time t . An implication of our assumptions on ρ

is that for any interval of fixed length, the earlier in time it is placed, the greater the expected number of occurrences in the interval.

The significance of $p(t)$ is more strongly emphasized in the case where we have n independent processes $X_i(t)$ such that

$$X_i(t) \sim P_0(\lambda_i(t)) \qquad i = 1,\ldots,n$$

where

(1) $$\lambda_i(t) = \delta_i \rho(t) \ .$$

In this case we would like to synthesize information from the n processes to estimate the common $p(t)$. In other words, for any

4

individual process there would be little value in factoring a δ from the intensity function. However, for some phenomena that we might sample (see the application in section 4) the behaviour might be such that from sample to sample

i)   the absolute intensity (i.e. expected rate of occurrence) varies
      but
ii)   the relative intensity (i.e. ratio of expected rate of
      occurrence at two different time points) does not.

In such a case the form for $\lambda_i(t)$ in (1) is appropriate. In our application we envision considerable variation in the δ_i . For some sampled processes, occurrences will be very sparse while for others they will be quite numerous. But again our interest is not in the δ_i but in $p(t)$.

Suppose then that we observe the n processes during the interval [0,T]. How may we estimate $p(t)$? (The reader may be curious as to whether differing lengths of observation time, for differing sampled processes can be accommodated. We will return to this question shortly.)

3. Estimation Approaches

In attempting to estimate $p(t;\alpha)$ using data in $[0,T]$ we consider two approaches.

A. Use the actual arrival times of the occurrences

B. Partition $[0,T]$ into intervals and record numbers of arrivals in each interval.

Clearly the latter approach results in some loss of information compared with the former. We shall obtain a measure of this loss and apply it to our ρ's of interest. In applications, it may be the case that data is gathered categorically for reasons of convenience or reliability so that only approach B is available. Hence an indication of the loss of information while still of interest may no longer be relevant.

Consider Approach A. From Cox & Hinkley (p.15-16) if $X(t) \sim P_0(\delta\rho(t))$ and in the interval $[0,T]$ (ordered) occurrences are observed at $w_1, w_2, \ldots, w_m$ then the unconditional likelihood becomes

$$L(\rho, \delta, m, w_1, w_2, \ldots, w_m)$$

$$(2) \qquad = \delta^m \prod_{i=1}^{m} \rho(w_i) \, e^{-\delta \int_0^T \rho(u)du} .$$

The conditional likelihood given $X(t) = m$ thus becomes

6

$$L(\rho, w_1, w_2, w_3, \ldots, w_m | X(T) = m)$$

$$= \frac{m! \, \delta^m \prod_{i=1}^{m} \rho(w_i) \, e^{-\delta \int_0^T \rho(u)\,du}}{\delta^m \left[\int_0^T \rho(u)\,du\right]^m e^{-\delta \int_0^T \rho(u)\,du/m}}$$

$$(3) \qquad = m! \prod_{i=1}^{m} \rho(w_i) \, (p(T))^{-m} \quad ,$$

free of δ .

For n samples with occurrences time w_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, m_i$ (that is $X_i(T) = m_i$) , $\Sigma m_i = m$, we similarly obtain

$$L(\rho, \delta_i, \, m_i, \, w_{ij})$$

$$(4) \qquad = \left\{ \prod_{i=1}^{n} \delta_i^{m_i} \right\} \prod_{i=1}^{n} \prod_{j=1}^{m_i} \rho(w_{ij}) \, e^{-(\Sigma \delta_i) p(T)}$$

and $L(\rho, w_{ij} | X_i(T) = m_i , \ i = 1, \ldots, n)$

$$(5) \qquad = \prod m_i! \prod_{i=1}^{n} \prod_{j=1}^{m_i} \rho(w_{ij}) (p(T))^{-m} .$$

(Note that if sample i is observed over $[0, T_i]$ the only change to be made in (5) is that $(p(T))^{-m}$ is replaced by $\prod_{i=1}^{n} (P(T_i))^{-m_i}$) .

Suppose ρ is of the form $\rho(t;\alpha)$ (p of the form $p(t;\alpha)$) so that with estimation of α, ρ is specified and thus p is

7

estimated. Then, in (2) the unconditional likelihood equations for α and δ are

$$\frac{\partial \log L}{\partial \alpha} = \sum \frac{\partial \log \rho(w_i;\alpha)}{\partial \alpha} - \delta \frac{\partial p(T;\alpha)}{\partial \alpha} = 0$$

$$\frac{\partial \log L}{\partial \delta} = \frac{m}{\delta} - p(T;\alpha) = 0$$

whence

$$(6) \qquad \frac{\partial \log L}{\partial \alpha} = \sum \frac{\partial \log \rho(w_i;\alpha)}{\partial \alpha} - m \frac{\partial p(T;\alpha)}{\partial \alpha} \Big/ p(T;\alpha) = 0 \; .$$

In (3) this conditional likelihood equation for α is

$$(7) \qquad \frac{\partial \log L/m}{\partial \alpha} = \sum \frac{\partial \log \rho(w_i;\alpha)}{\partial \alpha} - m \frac{\partial \log p(T;\alpha)}{\partial \alpha} = 0 \; .$$

Hence as is well known (e.g. Cox and Lewis, Chap 3) the conditional likelihood equation for α is the same as the unconditional one. We may rewrite (7) as

$$(8) \qquad \sum \frac{\partial \log(\rho(w_i;\alpha)/p(T;\alpha))}{\partial \alpha} = 0 \; .$$

In (8) the conditional nature of the likelihood equation is most apparent i.e. $\rho(u;\alpha)/p(T;\alpha)$ is the density function for an occurrence given it occurred by time T. Thus if $\hat{\alpha}$ is a solution (8) and the usual regularity conditions are satisfied, $\hat{\alpha}$ will be asymptotically normal with mean the true α and asymptotic variance $(I(\alpha))^{-1}$ where

$$I(\alpha) = E\left[\frac{\partial \log(\rho(W_1;\alpha)/p(T;\alpha))}{\partial \alpha}\right]^2$$

$$(9) \qquad = E\left(\frac{\partial \log \rho(W;\alpha)}{\partial \alpha}\right)^2 - \left(\frac{\partial \log p(T;\alpha)}{\partial \alpha}\right)^2 .$$

Computationally a more convenient form may be

$$I(\alpha) \equiv - E\left[\frac{\partial^2 \log(\rho(W;\alpha)/p(T;\alpha))}{\partial \alpha^2}\right]$$

$$(10) \qquad = - E \frac{\partial^2 \log \rho(W;\alpha)}{\partial \alpha^2} + \frac{\partial^2 \log p(T;\alpha)}{\partial \alpha} .$$

For the case of n samples, α is a solution to

$$(11) \qquad \sum_{i=1}^{n} \sum_{j=1}^{m_i} \frac{\partial \log \rho(W_{ij};\alpha)}{\partial \alpha} = m \frac{\partial \log p(T;\alpha)}{\partial \alpha}$$

from which we see that the separation of the occurrence times by samples is superfluous. The total of m occurrence times may be considered as a single sample. This may be noted directly by considering the occurrence times of the process $Y(t) = \sum_{i=1}^{n} X_i(t)$ in $[0,T]$.

Returning to (7), consider $\rho_1(t;\alpha) = \alpha e^{-\alpha t}$. We obtain

$$\sum \left(\frac{1}{\alpha} - w_i\right) - \frac{m T e^{-\alpha T}}{1 - e^{-\alpha T}} = 0$$

or

$$\sum w_i = \frac{m}{\alpha} - \frac{m T e^{-\alpha T}}{1 - e^{-\alpha T}}$$

and

(12) $$I(\alpha) = \frac{1}{\alpha^2} - \frac{T^2 e^{-\alpha T}}{(1-e^{-\alpha T})^2} \, .$$

(Again see Cox and Lewis, p.47).

For $\rho_2(t;\alpha) = \alpha/(\alpha+t)^2$ we obtain

$$\sum \left\{ \frac{1}{\alpha} - \frac{2}{\alpha+w_i} \right\} + \frac{m}{\alpha+T} = 0$$

or

$$\sum \left\{ \frac{1}{\alpha+w_i} \right\} \quad \frac{m(2\alpha+T)}{2\alpha(\alpha+T)}$$

and

(13) $$I(\alpha) = T^2/3\alpha^2(\alpha+T)^2 \, .$$

We now consider approach B . Suppose we define k intervals by $0 < t_1 < t_2 < \ldots t_k = T$. Then for one sample

$$X(t_1), \ X(t_2) - X(t_1) \ ,\ldots, \ X(t_{k-1}) - X(t_{k-2}) \,|\, X(T) = m$$

$$\sim \text{Multinomial} \quad (m, \ q_1, \ q_2, \ldots, q_k)$$

where

$$q_j = \frac{p(t_j) - p(t_{j-1})}{p(T)} \ , \quad j = 1,\ldots,k$$

$$(t_0 \equiv 0)$$

For n samples let

$$z_j = \sum_{i=1}^{n} (X_i(t_j) - X_i(t_{j-1}))$$

10

Then $Z = (Z_1, Z_2, \ldots, Z_{k-1})$ is complete and sufficient given $X_i(T) = m_i$ and in fact is multinomially distributed as well i.e. $(\Sigma m_i = m)$

$$Z \sim \text{Multinomial} \quad (m, q_1, \ldots, q_k) \quad .$$

Two natural estimation approaches arise at this point: the method of minimum chi square and the method of maximum likelihood (method of moments estimation might be considered as well). The former leads to consideration of the function

$$\sum_{j=1}^{k} (Z_j - mq_j)^2 / mq_j$$

while the latter leads to the (conditional) likelihood function

$$L(\rho, Z_1, \ldots, Z_k / m)$$

$$= m! \prod_{j=1}^{k} q_j^{Z_j} / Z_j! \qquad \left(Z_k \equiv m - \sum_{1}^{k-1} Z_j\right) \quad .$$

Again presuming ρ is of the form $\rho(t; \alpha)$, the resulting likelihood equation for α is

$$\frac{\partial \log L/m}{\partial \alpha} = \sum \frac{\partial \log q_j(\alpha)}{\partial \alpha} = 0$$

or

$$(14) \qquad \sum Z_j \frac{\partial \log (p(t_j; \alpha) - p(t_{j-1}; \alpha))}{\partial \alpha} - \frac{m \partial \log p(T; \alpha)}{\partial \alpha} = 0 \quad .$$

Again if $\hat{\alpha}_G$ (G for grouped) is a solution to (14), under the usual regularity conditions $\hat{\alpha}_G$ will be asymptotically normal with mean

11

the true α and asymptotic variance $(I_G(\alpha))^{-1}$ where

$$I_G(\alpha) = -\frac{1}{m} E\left[\sum z_j \frac{\partial^2 \log q_j(\alpha)}{\partial \alpha^2}\right]$$

(15)

$$= -\sum q_j(\alpha) \frac{\partial^2 \log q_j(\alpha)}{\partial \alpha^2}$$

or in terms of the $p_j(\alpha)$

(16) $(p(T;\alpha))^{-1} \sum \left(p(t_j;\alpha) - p(t_{j-1};\alpha)\right)^{-1}\left[\frac{\partial\left(p(t_j;\alpha)-p(t_{j-1};\alpha)\right)}{\partial \alpha}\right]^2 - \left(\frac{\partial \log p(T;\alpha)}{\partial \alpha}\right)^2$

The similarity between expressions (7) and (14) and between (9) and (16) is not unexpected. In fact, if as $k \to \infty$, $\sup\limits_{1 \le j \le k}(t_j - t_{j-1}) \to 0$

it is straightforward to show that the "categorical" expressions converge to the corresponding "continuous" ones .

In (14), consider $\rho_1(t;\alpha) = \alpha e^{-\alpha t}$. Since

$$q_j(\alpha) = \frac{e^{-\alpha t_{j-1}} - e^{-\alpha t_j}}{1-e^{-\alpha T}}$$

we obtain

(17) $\quad -\sum z_j \dfrac{t_{j-1}e^{-\alpha t_{j-1}} - t_j\, e^{-\alpha t_j}}{e^{-\alpha t_{j-1}} - e^{-\alpha t_j}} - \dfrac{mTe^{-\alpha T}}{1-e^{-\alpha T}} = 0$

and

$$(18) \qquad I_G(\alpha) = \sum \frac{e^{-\alpha(t_{j-1}+t_j)}(t_j-t_{j-1})^2}{(1-e^{-\alpha T})(e^{-\alpha t_{j-1}}-e^{-\alpha t_j})} - \frac{T^2 e^{-\alpha T}}{(1-e^{-\alpha T})^2} \quad .$$

Expressions (17) and (18) can be simplified considerably if we use equally spaced intervals. Let $\bar{t} = T/k$ whence $t_j = j\bar{t}$. Then (17) becomes

$$(19) \qquad \sum j Z_j = m(1-e^{-\alpha\bar{t}})^{-1} - m k e^{-\alpha T}(1-e^{-\alpha T})^{-1}$$

and (18) becomes

$$(20) \qquad \bar{t}^2 e^{-\alpha\bar{t}}(1-e^{-\alpha\bar{t}})^{-2} - T^2 e^{-\alpha T}(1-e^{-\alpha T})^{-2} \quad .$$

In (14) consider $\rho_2(t;\alpha) = \alpha/(\alpha+t)^2$. Now

$$q_j(\alpha) = \frac{\alpha(\alpha+T)}{T} \frac{(t_j-t_{j-1})}{(\alpha+t_j)(\alpha+t_{j-1})}$$

and we obtain

$$(21) \qquad \sum \frac{Z_j(t_j t_{j-1}-\alpha^2)}{\alpha(\alpha+t_j)(\alpha+t_{j-1})} + \frac{m}{\alpha+T} = 0$$

and

$$(22) \qquad I_G(\alpha) = \frac{\alpha+T}{T} \sum \left(\frac{t_j}{\alpha+t_j} - \frac{t_{j-1}}{\alpha+t_{j-1}}\right)^{-1} \left[\frac{t_j}{(\alpha+t_j)^2}-\frac{t_{j-1}}{(\alpha+t_{j-1})^2}\right]^2 - \frac{1}{(\alpha+T)^2} \quad .$$

Again, equally spaced intervals simplify matter Expression (21) becomes

$$(23) \qquad \sum Z_j \frac{(2\alpha+2j-1)\bar{t}}{(\alpha+j\bar{t})(\alpha+(j-1)\bar{t})} = \frac{m(2\alpha+T)}{\alpha(\alpha+T)}$$

and (22) becomes

$$(24) \qquad \frac{(\alpha+T)}{T} \frac{\bar{t}}{\alpha} \sum \frac{(\alpha^2-j(j-1)\bar{t}^2)^2}{(\alpha+jt)^3(\alpha+(j-1)\bar{t})^3} - \frac{1}{(\alpha+T)^2} \quad .$$

13

We now attempt comparison of the estimation approaches. Again $\hat{\alpha}$ is the MLE from (7) while $\hat{\alpha}_G$ is the MLE from (14). We employ ideas similar to those of Lindley . Given $\hat{\alpha}_G$ how may we "correct" to $\hat{\alpha}$. We obtain an approximation for $\Delta \equiv \hat{\alpha}_G - \hat{\alpha}$. Lindley considers the reverse situation; from the functional form for the ungrouped MLE he obtains a correction to the grouped MLE. In our applications the grouped MLE is all that is available. Lindley also considers an i.i.d. sample while our W_j are clearly not so.

We also seek to estimate the loss of efficiency due to grouping. Since $\hat{\alpha}$ is most efficient we calculate the (asymptotic) relative efficiency of $\hat{\alpha}_G$ to $\hat{\alpha}$ by $I(\alpha)/I_G(\alpha)$. In fact the proportion of information lost is $(I(\alpha) - I_G(\alpha))/I(\alpha)$.

To approximate Δ we assume equally spaced intervals. Using Taylor's theorem

$$p(j\bar{t};\alpha) - p((j-1)\bar{t};\alpha) \approx \bar{t}\,\rho((j-\tfrac{1}{2})\bar{t};\alpha) + \frac{\bar{t}^3}{24}\,\rho''((j-\tfrac{1}{2})\bar{t};\alpha)$$

whence

$$\frac{\partial\,\log(p(j\bar{t};\alpha) - p((j-1)\bar{t};\alpha))}{\partial\alpha} \approx \frac{\partial\,\log(\rho(j-\tfrac{1}{2})\bar{t};\alpha)}{\partial\alpha} + \frac{\bar{t}^2}{24}\,\frac{\partial}{\partial\alpha}\left(\frac{\rho''((j-\tfrac{1}{2})\bar{t};\alpha)}{\rho((j-\tfrac{1}{2})\bar{t};\alpha)}\right).$$

Hence using (14) and suppressing the arguments of ρ and p we have

$$\sum z_j\,\left.\frac{\partial\,\log\rho}{\partial\alpha}\right|_{\hat{\alpha}_G} + \frac{\bar{t}^2}{24}\sum z_j\,\left.\frac{\partial}{\partial\alpha}\left(\frac{\rho''}{\rho}\right)\right|_{\hat{\alpha}_G} - \left.\frac{m\,\partial\,\log p}{\partial\alpha}\right|_{\hat{\alpha}_G} \approx 0 \ .$$

14

In (7), presume that if $w_1 \in ((j-1)\bar{t}, j\bar{t}]$ then

$$\rho(w_1;\alpha) \approx \rho((j-\tfrac{1}{2})\bar{t};\alpha)$$

so that

$$\sum z_j \left.\frac{\partial \log \rho}{\partial \alpha}\right|_{\hat{\alpha}} - \left.\frac{m\,\partial \log p}{\partial \alpha}\right|_{\hat{\alpha}} \approx 0 \quad .$$

To a first approximation, then

$$(25) \qquad \Delta = \frac{-\dfrac{\bar{t}^2}{24} \sum z_j \dfrac{\partial}{\partial \alpha}\left.\left(\dfrac{\rho''}{\rho}\right)\right|_{\hat{\alpha}_G}}{\sum z_j \left.\dfrac{\partial^2 \log \rho}{\partial \alpha^2}\right|_{\hat{\alpha}_G} - \left.\dfrac{m\,\partial^2 \log p}{\partial^2 \alpha}\right|_{\hat{\alpha}_G}} \quad .$$

Iteration can be employed to better the approximation. For $\rho_1(t;\alpha) = \alpha e^{-\alpha t}$ (25) becomes

$$\frac{\bar{t}^2 \hat{\alpha}_G}{12} \left[\frac{1}{\hat{\alpha}_G^{\,2}} - \frac{T^2 e^{-\hat{\alpha}_G T}}{\left(1-e^{-\hat{\alpha}_G T}\right)^2} \right]^{-1} \quad .$$

For $\rho_2(t;\alpha) = \alpha/(\alpha+t)^2$, (25) becomes

$$\frac{\bar{t}^2}{2} \sum z_j b_j^{-3} \left[\frac{m\left((\hat{\alpha}_G+T)^2 + \hat{\alpha}_G^2\right)}{\hat{\alpha}_G^2\,(\hat{\alpha}_G+T)^2} - 2 \sum z_j b_j^{-2} \right]^{-1}$$

where $b_j = \hat{\alpha}_G + (j-\tfrac{1}{2})\bar{t}$.

15

For the relative efficiency of $\hat{\alpha}_G$ to $\hat{\alpha}$, from (9) and (16) we obtain

$$\text{Eff}(\hat{\alpha}_G|\hat{\alpha}) = \frac{E\left[\frac{\partial \log \rho(w;\alpha)}{\partial \alpha}\right]^2 - \left[\frac{\partial \log p(T;\alpha)}{\partial \alpha}\right]^2}{(p(T;\alpha))^{-1}\Sigma(p(t_j;\alpha)-p(t_{j-1};\alpha))^{-1}\left[\frac{\partial}{\partial \alpha}(p(t_j;\alpha)-p(t_{j-1};\alpha))\right]^2 - \left[\frac{\partial \log p(T;\alpha)}{\partial \alpha}\right]^2}$$

Let us consider our illustrative ρ_1, ρ_2 using equally spaced intervals. For ρ_1, we have

$$(26) \quad \text{Eff}(\hat{\alpha}_G|\hat{\alpha}) = \frac{\frac{1}{\alpha^2} - \frac{T^2 e^{-\alpha T}}{(1-e^{-\alpha T})^2}}{\frac{\bar{t}^2 e^{-\alpha \bar{t}}}{(1-e^{-\alpha t})^2} - \frac{T^2 e^{-\alpha T}}{(1-e^{-\alpha T})^2}}$$

For fixed T, the following limits may be easily discerned in (26).

i) $\bar{t}$ i.e. k fixed, $\alpha \to 0$, $\text{Eff} \to k^2/(k^2-1)$

ii) $\bar{t}$ i.e. k fixed, $\alpha \to \infty$, $\text{Eff} \to \infty$

iii) α fixed, $\bar{t} \to 0$ i.e. $k \to \infty$, $\text{Eff} \to 1$

iv) α fixed, $\bar{t} \to T$ i.e. $k \to 1$, $\text{Eff} \to \infty$

Letting $s = \alpha T$, with $k\bar{t} = T$ (26) becomes

$$(27) \quad \frac{k^2 g(s/k)(g(s) - s^2)}{s^2(g(s) - k^2 g(s/k))}$$

where

$$g(s) = (e^s-1)(1-e^{-s}) = \sum_{j=1}^{\infty} \frac{2s^{2j}}{(2j)!} \ .$$

16

Table 1 evaluates (27) over a range of s and k values. For ρ_2

$$(28) \qquad \text{Eff}(\hat{\alpha}_G | \alpha) = \frac{\dfrac{T^2}{3\alpha^2(\alpha+T)^2}}{\dfrac{\bar{t}(\alpha+t)}{\alpha T} \sum \dfrac{(\alpha^2-j(j-1)\bar{t}^2)^2}{(\alpha+j\bar{t})^3(\alpha+(j-1)\bar{t}^{-1})^3} - \dfrac{1}{(\alpha+T)^2}} \quad .$$

Letting $s = \alpha/T$ with $k\bar{t} = T$ (28) becomes

$$(29) \qquad \left\{ 3sk(s+1)^3 \sum \frac{(s^2k^2-j(j-1))^2}{(sk+j)^3(sk+(j-1))^3} - 3s^2 \right\}^{-1} \quad .$$

Table 2 evaluates (29) over a range of s and k values. The limiting behavior noted above for (26) also occurs for (28) with "$\alpha = 1/\alpha$".

Regardless of what parametric form was adopted for ρ and of what estimation approach was used the following offers a simple way of assessing fit. For any two time points t_1, t_2, $t_1 < t_2$, $X(t_1) \mid X(t_2) \sim B_1(X(t_2), p(t_1)/p(t_2))$. Hence (if $X(t_2) > 0$) $X(t_1)/X(t_2)$ is the UMVU of $p(t_1)/p(t_2)$. For any estimate $\hat{p}(t) \equiv p(t,\hat{\alpha})$ we can compare $p(t_1;\hat{\alpha})/(p(t_2;\hat{\alpha})$ with the UMVU ratio.

17

4. An Application

One attractive application of the above methodology is to the
matter of describing the active lifetime of a court case. More
specifically, associated with court opinions (particularly higher court
opinions such as those of the U.S. appellate or supreme courts) is the
issue of precedential value. How much effect will an opinion have on
later opinions? One way to study this point is to record citations
of this opinion in later cases. It is thus of interest to examine
the pattern of subsequent citation of a case. It seems reasonable
to assume that subsequent citations arrive according to a Poisson
process. Moreover, although the intensity function will vary greatly
from case to case the expected proportion of citations to a given time
need not. (For the data to be presented, some preliminary empirical
work using fixed points t_1, t_2 and comparing binomials has
supported this premise.) That is, sparsely cited cases will receive
about the same relative proportion of total cites by a fixed time as will
heavily cited cases. Our assumptions regarding $p(t)$ made in section
2 seem appropriate in this context since an opinion is a bit like wine.
At first it improves with age (i.e. is more often cited) but then it
deteriorates (i.e. is less often cited as newer opinion diminish its value).
Estimation approach B was forced in this situation. Data collection
recorded the date an opinion was handed down along with the dates of
subsequent citation. It was decided more convenient and reliable to
categorize subsequent citations into intervals after decision date than
to convert to the exact number of days after the decision date.

18

The court cases studied in this application are all from the U.S. Court of Appeals. They are drawn from a larger study (see Schuchman and Gelfand) of the Federal appellate courts. Two independent samples were analyzed. One consists of 100 cases from the Fifth circuit court of appeals while the other consists of 100 cases from the remaining ten circuits. In each sampled case the date of decision was recorded (all dates are between 1960 and 1962). By use of Shepard's Index we may find all later cases citing any sampled case. By examining these citing opinions, we may obtain the total number of citations for any sampled case and a categorical time from date of decision to each such citation. Subsequent citation of each case in each sample was considered through December 1976 thus achieving a minimum of 13 full years of observation with numerous cases reaching to nearly 16 full years.

For convenience we set $T = 12$ and consider $\bar{t} = 2, k = 6$.

The resulting Z vectors are given in Table 3. Note that the fifth circuit opinions are more frequently cited than their other circuit counterparts. Furthermore, we find an inexplicable, but common, bump in both samples occurring at Z_4, i.e. for the period between 6 and 8 years after date of decision. A test of homogeneity between the samples yields $\chi^2 = 1.290$ with descriptive level $= .93$.

Table 4 presents the estimation of α using both $p_1(t;\alpha)$ and $p_2(t;\alpha)$. We see that

(i)   Both functions p_1 and p_2 fit the data well

(ii)   Estimation by MLE and $MM\chi^2$ yield very similar estimates of α

(iii) The estimates of α for the two samples are fairly close

       (as expected from the test of homogeneity)

(iv)   For the 5th circuit data using p_1 and $\hat{\alpha}_G \approx .09, \Delta = 0.0027$

       while for p_2 and $\hat{\alpha}_G \approx 18$, $\Delta = 1.038$.

(v)   From Tables 1 and 2 the relative efficiency for p_1 under the given

       grouping is 1.03, for p_2, 1.04.

In Table 5 we compare the fits of p_1 and p_2 with the "best"
(i.e. UMVU) estimate as described at the end of the previous section.
We look at t_i relative to t_{i+1}, $t_i = \overline{it}$. Again the fits
look quite good i.e. the maximum relative error for the exponential
model is .03, for the bilinear model .04.

How can it be that two such different functions p_1 and p_2

fit the data almost identically? Table 6 clarifies the matter.
It also provides representative values of the function over time
using $\alpha = .09$ in p_1 and $\alpha = 18$ in p_2 . We see that
during the time interval $t = 2$ through 12 years $p_1(t)/p_2(t)$
is virtually constant. Since the assessment of fit involves $p(t)$
in a relative way, over the 12 years of observation p_1 and p_2

will be virtually indistinguishable. However in an absolute sense

p_1 is preferred to p_2.   This follows by asking how many years
we expect to wait for a specified proportion of subsequent citations
e.g. what is t such that $p_1(t) = .9$, such that $p_2(t) = .9$?

For p_1 the answer is 25.6 years, for p_2, 162 years. Additionally $p_1 (33.3) = .95$ and $p_1(51.1) = .99$. These proportions are reasonable both empirically and intuitively. Thus an exponential intensity function seems an effective descriptor for subsequent citation of an opinion.

s \ k	2	3	4	6	10	20	50
.1	1.333	1.125	1.067	1.029	1.010	1.003	1
.5	1.337	1.127	1.068	1.029	1.010	1.003	1
1	1.350	1.131	1.070	1.030	1.011	1.003	1
5	1.893	1.323	1.170	1.072	1.025	1.006	1.001
10	5.990	2.361	1.647	1.255	1.087	1.021	1.003
50	$> 10^8$	$> 10^4$	$> 10^3$	59.879	5.857	1.642	1.086

Table 1: Relative Efficiency for $p_1(t;\alpha) = 1-e^{-\alpha t}$

s \ k	2	4	6	10	20	50
.01	34.337	11.896	7.412	4.428	2.480	1.445
.05	7.683	3.034	2.119	1.529	1.178	1.035
.1	4.364	1.951	1.491	1.210	1.061	1.010
.5	1.778	1.170	1.075	1.027	1.007	1.001
1	1.500	1.102	1.044	1.016	1.004	1.001
5	1.344	1.067	1.303	1.010	1.003	1
10	1.336	1.067	1.029	1.010	1.003	1

Table 2: Relative Efficiency for $p_2(t;\alpha) = t/(\alpha+t)$

Sample	z_1	z_2	z_3	z_4	z_5	z_6
Fifth Circuit	225	191	145	152	112	85
Other Circuits	182	144	114	123	84	78

Table 3 :   The Fifth Circuit and Other Circuit Data

	$\hat{\alpha}_G$(MLE)	$\hat{\alpha}_G$(MMχ^2)	MMχ^2(D.L.)
$P_1(t;\alpha)$ $= 1 - e^{-\alpha t}$			
Fifth Circuit	.0893	.0888	4.679(.32)
Other Circuits	.0831	.0827	4.480(.35)
$P_2(t;\alpha)$ $= \dfrac{t}{\alpha+t}$			
Fifth Circuit	17.455	17.646	5.681(.22)
Other Circuits	18.753	18.905	4.386(.36)

Table 4 : Estimation of α .

24

	t_1/t_2	t_2/t_3	t_3/t_4	t_4/t_5	t_5/t_6
Fifth Circuit					
"Best" (UMVU)	1.849	1.349	1.271	1.157	1.103
P_1 $\hat{\alpha}(MM\chi^2)$	1.837	1.382	1.231	1.157	1.114
P_1 $\hat{\alpha}(MLE)$	1.836	1.381	1.231	1.1157	1.113
P_2 $\hat{\alpha}(MM\chi^2)$	1.815	1.373	1.229	1.160	1.119 .
P_2 $\hat{\alpha}(MLE)$	1.814	1.372	1.229	1.160	1.119
Other Circuits					
"Best" (UMVU)	1.791	1.350	1.280	1.149	1.121
P_1 $\hat{\alpha}(MM\chi^2)$	1.849	1.389	1.237	1.163	1.118
P_1 $\hat{\alpha}(MLE)$	1.847	1.388	1.237	1.162	1.118
P_2 $\hat{\alpha}(MM\chi^2)$	1.825	1.380	1.234	1.164	1.122
P_2 $\hat{\alpha}(MLE)$	1.824	1.379	1.234	1.163	1.122

Table 3 : Examining the Exponential & Bilinear Fits.

25

t	2	4	6	8	10	12	15	20	25	30
$P_1(t;.09)$ $= 1-e^{-.09t}$	.165	.302	.417	.513	.593	.660	.741	.835	.895	.933
$P_2(t;18)$ $= \dfrac{t}{18+t}$	.1	.182	.25	.308	.357	.40	.455	.526	.581	.625
P_1/P_2	1.65	1.66	1.67	1.67	1.66	1.65	1.63	1.58	1.54	1.49

Table 6 : Comparing P_1 and P_2 .

References

Cox, D.R. and Hinkley, D. (1974). *Theoretical Statistics*, Chapman and Hall, London.

Cox, D.R. and Lewis, P.A.W. (1966). *The Statistical Analysis of Series of Events*, Methuen, London.

Lindley, D.V. (1950). "Grouping Corrections and Maximum Likelihood Equations", *Proc. Camb. Philos. Soc.* 46, 106-110.

Schuchman, P. and Gelfand, A.E. (1980). "Can Judges Select Cases of 'No Precedential Values'?", *Emery University Law Review*.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 317	2. GOVT ACCESSION NO. AD-A115 346	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AN ESTIMATION PROBLEM WITH POISSON PROCESSES		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) ALAN E. GELFAND		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0475
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office Of Naval Research Statistics & Probability Program Code 411SP Arlington, VA 22217		12. REPORT DATE April 20, 1982
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Poisson Processes, Intensity Function, Maximum Likelihood Estimation, "Correction" Factor, Relative Efficiency.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE REVERSE SIDE.

#317

AN ESTIMATION PROBLEM WITH POISSON PROCESSES

This paper considers an estimation problem involving n independent Poisson processes such that the i-th process has intensity function $\lambda_i(t) = \delta_i \, \rho(t;\alpha)$. It is of interest to estimate $\rho(t;\alpha)$. Two estimation procedures are developed, one using the exact arrival times of observations, the second using categorical arrival times of observations. Two specific instances of $\rho(t)$, an exponential and a bilinear form are investigated further. An example applying the methodology to the active life of a judicial opinion is described.