Regression Estimates of Inputs to an $M(t)/G/\infty$ Service System

Richard L. Patterson

School of Natural Resources
The University of Michigan
Ann Arbor, Michigan 48109

ABSTRACT

Three problems are solved: proof of convergence in distribution of a sequence $\{N^k(t)\}$ of nonhomogeneous Poisson random variables to a nonhomogeneous Poisson random variable $N(t)$; construction of a sequence of multiple linear regression models whose conditional expectations equal $E(N^k(t))$ $(k=1,2,\dots)$ aside from an additive constant; exploration of L_1 and L_2 criteria for estimating parameters of the regression models. Presentation of a numerical analysis of a case study involving a $M(t)/M/\infty$ system of reproduction (arrivals) and mortality (services) within a biological population concludes the paper. Potential problems of interpreting parameter estimates obtained from linear programming implementations of the L_1 fitting criterion are analyzed.

INTRODUCTION AND OBJECTIVE

Models of $M(t)/G/\infty$ stochastic service systems, particularly $M/M/\infty$ models, have been used widely in analyses of flows of populations of subassemblies during manufacture, movements of patients in medical facilities, messages in communications networks, and animals of given species in a variety of habitats. In many instances direct counts of arrivals or departures of units are costly, inconvenient, or impossible to obtain, whereas counts or measurements related to numbers of incomplete services at points in time are available or can be obtained at relatively low cost.

The problem addressed is to numerically approximate an unknown input intensity function $\lambda(t)>0$ in an interval $(0<t\leqslant T)$ of a nonhomogeneous Poisson arrival process to a $M(t)/G/\infty$ service system using error prone measurements $\{n(t_{i1}), n(t_{i2}),\dots, n(t_{iM})\}$ which are counts or averages of counts of units with incomplete service at time instants $0<t_{i1}<t_{i2}<\cdots<t_{iM}\leqslant T$.

APPLIED MATHEMATICS AND COMPUTATION 24:47–63 (1987) 47

0096-3003/87/$03.50

Three questions are addressed, called Problems 1, 2, and 3:

PROBLEM 1. Construction of a sequence $\{N^1(t), N^2(t), \ldots, N^k(t)\}$ of Poisson distributed random variables which converge *in distribution* to the Poisson distributed random variable $N(t)$ = number of incompletely serviced units in the $M(t)/G/\infty$ system at time t $(0 < t \leqslant T)$.

PROBLEM 2. Construction of a sequence of linear regression models which correspond to the sequence $\{N^j(t)\}$ $(j = 1, 2, \ldots, k)$ and where the expectations of the dependent variables, apart from additive constant terms representing nonsampling errors, are equal respectively to the expectation functions $E(N^1(t)), \ldots, E(N^k(t))$ for $0 < t \leqslant T$.

PROBLEM 3. Numerical estimation of the coefficients of the regression models using L_1 and L_2 criteria.

METHODOLOGY

The $M(t)/G/\infty$ model assumes:

(i) a common service time distribution $B(x)$ of all arrivals, in which all service times are i.i.d. random variables, independent of all arrival times;

(ii) a nonhomogeneous Poisson process of arriving units $\{Y(t);\ t > 0,\ Y(0) = 0,\ \lambda(t) > 0\}$, where

$$\lambda(t) = \frac{d}{dt} E(Y(t)) > 0$$

is an arrival intensity function. $Y(t)$ is the number of arrivals in $(0, t]$.

From (i) and (ii) it follows that the number of units in the system at time t with incomplete service is described by a nonhomogeneous Poisson process $\{N(t);\ t > 0,\ N(0) = 0\}$ with expectation function

$$E(N(t)) = \int_0^t \lambda(x)[1 - B(t - x)]\, dx,$$

where:

(i) the initial condition $P(N(0) = 0) = 1$ is assumed for convenience;

(ii) $\lambda(x)\, dx\, [1 - B(t - x)]$ is the approximate probability of the joint event that an arrival occurs in $(x, x + dx)$ and does not complete service in the ensuing time interval of length $t - x$.

Let $\{n(t_{i1}), n(t_{i2}),\dots, n(t_{iM});\ 0 < t_{i1} < \cdots < t_{iM} < T\}$ denote a sequence of M measurements obtained at instants $t_{i1},\dots, t_{iM}$. The measurements:

(i) may be statistically correlated;
(ii) may contain both sampling and nonsampling errors;
(iii) may be obtained when the process is in an evolutionary phase.

Problem 1 is solved by constructing a sequence of non-homogeneous Poisson processes $\{N^j(t);\ t > 0,\ N^j(0) = 0,\ j = 1,2,\dots, k\}$ $(k = 1,2,\dots)$ whose expectation functions $g^1(t) = E(N^1(t)),\dots, g^k(t) = E(N^k(t))$ converge under appropriate conditions to the expectation function $E(N(t))$ of the Poisson random variable $N(t)$ throughout an interval $(0, T]$ containing the instants $t_{i1},\dots, t_{iM}$ of measurements. Each function $g^j(t)$ is a piecewise linear approximation to $E(N(t))$.

Problem 2 is solved by constructing, for each k, a multivariate linear regression model

$$\mathbf{n} = \mathbf{X}\cdot\boldsymbol{\lambda} + \mathbf{E}, \tag{1}$$

where:

(i) $\mathbf{n}$ is an $M\times 1$ column vector of random variables: $\mathbf{n} = (n(t_{i1}),\dots, n(t_{iM}))^T$.

(ii) $\mathbf{X}$ is an $M\times k$ matrix of mathematical variables.

(iii) $\boldsymbol{\lambda}$ is a $k\times 1$ column vector of unknown constants $(\lambda_1,\dots \lambda_k)^T$.

(iv) $\mathbf{E}$ is an $M\times 1$ column vector of random variables having (for fixed k) constant means μ and variances σ^2, where nonsampling error is implied if $\mu \neq 0$.

(v) $E(\mathrm{n}/\mathbf{X}) = [g^k(t_{i1}) + u,\dots, g^k(t_{iM}) + \mu]^T$ (M components) for $k = 1,2,\dots$.

Problem 3 is solved by estimating the $k+1$ unknown constants $\mu, \lambda_1,\dots, \lambda_k$ (or k constants $\lambda_1,\dots, \lambda_k$ if μ is assumed to be zero) appearing in (1) by means of:

(i) a linear goal programming implementation of an L_1 fitting criterion, and

(ii) a standard unconstrained least squares minimization technique implementing the L_2 fitting criterion.

SEQUENTIAL APPROXIMATIONS TO $E(N(t))$

Construction of the sequence stated as Problem 1 proceeds in four steps:

Step 1

Let $(0, T]$ be a fixed time interval, and let $0 < t_1 < T$ be an instant in the interval. Let $\{\mathbf{X}(t);\ 0 < t \leqslant T,\ \mathbf{X}(0) = 0,\ \lambda > 0)$ be a homogeneous Poisson arrival process to a $M/G/\infty$ system. Modify the random variable $\mathbf{X}(t)$ in the subinterval $(t_1, T]$ by redefining $\mathbf{X}(t)$ as: $\mathbf{X}(t) = \mathbf{X}(t_1)$ for $t_1 < t \leqslant T$. Define a new random process in $(0, T]$ as

$$\{\mathbf{X}_1(t);\, 0 < t \leqslant T,\, \mathbf{X}(0) = 0,$$

$$\mathbf{X}_1(t) = \mathbf{X}(t) \text{ for } 0 < t \leqslant t_1$$

$$\text{and } \mathbf{X}_1(t) = \mathbf{X}(t_1) \text{ for } t_1 < t \leqslant T\}.$$

The Poisson distributed random variable $\mathbf{X}_1(t)$ has expectation function

$$E(\mathbf{X}_1(t)) = \begin{cases} 0 & (t \leqslant 0), \\ \lambda_1 t & (0 < t \leqslant t_1), \\ \lambda_1 t_1 & (t_1 < t \leqslant T). \end{cases}$$

Using the process $\{\mathbf{X}_1(t)\}$ as input, define a random process (which can be shown to be nonhomogeneous Poisson) $\{N_1(t);\ 0 < t \leqslant T,\ N_1(0) = 0\}$ describing the number of incomplete services at time t. The expectation function of the random variable $N_1(t)$ is

$$E(N_1(t)) = \begin{cases} 0 & (t \leqslant 0), \\ \lambda_1 \int_0^t [1 - B(x)]\, dx & (0 < t \leqslant t_1), \\ \lambda_1 \int_0^{t_1} [1 - B(t - x)]\, dx & (t_1 < t \leqslant T). \end{cases} \tag{2}$$

The Poisson properties of $N_1(t)$ with expectation function (2) are most easily demonstrated using the compound distribution approach outlined by Newell [1] and others.

Define next a random process $\{N^1(t) = N_1(t);\ 0 < t \leqslant T,\ N^1(0) = 0\}$. $\{N^1(t)\}$ is taken as a first approximation ($k = 1$) to $\{N(t);\ t > 0,\ N(0) = 0\}$, describing the number of incomplete services in a $M(t)/G/\infty$ system with nonhomogeneous arrival process $\{\mathbf{Y}(t);\ 0 < t \leqslant T,\ \mathbf{Y}(0) = 0,\ \lambda(t)\}$ defined over $(0, T]$, where $0 < t_1 < T$. The process $\{N^1(t)\}$ may provide only a crude approximation of the number of incomplete services in $M(t)/G/\infty$. Its accuracy depends upon the behavior of $\lambda(t)$ over $0 < t \leqslant T$, the choice of time instant t_1 $(0 < t_1 < T)$, and the choice of λ_1. Differences between the random variables $N^k(t)$ and $N_k(t)$ become obvious for $k = 2, 3, \ldots$.

Step 2

Let $0 < t_1 < t_2 < \cdots < t_k \leqslant T$ be k $(k = 1, 2, \ldots)$ arbitrary but fixed instants in time. Define a sequence of mutually independent arrival processes

$\{\mathbf{X}_1(t)\},\ldots,\{\mathbf{X}_k(t)\}$ which have the following properties:

(i) $\mathbf{X}_i(t)=0$ for $0\leqslant t\leqslant t_{i-1}$, $i=1,2,\ldots,k$ $(t_0=0)$.

(ii) $\mathbf{X}_i(t)=\mathbf{X}(t_i)$ for $t_i<t\leqslant T$.

(iii) $\mathbf{X}_i(t)$ is a Poisson distributed random variable with expectation function

$$E(\mathbf{X}_i(t))=\begin{cases}0 & (t\leqslant t_{i-1}),\\ \lambda_i\cdot(t-t_{i-1}) & (t_{i-1}<t\leqslant t_i),\\ \lambda_i\cdot(t_i-t_{i-1}) & (t_i<t\leqslant T).\end{cases}\tag{3}$$

Define a sequence of random processes

$$\left\{\mathbf{X}^k(t);\,0<t\leqslant T,\mathbf{X}^k(0)=0;\,\lambda_1,\lambda_2,\ldots,\lambda_k\right\}\qquad(k=1,2,\ldots)$$

which have the following properties:

(i) $\mathbf{X}^k(t)=\mathbf{X}_1(t)+\cdots+\mathbf{X}_k(t)$ $(k=1,2,\ldots)$.

(ii) $\mathbf{X}^k(t)$ is a Poisson distributed r.v. with continuous, piecewise linear expectation function

$$E(\mathbf{X}^k(t))=\begin{cases}0 & (t\leqslant 0),\\ \lambda_1 t & (0<t\leqslant t_1),\\ \lambda_1t_1+\lambda_2\cdot(t-t_1) & (t_1<t\leqslant t_2),\\ \vdots & \\ \lambda_1t_1+\lambda_2\cdot(t_2-t_1)+\cdots+\lambda_k\cdot(t-t_{k-1}) & (t_{k-1}<t\leqslant t_k),\\ \lambda_1t_1+\lambda_2\cdot(t_2-t_1)+\cdots+\lambda_k\cdot(t_k-t_{k-1}) & (t_k<t\leqslant T).\end{cases}\tag{4}$$

For each $k=1,2,\ldots$, $\mathbf{X}^k(t)$ is a r.v. that approximates the nonhomogeneous Poisson distributed r.v. $\mathbf{Y}(t)$ denoting the cumulative number of arrivals to $M(t)/G/\infty$ in $(0,t]$. The proximity of $\mathbf{X}^k(t)$ to $\mathbf{Y}(t)$ depends upon $\lambda(t)$, the choice of k, the choices of instants $t_1,\ldots,t_k$, and the choices of $\lambda_1,\ldots,\lambda_k$.

Step 3

Construct k independent Poisson processes $\{N_i(t);\,0<t\leqslant T,\ N_i(t)=0$ for $t\leqslant t_{i-1}\}$ $(i=1,2,\ldots,k)$ describing the numbers $N_i(t)$ of incomplete services at time t, where:

(i) arrivals are generated by the respective processes $\{\mathbf{X}_i(t)\}$ $(i=1,2,\ldots,k)$ defined above;

(ii) $B(x)$ is the common c.d.f. for all i.i.d. service times, independent of all arrival times;

(iii) the random variables $N_i(t)$ have expectation functions

$$E(N_i(t)) = \begin{cases} 0 & (t \leqslant t_{i-1}), \\ \lambda_i \int_0^{t-t_{i-1}} [1 - B(x)]\, dx & (t_{i-1} < t \leqslant t_i), \\ \lambda_i \int_0^{t_i - t_{i-1}} [1 - B(t - t_{i-1} - x)]\, dx & (t_i < t \leqslant T) \end{cases}$$

$$(i = 1, 2, \ldots, k). \quad (5)$$

Step 4

Define Poisson distributed random processes $\{N^k(t); 0 < t \leqslant T, N^k(0) = 0; \lambda_1, \ldots, \lambda_k\}$ $(k = 1, 2, \ldots)$ where

(i) $N^k(t) = N_1(t) + \cdots + N_k(t)$;

(ii) the random variable $N^k(t)$ has the continuous, piecewise linear expectation function

$$g^k(t) = E(N^k(t)) =$$

$$\begin{cases} 0 & (t \leqslant 0), \\ \lambda_1 \int_0^t [1 - B(x)]\, dx & (0 < t \leqslant t_1), \\ \lambda_1 \int_0^{t_1} [1 - B(t - x)]\, dx & \\ \quad + \lambda_2 \int_0^{t - t_1} [1 - B(x)]\, dx & (t_1 < t \leqslant t_2), \\ \vdots & \\ \lambda_1 \int_0^{t_1} [1 - B(t - x)]\, dx & \\ \quad + \lambda_2 \int_0^{t_2 - t_1} [1 - B(t - t_1 - x)]\, dx & \\ \quad + \cdots + \lambda_{k-1} \int_0^{t_{k-1} - t_{k-2}} [1 - B(t - t_{k-2} - x)]\, dx & \\ \quad + \lambda_k \int_0^{t - t_{k-1}} [1 - B(x)]\, dx & (t_{k-1} < t \leqslant t_k), \\ \lambda_1 \int_0^{t_1} [1 - B(t - x)]\, dx & \\ \quad + \cdots + \lambda_k \int_0^{t_k - t_{k-1}} [1 - B(t - x)]\, dx & (t_k < t \leqslant T). \end{cases}$$

$$(6)$$

It remains to show that the sequence of expectation functions $g^1(t),\ldots,g^k(t)$ converges *in distribution* to the expectation function $E(N(t))$ of the Poisson process describing the number of incomplete services in $M(t)/G/\infty$ at time t $(0<t\leqslant T)$.

CONVERGENCE OF THE SEQUENCE $\{g^i(t)\}$

Let $\lambda(t)$ be a continuous, positive function defined on $(0,T]$. Let the sequence of random variables $N^1(t),\ldots,N^k(t)$ be given as defined above. Let the service time distribution $B(x)$ be differentiable on $(0,T]$, and assume $B(0)=0$. Let $t_1^*, t_2^*,\ldots,t_k^*$ be instants in time specified so that

(i) $0<t_1^*<t_1<t_2^*<t_2<\cdots<t_{k-1}<t_k^*<t_k\leqslant T$;
(ii) $t_1,\ldots,t_k$ are the instants defined in step 2 above;
(iii) $\lambda_1=\lambda(t_1^*)$, $\lambda_2=\lambda(t_2^*),\ldots,$ $\lambda_k=\lambda(t_k^*)$.

In the equations (6) rewrite the terms

$$\int_0^{t_j-t_{j-1}}\left[1-B(t-t_{j-1}-x)\right]dx$$

in the equivalent form

$$\int_{t_{j-1}}^{t_j}[1-B(t-x)]\,dx.$$

Making the change of variable $t-x=y$ and substituting into the equations (6), they assume the form

$$E\left(N^j(t)\right)=\sum_{i=1}^{j}\lambda(t_i^*)\left(t_i-t_{i-1}-\int_{t-t_i}^{t-t_{i-1}}B(y)\,dy\right) \tag{7}$$

By the law of the mean:,

$$\int_{t-t_i}^{t-t_{i-1}}B(y)\,dy=B(\bar{t}_i)(t_i-t_{i-1}),$$

where $\bar{t}_i$ is a number in the subinterval $t-t_i<\bar{t}_i<t-t_{i-1}$. Substituting into (7) and letting $j=k$, we have

$$g^k(t)=E\left(N^k(t)\right)=\sum_{i=1}^{k}\lambda(t_i^*)(t_i-t_{i-1})\left[1-B(t_i)\right].$$

As $k \to \infty$ and $\max(t_i - t_{i-1}) \to 0$,

$$g^k(t) \to \int_0^t \lambda(x)\,dx\,[1 - B(t-x)] = E(N(t)) \tag{8}$$

for $0 < t \leqslant T$, which proves convergence in distribution of the Poisson sequence $N^1(t),\ldots,N^k(t)$ to $N(t)$ on $(0,T]$.

The equations (6) are thus established under the stated conditions as a sequence of piecewise linear approximations converging to the mean and variance of the random variable $N(t)$ describing the number of incomplete services in the system $M(t)/G/\infty$ at time t.

REGRESSION ESTIMATES OF $\lambda_1,\ldots,\lambda_k$

Assume:

(1) the multivariate linear regression model specified by Equation (1) and conditions (i)–(v) below it;

(2) measurements $n(t_{i1}),\ldots,n(t_{iM})$ obtained at times $t_{i1},\ldots,t_{iM}$ [in particular, $n(t_{ij})$ may be a count or an average of repeated, error prone counts of incomplete services at time t_{ij}];

(3) choices of an integer k and time instants $0 < t_1 < \cdots < t_k$ which define subintervals of lengths $t_1, t_2 - t_1,\ldots,t_k - t_{k-1}$.

Assume also that the M measurements are distributed across the k subintervals $(0,t_1),(t_1,t_2),\ldots,(t_{k-1},t_k)$ as follows:

$$\begin{aligned}
\left(n(t_{i1}),\ldots,n(t_{ij_1})\right) &\in (0,t_1] && (j_1 \text{ measurements}),\\
\left(n\left(t_{i\,j_1+1}\right),\ldots,n\left(t_{ij_2}\right)\right) &\in (t_1,t_2] && (j_2 - j_1 \text{ measurements}),\\
&\ \ \vdots\\
\left(n\left(t_{i\,j_{k-1}+1}\right),\ldots,n(t_{iM})\right) &\in (t_{k-1},t_k] && (M - j_{k-1} \text{ measurements}),
\end{aligned}$$

where $j_1 \geqslant 1,\ldots,M - j_{k-1} \geqslant 1$.

If the mean μ of the common c.d.f. of random variables $\mathbf{E} = (E_1,\ldots E_k)^T$ is hypothesized to be nonzero (nonsampling error present), the coefficients $\lambda_1,\ldots,\lambda_k,\mu$ are estimated by solving the equations (9) [which follow from the conditional expectation of (1)] for $\lambda_1,\ldots,\lambda_k,\mu$:

$$\begin{aligned}
\lambda_1 x_{11} + \cdots + \lambda_k x_{1k} + \mu &= n(t_{i1}),\\
&\ \ \vdots\\
\lambda_1 x_{M1} + \cdots + \lambda_k x_{Mk} + \mu &= n(t_{iM}),
\end{aligned} \tag{9}$$

where the x_{ij}'s are nonnegative coefficients specified in Table 1.

TABLE 1

MATHEMATICAL COEFFICIENTS OF THE EQUATIONS (9)

	Col. 1	Col. 2		Col. $k-1$	Col. k
Row 1	$\int_0^{t_{i1}}[1-B(x)]\,dx$	0	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
Row j_1	$\int_0^{t_{i,j_1}}[1-B(x)]\,dx$	0	$\cdots$	0	0
Row j_1+1	$\int_0^{t_1}[1-B(t_{i,j_1+1}-x)]\,dx$	$\int_0^{t_{i,j_1+1}-t_1}[1-B(x)]\,dx$	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
Row j_2	$\int_0^{t_1}[1-B(t_{i,j_2}-x)]\,dx$	$\int_0^{t_{i,j_2}-t_1}[1-B(x)]\,dx$	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
Row $j_{k-1}+1$	$\int_0^{t_1}[1-B(t_{i,j_{k-1}+1}-x)]\,dx$	$\int_0^{t_2-t_1}[1-B(t_{i,j_{k-1}+1}-t_1-x)]\,dx$	$\cdots$	$\int_0^{t_{k-1}-t_{k-2}}[1-B(t_{i,j_{k-1}+1}-t_{k-2}-x)]\,dx$	$\int_0^{t_{i,j_{k-1}+1}-t_{k-1}}[1-B(x)]\,dx$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
Row M	$\int_0^{t_1}[1-B(t_{iM}-x)]\,dx$	$\int_0^{t_2-t_1}[1-B(t_{iM}-t_1-x)]\,dx$	$\cdots$	$\int_0^{t_{k-1}-t_{k-2}}[1-B(t_{iM}-t_{k-2}-x)]\,dx$	$\int_0^{t_{iM}-t_{k-1}}[1-B(x)]\,dx$

If nonsampling error is assumed to be absent, then μ is set equal to zero in (9) prior to solving.

NUMERICAL ESTIMATION OF PARAMETERS

The case of interest is when the equations (9) are overdetermined ($M > k+1$). MSSD (minimizing sum of squared deviations: L_2 norm) and MSAD (minimizing sum of absolute deviations: L_1 norm) are both useful fitting criteria. Two numerical methods using the L_1 norm are descent techniques [2] and linear goal programming [3].

MSSD Techniques

Aside from problems of multicollinearity in (9), which is most likely to occur if measurements are obtained at times t_{ij} and $t_{i,j+1}$ close together (a good rule of thumb is to select times t_{ij} such that all differences $t_{i,j+1} - t_{ij}$ satisfy $t_{i,j+1} - t_{ij} \geqslant \int_0^\infty [1 - B(x)]\,dx$), one or more of the estimates $\hat{\lambda}_j$ may be nonpositive. An estimate $\hat{\lambda}_j = 0$ for parameter λ_j may be interpreted in one of two ways:

(i) λ_j is positive but may or may not be close to zero. Setting $\hat{\lambda}_j = 0$ simply guarantees no contribution of arrivals in (t_{j-1}, t_j) from the fitted model of arrivals (under MSAD solution methods there may also be implications of nonuniqueness of solution, as explained below.

(ii) λ_j is zero, implying a gap in the interval (t_{j-1}, t_j) in which no arrivals are generated. No mathematical problem occurs in the solution of the equations (9) if λ_j is set equal to zero.

If $\hat{\lambda}_j < 0$, there can be two interpretations:

(i) for $\hat{\lambda}_j$ close to zero, it may be inferred that the actual value of λ_j is likely to be nonnegative, subject to a test of the hypothesis $H_0: \lambda_j \geqslant 0$ against $H_a: \lambda_j < 0$. If the test supports the null hypothesis then the model's nonnegativity requirement on the λ_j's can be satisfied by $\hat{\lambda}_j = 0$, in which case there are two alternatives available with respect to the remaining $\hat{\lambda}$'s:

(1) Allow the $\hat{\lambda}_i$'s $(i \neq j)$ to remain equal to their numerical values obtained in the least squares solution, which means that solution is now an approximate one;

(2) recompute a least squares solution for the λ_i's where λ_j is set equal to zero, thereby preventing the jth column in Table 1 from affecting the remaining estimates of the λ_i's. The effect upon the revised $\hat{\lambda}_i$'s will depend upon which column is eliminated. A potential difficulty

with this alternative is that a recomputation of the least squares estimates may again give one or more negative estimates among the remaining $k-1$ λ_j's.

(ii) If $\hat{\lambda}_j < 0$, it may be inferred that λ_j is indeed negative, an inference which violates a model assumption. There are two options available:

(1) the solution, including negative estimates $\hat{\lambda}_j$, may be saved as a trial initial solution in another optimization code for computing estimates $\hat{\lambda}_1, \ldots, \hat{\lambda}_k, \hat{\mu}$;
(2) the intervals $(0, t_1], \ldots, (t_{k-1}, t_k]$ may be redefined, possibly requiring a different value of k, and the estimates $\hat{\lambda}_j$ are recomputed.

If the exercise of a least squares optimizing code yields k positive estimates $\hat{\lambda}_1, \ldots, \hat{\lambda}_k$, then the usual array of statistical evaluations of the solution may be applied. Clearly, nonnegativity constraints on the regression parameters are a potential source of difficulties in application of unconstrained least squares estimation techniques.

MSAD Techniques

Of the two main numerical approaches to parameter estimation under the L_1 norm, descent techniques and linear goal programming, only the latter is considered here.

Literature on the application of linear programming to the estimation of statistical model parameters using the MSAD norm have addressed mainly questions of model formulations, efficiency (rates of convergence) of solution codes, and uniqueness of solutions, and to a lesser extent statistical properties of estimates. Much of the existing literature considers cases in which the regression parameters are unconstrained. Reasonably general conclusions about statistical properties of estimates of model parameters in the equations (1) are available, but both theory and experimentation demonstrate their dependence upon the nature of the error term **E**. Rice and White [4] showed that the MSAD criterion yielded parameter estimates which had smaller variances than the MSSD criterion when components of **E** followed a Laplace or any long tailed distribution. They also noted that in the presence of outliers or "wild" points the L_1 norm appeared to be markedly superior among the class of L_q norms ($q \geq 1$). Taylor [5] showed that in the regression model specified by Equation (1), if the distribution of **E** is symmetrical about zero, then minimizing the sum of absolute values of residuals provides an unbiased estimate of μ. Taylor also showed that if Equation (1) contains a constant term and **E** is symmetrically distributed about **0**, then the L_1 estimator of $(\lambda_1, \ldots, \lambda_k, \mu)^T$ is unbiased. All of these results are relevant to the present problem of estimating the parameters $\lambda_1, \ldots, \lambda_k, \mu$ when a linear goal programming method is employed.

Other advantages of a linear programming implementation of an L_1 fitting criterion for estimating $(\lambda_1, \ldots, \lambda_k, \mu)^T$ are the ease with which nonnegativity constraints on the λ_j's can be included and the guarantee of a solution vector $(\hat{\lambda}_1, \ldots, \hat{\lambda}_k, \hat{\mu})^T$ even when multicollinearity exists within the rows in Table 1. Moreover, if all measurements $n(t_{ij})$ are strictly positive, a basic feasible solution can be found where all $\hat{\lambda}_j$'s are positive, due to the nonnegativity of the entries in Table 1 and the presence of a full complement of slack variables in the linear programming model. Yet another advantage is the use of sensitivity analyses in evaluating the robustness of a solution.

A potential difficulty inherent in the linear programming implementation of the L_1 metric is the lack of an *a priori* guarantee of uniqueness of an optimal solution vector. However, even when the full solution vector is not unique, the part giving $\hat{\lambda}_1, \ldots, \hat{\lambda}_k, \hat{\mu}$ may be unique. Consider the following goal program:

$$\min \quad \sum_{i=1}^{M} u_i + v_i$$

subject to

$$\begin{gathered}
\lambda_1 x_{11} + \cdots + \lambda_k x_{1k} + \mu - u_1 + v_1 = n(t_{i1}), \\
\vdots \\
\lambda_1 x_{M1} + \cdots + \lambda_k x_{Mk} + \mu - u_M + v_M = n(t_{iM}), \\
u_i, v_i, \lambda_i \geq 0; \qquad \mu \text{ unconstrained.}
\end{gathered} \tag{10}$$

There are $2M + k + 1$ "decision variables" and M equality constraints. The value of u_j in any tableau is the amount by which the left hand side of the jth equation (when the current solution is substituted) undershoots the value $n(t_{ij})$, while v_j is the amount of overshoot. Multiple optimal solutions occur when a relative cost coefficient of a nonbasis variable in the final tableau is zero. In particular, if one of the variables λ_j is nonbasic in the final tableau (i.e. $\hat{\lambda}_j = 0$) and its *relative cost* is zero, then the part of the solution of interest, i.e., $(\hat{\lambda}_1, \ldots, \hat{\lambda}_k, \hat{\mu})^T$, *may or may not be unique*, reflecting dual degeneracy in the model. If there are no nonbasic variables in the final tableau of (11) that have zero relative cost coefficients, there will be a unique optimum.

Since the multipliers $t_j - t_{j-1}$ of the $\hat{\lambda}_j$'s in the Poisson model may be large or small but are known, it can at least be roughly estimated whether nonuniqueness of a solution to (10) is likely to result in no errors, small errors, or serious errors in the (Poisson distributed) predictions of the cumulative numer of arrivals in $(0, t]$ and the number of incomplete services at time t.

The other form of degeneracy that can occur in a linear program is primal degeneracy (a basic variable in the final tableau has a value of zero). In theory, it can create two problems:

(i) cycling of solutions, which normally does not occur with modern solution codes;

(ii) difficulty of making a "market" interpretation of relative cost coefficients representing the solutions to the dual variables.

The importance of the market interpretation lies in its application to investigating the sensitivity of an optimal basis to perturbations in the measurements $n(t_{ij})$. If primal degeneracy occurs, the sensitivity analysis on right hand side coefficients of (10) cannot be reliably interpreted.

CASE STUDY EXAMPLE

A case study [6] was conducted in which seasonal reproduction (arrivals) to a biological population and mortality (services) were modeled by an $M(t)/M/\infty$ system. The distribution function governing the length of residence time of each individual in a targeted environment was assumed to be negative exponential. Reproduction occurred over a period of approximately 60 days, which was subdivided into $k = 6$ time subintervals. Seventeen ($M = 17$) measurements of average concentrations of individuals were available (Table 2). The lengths of time intervals between measurements varied from 1 to 21 days. A Julian calendar was adopted where days of a calendar year were numbered consecutively from 1 through 365. The period of reproduction was assumed to begin on day 126 and end on day 189. Based upon inspection of data in Table 2, the interval (126, 189) was subdivided into six contiguous segments by time instants $t_0 = 126$, $t_1 = 141$, $t_2 = 157$, $t_3 = 163$, $t_4 = 176$, $t_5 = 187$, $t_6 = 189$. A series of six models were solved, from which numerical estimates of the vector $(\lambda_1, \ldots, \lambda_6)^T$ were obtained. The coefficient μ was assumed to be zero in all cases. Computer runs were organized into groups according to: (i) the value of the mean residence time of an individual in the targeted environment (mean = 10 days, 50 days); (ii) the regression model fitting criterion employed (MSSD, MSAD); (iii) the computer and code employed (three combinations); (iv) the nonnegativity restraints placed upon the coefficients $\lambda_1, \ldots, \lambda_6$ (three combinations).

TABLE 2

WEIGHTED MEAN LARVAL CONCENTRATIONS OF DOROSOMA CEPEDIANUM IN WESTERN LAKE ERIE, 1977

Julian date	Mean concentration (per 100 m^3)	Standard error of mean
119	0	
122	0	
130	0.04	0.04
131	0	
132	0	
153	120.23	27.51
154	552.57	94.38
161	973.29	244.25
163	822.05	70.88
164	545.27	174.46
166	358.76	66.01
172	567.63	37.49
173	163.25	23.94
174	542.21	104.80
176	234.53	35.73
186	48.22	5.04
187	14.57	3.67
188	1.62	0.72
189	3.12	1.44

Table 3 lists computer run numbers and the corresponding combinations of criteria (i)–(iv). Optimized parameter estimates for each run are listed in Table 4.

Out of the six model runs, models 1 and 4 yielded unique (not dual degenerate), feasible (all $\hat{\lambda}_i$'s nonnegative) solutions $(\hat{\lambda}_1, \ldots, \hat{\lambda}_6)^T$. Model 1 is primal degenerate (the basic variables λ_1 and u_3 are equal to zero in final tableau), which sets a caution flag on sensitivity analysis of the optimal solution to variations in the vector of measurements appearing on the right hand side of (10). Although sensitivity analysis on the right hand side coefficients is helpful in checking robustness of specific model solutions to variations in measurements data, particularly when estimates of standard errors of the measurements are available (Table 2), it can be assumed from studies cited above that estimates of λ_j's obtained as a solution to the model (10) are likely to be no more variable than least squares estimates and will be unbiased. Optimal solution vectors to (10) are piecewise constant functions of the measurements data, and sensitivity analysis permits an examination of this function for specific cases.

TABLE 3

ORGANIZATION OF SIX MODELING RUNS

Model run number	Mean residence time (days)	Fitting criterion	Computer/ code	Nonnegativity restraints
1	10	MSAD	IBM PC/ MP1-LPROG	All $\lambda_j \geqslant 0$
2	10	MSAD	Michigan Amdahl/ SIMPLEX 2	$\lambda_1, \lambda_2, \lambda_3 \geqslant 0$; $\lambda_4, \lambda_5, \lambda_6$ unrestricted
3	10	MSSD	Michigan Amdahl/ MIDAS-REGRESS	All λ_j unrestricted
4	50	MSAD	IBM PC/ MPI-LPROG	All $\lambda_j \geqslant 0$
5	50	MSAD	Michigan Amdahl/ SIMPLEX 2	$\lambda_1, \lambda_2, \lambda_3 \geqslant 0$; $\lambda_4, \lambda_5, \lambda_6$ unrestricted
6	50	MSSD	Michigan Amdahl/ MIDAS-REGRESS	All λ_j unrestricted

TABLE 4

OPTIMIZED PARAMETER ESTIMATES FROM SIX MODELING RUNS

	Estimate					
Run No.	Parameter 1 126–141 days	2 141–157	3 157–163	4 163–176	5 176–187	6 187–189
1	0.00	76.01	108.46	1.47	0.00	0.00
2	0.00	76.01	108.46	1.47	− 9.57	− 4.22
3	− 2.98	54.94	117.49	12.11	− 11.57	− 10.11
4	0.01	31.35	0.00	0.00	0.00	0.00
5	0.01	48.25	41.81	− 34.88	− 17.58	− 5.53
6	10.79	26.96	57.17	− 29.08	− 17.34	− 10.14

Model 4 is not primal degenerate. The fit of model 1 is markedly superior to that of model 4 (63 percent increase in minimum value of objective function from model 1 to model 4), so that on the basis of goodness of fit the estimates provided by model 1 are preferred. From a biological basis model 1 is preferred in that a mean residence time of 10 days in the environment for an individual is more plausible than 50 days.

Model 1 was further checked against model 2, which relaxes the nonnegativity restraints on λ_4, λ_5, and λ_6. The optimal value of the objective function was reduced by only 10 percent, and the optimal estimates of $\lambda_1, \lambda_2, \lambda_3$ remained unchanged, indicating that the buildup in concentration of individuals during the first 30 days of reproduction is much more influential in estimating production coefficients than the last 10 days, during which it is probable that few or no new arrivals occurred. Model 2 was both primal and dual degenerate, the latter implying that other optimal solutions exist. Only a further exercise of the model can exhibit alternative optimum vectors, although there is no implication that the estimates $\hat{\lambda}_1, \ldots, \hat{\lambda}_6$ will necessarily differ in other optimal solutions from the values shown in Table 4.

Models 3 and 6 were standard unconstrained multivariate linear regressions from which least squares estimates of parameters were obtained. The overall fit of model 3 was significant ($P = 0.0002$, $R^2 = 0.76$), the difficulty being that negative parameters violate the basic assumption of Poisson arrivals. Estimates of individual parameters varied in significance, with only $\hat{\lambda}_2$ and $\hat{\lambda}_3$ being highly significant. In light of the insight gained from running model 3, variations of model 3 could be introduced by changing the points of subdivision over which the λ_j's are defined, or by leaving the subdivisions unchanged and setting one or more λ_j's equal to zero prior to estimating coefficients, or by a combination of both. The adequacy of fit of model 6 was judged inferior to that of model 3. Additional steps in validation for all models would include analysis of residuals.

CONCLUSIONS

The above method of estimating nonhomogeneous Poisson distributed inputs to an $M(t)/G/\infty$ system is feasible in all respects. Its accuracy in approximating the arrival process depends fundamentally upon the availability and quality of observations of incomplete services, as selection of the parameter k should only be made in light of the data.

REFERENCES

1 G. F. Newell, *Applications of Queuing Theory*, Chapman and Hall, New York, 1982, p. 179.

2 L. A. Josvanger and V. A. Sposito, L_1 norm estimates using a descent approach, in *Computer Science and Statistics: Proceedings of the 15th Symposium on the Interface* (J. E. Gentle, Ed.), North-Holland, 1983.
3 V. Sposito, W. Smith, and G. McCormick, *Minimizing the Sum of Absolute Deviations*, Vandenhoeck and Ruprecht, Göttingen, 1978.
4 J. R. Rice and J. S. White, Norms for smoothing and estimation, *SIAM Rev.* 6:243–256 (1964).
5 L. D. Taylor, Estimation by minimizing the sum of absolute errors, in *Frontiers of Econometrics* (P. Zarembka, Ed.), Academic, 1974.
6 R. L. Patterson, Estimation of seasonal larval production from a stochastic model of abundance, *J. Theoret. Biol.*, submitted for publication.