

LEAST SQUARES ESTIMATION OF NONHOMOGENEOUS POISSON PROCESSES

MICHAEL E. KUHL

*Department of Industrial and Manufacturing Systems Engineering,
Louisiana State University, Baton Rouge, LA 70803-6409, U.S.A.*

JAMES R. WILSON

*Department of Industrial Engineering,
North Carolina State University, Raleigh, NC 27695-7906, U.S.A.*

(September 23, 1999)

We formulate and evaluate weighted least squares (WLS) and ordinary least squares (OLS) procedures for estimating the parametric mean-value function of a nonhomogeneous Poisson process. We focus the development on processes having an exponential rate function, where the exponent may include a polynomial component or some trigonometric components. Unanticipated problems with the WLS procedure are explained by an analysis of the associated residuals. The OLS procedure is based on a square root transformation of the “detrended” event (arrival) times—that is, the fitted mean-value function evaluated at the observed event times; and under appropriate conditions, the corresponding residuals are proved to converge weakly to a normal distribution with mean 0 and variance 0.25. The results of a Monte Carlo study indicate the advantages of the OLS procedure with respect to estimation accuracy and computational efficiency.

Keywords: Counting processes; least squares; parameter estimation; simulation

1. INTRODUCTION

In many simulation studies, we encounter arrival processes having a long-term trend or multiply periodic behavior. A prominent recent example is found in a large-scale simulation model of the organ procurement and transplantation network of the United States that was developed for the United Network for Organ Sharing (UNOS) (Pritsker, 1998). The UNOS Liver Allocation Model (ULAM) is currently being used by UNOS to evaluate alternative liver-allocation policies for the United States. In analyzing ULAM’s arrival streams of liver-transplant donors and patients, we found some arrival rates to exhibit significant growth over time as well as daily, weekly, or annual effects—that is, cyclic patterns of behavior with periods of 1, 7, or 365 days, respectively (Pritsker et al., 1995).

In building ULAM, Pritsker et al. (1995) implemented the modeling, estimation, and simulation procedures introduced by Kuhl, Wilson, and Johnson (1997) for representing a nonhomogeneous Poisson process (NHPP) having an exponential rate function, where the exponent may include a polynomial component or some trigonometric components; and throughout the rest of this paper,

we use the acronym EPTMP to describe rate functions of the form called “Exponential-Polynomial-Trigonometric with Multiple Periodicities.” Kuhl, Wilson, and Johnson (1997) used a maximum likelihood procedure to estimate the parameters of an EPTMP-type rate function. Although this methodology has been found to yield accurate representations of some nonstationary arrival processes, we have encountered the following problems with the maximum likelihood estimation procedure in certain large-scale practical applications such as the development of ULAM:

1. It does not always provide a sufficiently accurate estimate of the underlying arrival process as represented by the historical buildup of arrivals over the relevant observation interval—that is, in some applications the fitted mean-value function does not adequately represent prominent features of the empirical mean-value function over the observation interval.
2. It can require extraordinarily long execution times to yield final parameter estimates, especially in situations requiring the user to fit a large number of separate arrival processes each with a high arrival rate over the observation interval.

Item 1 above is particularly important in the validation of large-scale simulations such as ULAM, where fidelity to the historical record of the stochastic input models driving the simulation is often critical to acceptance of the overall study. Item 2 above can become a dominant consideration with increasing scope and complexity of the simulation input-modeling task. For example, in the development of ULAM we had to estimate separate arrival streams of: (a) organ donors at each of 63 Organ Procurement Organizations, and (b) liver patients at each of 106 transplant centers; furthermore, many of the corresponding data sets consisted of 500 to 1,500 arrivals recorded over a five-year observation interval. These considerations motivated us to seek alternatives to the maximum likelihood procedure for estimating the parameters of an NHPP having an EPTMP-type rate function.

Least squares procedures have been widely used to fit distribution functions to sample data. For example, Swain, Venkatraman, and Wilson (1988) used least squares procedures to estimate the parameters of a cumulative distribution function (c.d.f.) in the univariate Johnson translation system of distributions. Similarly, Wagner and Wilson (1996) found least squares to be an effective and computationally efficient method for fitting a univariate Bézier c.d.f. to sample data. Fitting a mean-value function and a c.d.f. are similar in that both are increasing functions which are fitted to the (possibly rescaled) cumulative frequency of occurrence of relevant observations. Since certain variants of least squares have been used effectively for estimating c.d.f.’s, we sought to develop appropriate least squares procedures for estimating NHPPs. Härtler (1989) provided additional motivation for developing such alternatives to maximum likelihood estimation of NHPPs.

The main objective of this paper is to develop and evaluate computationally efficient weighted least squares (WLS) and ordinary least squares (OLS) procedures for fitting the mean-value function of an NHPP with an EPTMP-type rate function. An analysis of the small- and large-sample properties of the residuals arising in the WLS procedure explains the unanticipated problems we en-

countered in practical applications of this procedure. As an alternative least squares approach that avoids the problems of the WLS procedure, we present an OLS procedure based on a square root transformation of the “detrended” event (arrival) times—that is, the fitted mean-value function evaluated at the observed event times; and under appropriate conditions, we prove that the corresponding residuals converge weakly to a $N(0, \frac{1}{4})$ distribution as their event indexes tend to infinity. The results of a Monte Carlo study are summarized to indicate the advantages and disadvantages of the OLS procedure. Based on the Monte Carlo results as well as our practical experience in using the OLS procedure to fit NHPPs with EPTMP-type rate functions to the UNOS arrival streams of liver- and kidney-transplant donors and patients, we believe that the techniques presented in this paper are capable of adequately modeling a large class of arrival process encountered in large-scale simulation applications.

The rest of this paper is organized as follows. In Section 2, we formulate the WLS and OLS estimation procedures. The numerical methods used to implement these estimation procedures are discussed in Section 3. Section 4 contains a summary of the results of the experimental performance evaluation. In Section 5 we recapitulate the main findings of this research and discuss our recommendations for future work. Although this paper is based on Kuhl (1997), some of our results were also presented in Kuhl, Damerdji, and Wilson (1998).

2. METHODOLOGY

2.1 Basic Nomenclature

We consider counting processes that represent the buildup of events (arrivals) over time. For such processes we are able to observe each arrival time exactly, and in general the arrival intensity (rate) changes over time. Under certain assumptions a nonstationary arrival process can be represented as an NHPP $\{N(t) : t \geq 0\}$, where $N(t)$ is the number of arrivals in the time interval $(0, t]$ for all $t \geq 0$, such that the instantaneous arrival rate at time t , $\lambda(t)$, is a nonnegative integrable function of time; and the corresponding mean-value function is

$$\mu(t) \equiv E[N(t)] = \int_0^t \lambda(z) dz \quad \text{for all } t \geq 0.$$

The probabilistic behavior of an NHPP is completely characterized by its rate or mean-value function.

We seek to develop computationally efficient least squares methods for fitting NHPPs to arrival processes having parametric rate functions such as the EPTMP-type rate function

$$\lambda(t) = \exp\{h(t; m, p, \Theta)\} \quad \text{for all } t \in [0, S], \tag{1}$$

with exponent

$$h(t; m, p, \Theta) = \sum_{i=0}^m \alpha_i t^i + \sum_{k=1}^p \gamma_k \sin(\omega_k t + \phi_k),$$

where

$$\Theta = [\alpha_0, \alpha_1, \dots, \alpha_m, \gamma_1, \dots, \gamma_p, \phi_1, \dots, \phi_p, \omega_1, \dots, \omega_p]$$

is the vector of continuous parameters. When we want to emphasize the dependence on Θ of the rate function (1) and its associated mean-value function, we write these functions as $\lambda(t; \Theta)$ and $\mu(t; \Theta)$, respectively. The primary objective of the least squares procedures is to estimate $\mu(t; \Theta)$ for all $t \in [0, S]$.

For an NHPP $\{N(t) : t \geq 0\}$, we let $\{\tau_i : i = 1, 2, \dots\}$ denote the corresponding (random) arrival times; and we let $\{t_i : i = 1, 2, \dots\}$ denote a (fixed) realization of the arrival-time process (that is, an observed sequence of specific arrival times). If we know the exact value of the parameter vector Θ as well as the form of the mean-value function $\mu(t; \Theta)$, then the error term

$$\varepsilon_i(\Theta) \equiv \mu(\tau_i; \Theta) - E[\mu(\tau_i; \Theta)] \quad \text{for } i = 1, 2, \dots \quad (2)$$

represents random variation of the “detrended” arrival time $\mu(\tau_i; \Theta)$ about its mean; and thus $E[\varepsilon_i(\Theta)] = 0$ for $i = 1, 2, \dots$. If the errors in (2) were independent and identically distributed (i.i.d.), then the usual approach to computing the OLS estimator $\tilde{\Theta}_{\text{OLS}}$ of the parameter vector Θ based on the arrivals in the observation interval $[0, S]$ would be to minimize the error sum of squares

$$\text{SSE}(\hat{\Theta}) = \sum_{i=1}^{N(S)} \left\{ \mu(\tau_i; \hat{\Theta}) - E[\mu(\tau_i; \hat{\Theta})] \right\}^2$$

over all values of $\hat{\Theta}$ so that we would take $\tilde{\Theta}_{\text{OLS}} = \arg \min_{\hat{\Theta}} \text{SSE}(\hat{\Theta})$ (Seber and Wild, 1989).

In the case of an NHPP, the errors in (2) are neither independent nor identically distributed; and in particular, an NHPP has the following probabilistic characterization. If the counting process $\{N(t) : t \geq 0\}$ is an NHPP with rate function $\lambda(t; \Theta)$ and mean-value function $\mu(t; \Theta)$ for all $t \geq 0$, then the epochs $\{\tau_i : i = 1, 2, \dots\}$ are the arrival times of this NHPP if and only if the corresponding “detrended” epochs $\{\tau_i^* = \mu(\tau_i; \Theta) : i = 1, 2, \dots\}$ are the arrival times of a homogeneous Poisson process with rate 1; and in this situation the detrended arrival times $\{\tau_i^* : i = 1, 2, \dots\}$ satisfy

$$E[\tau_i^*] = i \quad \text{for } i = 1, 2, \dots, \quad (3)$$

and

$$\text{Cov}[\tau_i^*, \tau_j^*] = \min\{i, j\} \quad \text{for } i, j = 1, 2, \dots \quad (4)$$

(Çinlar, 1975).

Clearly the error terms defined by (2) have the covariance structure (4) of the detrended arrival times. To exploit this known covariance structure, we developed WLS and OLS procedures for estimating the mean-value function of an NHPP. These methods are examined in the following subsections.

2.2 Weighted Least Squares Estimation of NHPPs

For the ideal situation in which an NHPP with rate function of the form (1) has known parameter vector Θ , the corresponding ideal residuals $\{\varepsilon_i(\Theta) : i = 1, \dots, n\}$ for a fixed sample size n have the covariance matrix \mathbf{V} defined by (4) with inverse $\mathbf{V}^{-1} = [G_{ij}]$ given by

$$G_{ij} = \begin{cases} 2, & \text{if } 1 \leq i = j \leq n - 1, \\ -1, & \text{if } 2 \leq j = i + 1 \leq n \text{ or } 1 \leq j = i - 1 \leq n - 1, \\ 1, & \text{if } i = j = n, \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

(Kuhl, 1997). In the WLS approach to estimating the mean-value function of a target NHPP over the observation interval $[0, S]$, we take $n = N(S)$ so that the error sum of squares

$$\text{SS}_E(\hat{\Theta}) = \varepsilon^T(\hat{\Theta})\mathbf{V}^{-1}\varepsilon(\hat{\Theta}) = \sum_{i,j=1}^{N(S)} G_{ij}\varepsilon_i(\hat{\Theta})\varepsilon_j(\hat{\Theta}) \quad (6)$$

is to be minimized over all feasible values of $\hat{\Theta}$, where the vector of actual residuals $\varepsilon(\hat{\Theta})$ has i th component $\varepsilon_i(\hat{\Theta}) \equiv \mu(\tau_i; \hat{\Theta}) - E[\mu(\tau_i; \hat{\Theta})]$ for $i = 1, 2, \dots$. Assuming that $\hat{\Theta}$ is close to Θ , we use the approximation $E[\mu(\tau_i; \hat{\Theta})] \approx E[\mu(\tau_i; \Theta)] = i = N(\tau_i)$ for $i = 1, 2, \dots$; thus in practice, the i th actual residual represents the discrepancy between the fitted mean-value function $\mu(\tau_i; \hat{\Theta})$ and the empirical mean-value function $N(\tau_i)$ for $i = 1, 2, \dots$.

The error sum of squares (6) may be reexpressed as

$$\text{SS}_E(\hat{\Theta}) = \left| \mathbf{u}(\hat{\Theta}) \right|^2 = \sum_{i=1}^{N(S)} u_i^2(\hat{\Theta}), \quad (7)$$

where: $\mathbf{u}(\hat{\Theta}) \equiv \mathbf{L}^T \varepsilon(\hat{\Theta})$ is a vector of transformed residuals that in ideal circumstances should be uncorrelated with unit variance; and for $n = N(S)$, the $n \times n$ matrix $\mathbf{L} = [L_{ij}]$ defined by

$$L_{ij} = \begin{cases} \sqrt{2}, & \text{if } i = j = 1, \\ -\sqrt{(i-1)/i}, & \text{if } 1 \leq j = i - 1 \leq n - 1, \\ \sqrt{(i+1)/i}, & \text{if } 2 \leq j = i \leq n - 1, \\ 1/\sqrt{n}, & \text{if } i = j = n, \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

yields Cholesky decomposition $\mathbf{V}^{-1} = \mathbf{L}\mathbf{L}^T$. Appendix A details some relevant properties of the transformed residuals that result in unanticipated anomalous behavior of the WLS estimator

$$\tilde{\Theta}_{\text{WLS}} = \arg \min_{\hat{\Theta}} \sum_{i=1}^{N(S)} u_i^2(\hat{\Theta}) \quad (9)$$

for the NHPP parameter vector Θ .

It is clear from Appendix A (in particular, displays (A.3) and (A.4)) that information about the discrepancy between the empirical mean-value function $N(t)$ and the fitted mean-value function $\mu(t; \hat{\Theta})$ has been completely eliminated from all but the last element of $\mathbf{u}(\hat{\Theta})$, the vector of transformed residuals. It follows that even in the ideal situation for which the WLS estimation procedure starts with perfect (error-free) initial estimates of the unknown parameters so that $\hat{\Theta} = \Theta$ and the transformed residuals $\mathbf{u}(\Theta)$ are in fact uncorrelated with unit variance, the value of the objective function (7) contains relatively little information about how closely the current estimate of the mean-value function approximates the empirical mean-value function. Furthermore, in Appendix A we see that each ideal transformed residual $u_i(\Theta)$ converges in distribution to a standard reversed exponential variate W with $E[W] = 0$ and $\text{Var}[W] = 1$ as $i \rightarrow \infty$ provided $i \neq n$; and the last ideal transformed residual $u_n(\Theta)$ converges in distribution to a standard normal variate as $n \rightarrow \infty$. These properties of the WLS residuals explain the problems we have encountered in practical applications of the WLS procedure based on (9).

Kuhl, Damerdjji, and Wilson (1998, p. 639) and Kuhl (1997, p. 53) describe a striking example of the anomalous behavior that can result from using the WLS procedure (9) to fit an NHPP to an observed arrival process. The divergence between the fitted and empirical mean-value functions in this application exemplifies the way in which the WLS estimation procedure can fail in practice.

The first- and second-order moment structure arising in WLS estimation of the mean-value function of an NHPP is similar to the corresponding moment structure that arises in WLS estimation of c.d.f.'s. In particular, compare equations (5) and (8) above to the corresponding results (7) and (8) of Swain, Venkatraman, and Wilson (1988), wherein the latter results were used to formulate a WLS procedure for estimating a c.d.f. in the Johnson translation system of distributions. However, in the case of fitting a Johnson c.d.f., certain constraints must be imposed on the WLS estimation procedure to ensure that the fitted c.d.f. is monotonically increasing with lower and upper limits of 0 and 1, respectively. These constraints prevent the errors observed in WLS estimation of a Johnson c.d.f. from being as pronounced as the errors observed in WLS estimation of an NHPP. Nevertheless, the analysis given in this section can be easily adapted to explain the problems arising in WLS estimation of Johnson c.d.f.'s that were reported by Swain, Venkatraman, and Wilson (1988).

2.3 Ordinary Least Squares Estimation of NHPPs

Because of the fundamental problems that we encountered in using the WLS procedure for estimating NHPPs, we developed an alternative approach based on a variance-stabilizing transformation together with an OLS estimation procedure. In building a statistical model for which the variance of the original response variable is proportional to its mean (as in (3) and (4)), a standard variance-stabilizing transformation is to take the square root of the original response (Box, Hunter, and Hunter, 1978). This suggests using a least-squares formulation based on ideal residuals having the form $\sqrt{\mu(\tau_i; \Theta)} - E[\sqrt{\mu(\tau_i; \Theta)}]$ for $i = 1, 2, \dots$. In Appendix B, we show that as $i \rightarrow \infty$,

$E[\sqrt{\mu(\tau_i; \Theta)}] - \sqrt{i - \frac{1}{4}} \rightarrow 0$ and $\text{Var}[\sqrt{\mu(\tau_i; \Theta)}] \rightarrow \frac{1}{4}$; moreover we show that ideal residuals of the form

$$\eta_i(\Theta) \equiv \sqrt{\mu(\tau_i; \Theta)} - \sqrt{i - \frac{1}{4}} \quad \text{for } i = 1, 2, \dots \quad (10)$$

converge in distribution to a normal distribution with mean zero and variance $\frac{1}{4}$,

$$\eta_i(\Theta) \xrightarrow[i \rightarrow \infty]{\mathcal{D}} N(0, \frac{1}{4}). \quad (11)$$

Thus, we see that the square root transformation does in fact stabilize the variance of the ideal residuals defined by (10); and (11) motivates our formulation of the variance stabilized–OLS estimate for the parameter vector Θ as

$$\tilde{\Theta}_{\text{OLS}} = \arg \min_{\hat{\Theta}} \text{SS}_E(\hat{\Theta}), \quad \text{where } \text{SS}_E(\hat{\Theta}) = \sum_{i=1}^{N(S)} \left(\sqrt{\mu(\tau_i; \hat{\Theta})} - \sqrt{i - \frac{1}{4}} \right)^2. \quad (12)$$

The next step is to identify an efficient numerical procedure for: (a) determining appropriate values for m and p in (1); and then (b) minimizing the sum of squared errors $\text{SS}_E(\hat{\Theta})$ defined in (12) for the given values of m and p .

3. PARAMETER ESTIMATION PROCEDURE

Given the degree m of the polynomial trend component and the number p of periodic components in the EPTMP-type rate function (1), we use the Nelder-Mead simplex search algorithm (Nelder and Mead, 1964; Olsson, 1974; Olsson and Nelson, 1975) to perform the OLS procedure, yielding the final estimator $\tilde{\Theta}_{\text{OLS}}$ in (12). The Nelder-Mead algorithm is a general-purpose, direct-search method for unconstrained optimization of continuous response functions that may be nondifferentiable. We investigated several other optimization methods including the Levenberg-Marquardt algorithm, which is a specialized gradient-search method for least squares problems (Kennedy and Gentle, 1980). On a suite of typical least-squares test problems, we found that the Nelder-Mead algorithm was faster than the Levenberg-Marquardt algorithm while yielding solutions with virtually the same accuracy. Moreover in an extensive evaluation of the performance of the Nelder-Mead algorithm versus the Powell method (1964), the Davidon-Fletcher-Powell method (Davidon, 1959; Fletcher and Powell, 1963), and the truncated-Newton method (Nash, 1984), Flanigan (1993) and Wagner and Wilson (1996) found the Nelder-Mead algorithm to be faster and more stable for least squares estimation of univariate Bézier distributions. These considerations motivated our use of the Nelder-Mead algorithm in implementing the OLS estimation procedure.

The number p of periodic components and the initial estimates for the frequencies $\{\omega_1, \dots, \omega_p\}$ can be obtained either from prior information about the process or from a standard spectral analysis of the series of events (Lewis, 1970). For some illustrative examples, see Lee, Wilson, and Crawford (1991) or Kuhl and Wilson (1996). In many applications, the values of p and the frequencies

$\{\omega_1, \dots, \omega_p\}$ are known at the outset; and throughout the rest of this paper, we assume that these quantities are known and therefore not subject to estimation.

To determine an appropriate value of the degree m of the polynomial trend component, we use a heuristic variant of the likelihood ratio test developed by Kuhl, Wilson, and Johnson (1997). For a fixed value of m , this heuristic likelihood ratio test uses the rapid approximation method detailed in Sections 2.2–2.3 of Kuhl, Wilson, and Johnson (1997) to compute an initial estimate $\hat{\Theta}_m$ of the hypothesized vector Θ_m of unknown continuous parameters. In practice we have found that $\hat{\Theta}_m$ is an excellent approximation to the associated maximum likelihood estimator of Θ_m (in the sense of nearly maximizing the likelihood function) while requiring much less computing time than the maximum likelihood estimator. We sought to exploit this property in our procedure for estimating m .

The approximate likelihood ratio test to determine the final estimate of the trend degree m operates as follows. Suppose n arrivals have been observed at the arrival times $t_1 < t_2 < \dots < t_n$ in the observation interval $[0, S]$ as a realization of an NHPP with a rate function of the form (1). Given $N(S) = n$ and $\mathbf{t} = (t_1, t_2, \dots, t_n)$, for each trial degree m we let $\mathcal{L}_m(\hat{\Theta}_m|n, \mathbf{t})$ denote the corresponding log-likelihood function evaluated at $\hat{\Theta}_m$, the initial estimate of Θ_m based on the rapid approximation technique of Kuhl, Wilson, and Johnson (1997). Under the null hypothesis that the current value of m is the true degree of the trend component of the underlying EPTMP-type rate function (1), the test statistic

$$2\left[\mathcal{L}_{m+1}(\hat{\Theta}_{m+1}|n, \mathbf{t}) - \mathcal{L}_m(\hat{\Theta}_m|n, \mathbf{t})\right] \quad (13)$$

has approximately a chi-squared distribution with one degree of freedom provided S and n are sufficiently large. Thus we use (13) to assess the importance of successive increments of the likelihood function as the degree of the estimated trend component is repeatedly incremented by one. The degree of the fitted EPTMP-type rate function is determined to be the smallest value of m for which the difference (13) is not significant at a prespecified level of significance. The corresponding vector $\hat{\Theta}_m$ provides the initial parameter estimates (starting values) for the Nelder-Mead simplex search algorithm, which in turn yields the final least squares estimate $\tilde{\Theta}_m$ of the vector of unknown continuous parameters.

Ultimately our only justification for the heuristic likelihood ratio test (13) is that we have found it to work reasonably well in a large class of practical applications. Some evidence of the reliability of this test can be found in Table IV of the experimental performance evaluation described in the next section. In general we have observed that (13) performs reliably in data sets of “moderate” to “large” size—that is, in applications with $n \geq 500$ arrivals; and this is precisely the situation that our overall OLS estimation procedure is designed to handle. We use the heuristic likelihood ratio test (13) in preference to Akaike’s Information Criterion (AIC) (Akaike, 1974) for two reasons:

1. In other work we found the AIC to yield overparameterized models (see also p. 175 of Venables and Ripley (1994)), whereas in all of our computational experience (13) largely avoids this

problem.

2. The AIC requires computation of the maximum likelihood estimator of the vector of unknown continuous parameters; and our main motivation for development of the least squares estimator (12) was to avoid the computational overhead inherent in using maximum likelihood estimation in this context.

We implemented the parameter estimation method described above in the public-domain software package `mp31s` (Kuhl and Wilson, 1998); and we used `mp31s` to carry out the experimental performance evaluation described in the next section.

4. EXPERIMENTAL PERFORMANCE EVALUATION

4.1 Generation of Experimental Data

To evaluate the performance of the OLS procedure in estimating an NHPP with rate function of the form (1), we employed seven target NHPPs that represent processes having up to four cyclic components or a general trend over time or both. As summarized in Table I, these cases are based on the set of experimental cases used by Kuhl, Wilson, and Johnson (1997) to evaluate their maximum likelihood estimation procedure for NHPPs with EPTMP-type rate functions. Notice that Cases 1 through 6 in Table I coincide with the similarly numbered cases in Kuhl, Wilson, and Johnson (1997). Case 0 is a EPTMP-type rate function with one periodic component. Cases 1 through 3 consist of exponential rate functions with two periodic components. Cases 0 and 1 do not contain a general trend over time. Cases 2, 3, and 4 contain general trends that are represented by polynomials of degree 1, 2, and 3, respectively. Rate functions of type EPTMP with three and four periodic components and no long-term trend are utilized in Cases 5 and 6, respectively.

In addition to these seven cases, we performed a “Base Case” experiment that is used as a benchmark for comparing the goodness-of-fit statistics computed in the each of the seven main experimental cases. The Base Case indicates the relative magnitude of the various goodness-of-fit statistics that we would obtain if instead of approximating the observed arrival process with an appropriate NHPP model, we were to use a (homogeneous) Poisson process having a constant arrival rate. In the Base Case we generated realizations of the NHPP specified for Case 0, but we simply fitted a homogeneous Poisson process to the observed series of events. Although many of the resulting goodness-of-fit statistics are only directly comparable to their counterparts in Case 0, the Base Case provides some means for judging the relative performance of the OLS procedure (12) when it is used in conjunction with the NHPP model (1) to estimate an arrival process.

As shown in Table I, the frequencies used in the experimentation are expressed in radians per unit time so that $\omega_1 = 2\pi$, $\omega_2 = 4\pi$, $\omega_3 = 8\pi$, and $\omega_4 = \pi$ radians per unit time. If the unit of time is taken to be one year, then these frequencies represent annual, semiannual, quarterly, and biennial effects, respectively.

TABLE I Parameters of NHPPs Used in the Experimental Evaluation.

Parameter	Case						
	0	1	2	3	4	5	6
α_0	3.6269	3.6269	3.6269	3.6269	3.6269	3.6269	3.6269
α_1	—	—	0.1000	-0.1000	-0.4743	—	—
α_2	—	—	—	0.0200	0.0873	—	—
α_3	—	—	—	—	-0.0041	—	—
γ_1	1.0592	1.0592	1.0592	1.0592	1.0592	1.0592	1.0592
ϕ_1	-0.6193	-0.6193	-0.6193	-0.6193	-0.6193	-0.6193	-0.6193
ω_1	6.2831	6.2831	6.2831	6.2831	6.2831	6.2831	6.2831
γ_2	—	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
ϕ_2	—	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
ω_2	—	12.5664	12.5664	12.5664	12.5664	12.5664	12.5664
γ_3	—	—	—	—	—	0.2500	0.2500
ϕ_3	—	—	—	—	—	0.2500	0.2500
ω_3	—	—	—	—	—	25.1327	25.1327
γ_4	—	—	—	—	—	—	0.7500
ϕ_4	—	—	—	—	—	—	0.7000
ω_4	—	—	—	—	—	—	3.1416

Realizations of the selected NHPPs were generated over the observation interval $(0, S]$ using the piecewise inversion algorithm `mp3sim` (Kuhl and Wilson, 1996). For each case, we set $S = 12$ and we generated $K = 100$ independent replications of the corresponding arrival process $\{N(t) : t \in (0, S]\}$. On each replication of each case, we used a significance level of 0.05 in the approximate likelihood ratio test (13) to determine the appropriate degree m of the polynomial trend, where m was limited to the range $0 \leq m \leq 6$; and then we applied the OLS scheme (12) to fit an NHPP to the observed series of event times. The software package `mp3ls` (Kuhl and Wilson, 1998) was invoked on each replication of each case to perform this parameter estimation procedure. The objective of the performance evaluation using `mp3sim` and `mp3ls` was to characterize the statistical goodness of fit and the computational efficiency achieved by the OLS procedure in the chosen suite of seven test problems.

4.2 Formulation of Performance Measures

To evaluate the performance of the OLS estimation procedure, we used both visual-subjective and numerical goodness-of-fit criteria. Some of these numerical performance measures were formulated by Kuhl, Wilson, and Johnson (1997) to evaluate their maximum likelihood estimation procedure for fitting an EPTMP-type rate function. Among these goodness-of-fit statistics are absolute measures of error that are specific to each experiment as well as relative performance measures that can be compared across different experiments. For replication k of a given case ($k = 1, \dots, K$), we let $\tilde{\lambda}_k(t)$ and $\tilde{\mu}_k(t)$ respectively denote the estimated rate and mean-value functions; and we let $\lambda(t)$

and $\mu(t)$ respectively denote the true underlying rate and mean-value functions, where we have suppressed the dependence of these latter functions on the true parameter vector Θ for notational simplicity.

We used the following goodness-of-fit statistics to measure the ability of the OLS procedure to estimate the theoretical rate and mean-value functions of the underlying NHPP in each experimental case. As defined in Kuhl, Wilson, and Johnson (1997), we let

$$\delta_k \equiv \frac{1}{S} \int_0^S \left| \tilde{\lambda}_k(t) - \lambda(t) \right| dt \quad \text{and} \quad \delta_k^* \equiv \max \left\{ \left| \tilde{\lambda}_k(t) - \lambda(t) \right| : 0 \leq t \leq S \right\}$$

respectively denote the average absolute error and maximum absolute error that occur in estimating the rate function $\lambda(t)$ for all $t \in [0, S]$ on the k th replication of the target NHPP; and we let $\bar{\delta}$ and V_δ respectively denote the sample mean and the sample coefficient of variation of the observations $\{\delta_k : k = 1, \dots, K\}$. The statistics $\bar{\delta}^*$ and V_{δ^*} are computed similarly from the observations $\{\delta_k^* : k = 1, \dots, K\}$. Moreover, we let

$$\Delta_k \equiv \frac{1}{S} \int_0^S \left| \tilde{\mu}_k(t) - \mu(t) \right| dt \quad \text{and} \quad \Delta_k^* \equiv \max \left\{ \left| \tilde{\mu}_k(t) - \mu(t) \right| : 0 \leq t \leq S \right\}$$

respectively denote the average absolute error and maximum absolute error that occur in estimating the mean-value function $\mu(t)$ for all $t \in [0, S]$ on the k th replication of the target NHPP; and we let $\bar{\Delta}$ and V_Δ respectively denote the sample mean and the sample coefficient of variation of the observations $\{\Delta_k : k = 1, \dots, K\}$. The statistics $\bar{\Delta}^*$ and V_{Δ^*} are computed similarly from the observations $\{\Delta_k^* : k = 1, \dots, K\}$. As in Kuhl, Wilson, and Johnson (1997), we also report the “normalized” statistics

$$Q_\delta \equiv \frac{\bar{\delta}}{\mu(S)/S}, \quad Q_{\delta^*} \equiv \frac{\bar{\delta}^*}{\mu(S)/S}, \quad Q_\Delta \equiv \frac{\bar{\Delta}}{\frac{1}{S} \int_0^S \mu(t) dt}, \quad \text{and} \quad Q_{\Delta^*} \equiv \frac{\bar{\Delta}^*}{\frac{1}{S} \int_0^S \mu(t) dt}$$

to facilitate comparison of results for different cases.

In addition to goodness-of-fit statistics that measure the ability of the OLS procedure to estimate the theoretical rate and mean-value functions of the underlying NHPP, we formulated statistics that measure the ability of the OLS procedure to approximate each observed arrival process. On the k th replication of a given NHPP ($k = 1, 2, \dots, K$), we let $\{t_{i,k} : i = 1, 2, \dots, N_k(S)\}$ denote the arrival epochs observed in the time interval $[0, S]$. Thus for $k = 1, 2, \dots, K$, the k th replication of the sum of squared OLS estimation errors and the mean squared OLS estimation error are respectively given by

$$\text{SS}_E(\tilde{\Theta})_k \equiv \sum_{i=1}^{N_k(S)} \left(\sqrt{\tilde{\mu}_k(t_{i,k})} - \sqrt{i - \frac{1}{4}} \right)^2 \quad \text{and} \quad \text{MS}_E(\tilde{\Theta})_k \equiv \text{SS}_E(\tilde{\Theta})_k / N_k(S).$$

We let $\overline{\text{SS}}_E$ and V_{SS_E} respectively denote the sample mean and the sample coefficient of variation of the observed values $\{\text{SS}_E(\tilde{\Theta})_k : k = 1, 2, \dots, K\}$. Similarly, we let $\overline{\text{MS}}_E$ and V_{MS_E} respectively

denote the sample mean and the sample coefficient of variation of the observed values $\{\text{MS}_E(\tilde{\Theta})_k : k = 1, 2, \dots, K\}$.

The average absolute error and maximum absolute error that occur in estimating the empirical mean-value function on the k th replication are respectively given by

$$D_k \equiv \frac{1}{N_k(S)} \sum_{i=1}^{N_k(S)} \left| \tilde{\mu}_k(t_{i,k}) - i \right| \quad \text{and} \quad D_k^* \equiv \max \left\{ \left| \tilde{\mu}_k(t_{i,k}) - i \right| : 1 \leq i \leq N_k(S) \right\}$$

for $k = 1, 2, \dots, K$. We let \overline{D} denote the sample mean of the observed values $\{D_k : k = 1, 2, \dots, K\}$; and we let \overline{D}^* denote the sample mean of the observed values $\{D_k^* : k = 1, 2, \dots, K\}$. We also formulated two types of aggregate performance measures to compare \overline{D} and \overline{D}^* across experiments. The first type uses the grand average level of the empirical mean-value functions computed over all K replications to normalize the average performance measures \overline{D} and \overline{D}^* so that we take

$$Q_D = \frac{\overline{D}}{\frac{1}{K} \sum_{k=1}^K \frac{1}{S} \int_0^S N_k(t) dt} \quad \text{and} \quad Q_{D^*} = \frac{\overline{D}^*}{\frac{1}{K} \sum_{k=1}^K \frac{1}{S} \int_0^S N_k(t) dt}.$$

The second type of aggregate performance measure is calculated by expressing each performance measure D_k and D_k^* observed on the k th replication as a percentage of the average level of the empirical mean-value function on that replication; then the resulting normalized statistics are averaged over all K replications, yielding

$$H_D = \frac{1}{K} \sum_{k=1}^K \frac{D_k}{\frac{1}{S} \int_0^S N_k(t) dt} \quad \text{and} \quad H_{D^*} = \frac{1}{K} \sum_{k=1}^K \frac{D_k^*}{\frac{1}{S} \int_0^S N_k(t) dt},$$

respectively.

In addition to numerical goodness-of-fit statistics, we used graphical methods to provide a visual means of determining the quality of the estimates. For each case, the underlying theoretical rate (respectively, mean-value) function was graphed along with a tolerance band for the estimated rate (respectively, mean-value) function. Approximate tolerance bands for the rate function $\lambda(t)$, $t \in (0, S]$, were obtained as follows. For each fixed time $t \in (0, S]$, let

$$\tilde{\lambda}_{(1)}(t) < \tilde{\lambda}_{(2)}(t) < \dots < \tilde{\lambda}_{(K)}(t)$$

denote the ordered estimates of $\lambda(t)$ obtained on all K replications of the estimation procedure. Thus an approximate $100(1 - \beta)\%$ tolerance interval for $\lambda(t)$ is given by

$$\left[\tilde{\lambda}_{(\lceil K\beta/2 \rceil)}(t), \tilde{\lambda}_{(\lceil K\{1-\beta\}/2 \rceil)}(t) \right],$$

where $\lceil z \rceil$ denotes the smallest integer greater than or equal to z . For example if $K = 100$ and $\beta = 0.10$, then the estimated 90% tolerance interval for $\lambda(t)$ at a *single fixed* time $t \in [0, S]$ is $[\tilde{\lambda}_{(5)}(t), \tilde{\lambda}_{(95)}(t)]$. Similarly, tolerance intervals are obtained for the mean-value function $\mu(t)$ at a fixed time $t \in (0, S]$.

4.3 Presentation of Results

For each of the experimental cases specified in Subsection 4.1, Tables II and III contain a summary of the goodness-of-fit statistics formulated in Subsection 4.2. The statistics in Table II describe the errors in estimating the underlying theoretical rate and mean-value functions. The statistics in Table III describe the errors in fitting the $K = 100$ empirical mean-value functions. These tables also include entries of the form $\pm \widehat{\text{SE}}[\cdot]$ (that is, plus-or-minus the estimated standard error) for the following performance measures: $\bar{\delta}$, $\bar{\delta}^*$, $\bar{\Delta}$, $\bar{\Delta}^*$, $\overline{\text{SSE}}$, and $\overline{\text{MSE}}$. Thus for example, we see from Table II that in Case 3, $\bar{\Delta} = 12.9$ and $\widehat{\text{SE}}[\bar{\Delta}] = 0.76$. The estimated standard errors reported in Tables II and III provide some evidence that taking $K = 100$ replications of the OLS estimation procedure for each case yields reasonably stable estimates of the selected goodness-of-fit statistics for that case. Moreover, Table IV shows the frequency distribution of the fitted degree of the polynomial trend taken over all 100 replications for each case. Figures 1 through 6 display the graphs of 90% tolerance bands for the rate functions and mean-value functions associated with cases 1, 3, and 6.

TABLE II Goodness-of-Fit Statistics for Estimating $\lambda(t)$ and $\mu(t)$, $t \in [0, 12]$, with \pm the Estimated Standard Errors for Selected Statistics.

Case	$\mu(S)$	$\bar{\delta}$	V_{δ}	Q_{δ}	$\bar{\delta}^*$	V_{δ^*}	Q_{δ^*}	$\bar{\Delta}$	V_{Δ}	Q_{Δ}	$\bar{\Delta}^*$	V_{Δ^*}	Q_{Δ^*}
Base	586	29.8 ± 0.06	0.02	0.61	58.3 ± 0.29	0.05	1.19	13.7 ± 1.00	0.73	0.046	31.7 ± 2.09	0.66	0.110
0	586	10.0 ± 0.65	0.65	0.21	23.7 ± 1.42	0.60	0.48	12.4 ± 1.04	0.84	0.043	25.1 ± 1.86	0.74	0.087
1	588	11.4 ± 0.54	0.47	0.23	29.4 ± 1.38	0.47	0.60	12.8 ± 0.93	0.73	0.043	25.4 ± 1.63	0.64	0.086
2	1126	16.0 ± 0.58	0.36	0.17	65.5 ± 2.75	0.42	0.70	15.6 ± 1.09	0.70	0.033	37.4 ± 2.43	0.65	0.081
3	967	12.2 ± 0.33	0.27	0.15	80.3 ± 1.61	0.32	1.00	12.9 ± 0.76	0.59	0.038	33.3 ± 1.73	0.52	0.097
4	396	12.9 ± 0.89	0.69	0.39	57.5 ± 3.74	0.65	1.44	10.7 ± 0.58	0.54	0.058	26.6 ± 1.36	0.51	0.144
5	599	13.9 ± 0.64	0.46	0.28	40.9 ± 1.88	0.46	0.82	11.1 ± 0.90	0.81	0.037	23.5 ± 1.72	0.73	0.079
6	714	13.0 ± 0.42	0.32	0.22	52.3 ± 1.99	0.38	0.88	13.7 ± 0.88	0.64	0.037	29.2 ± 1.72	0.59	0.078

4.4 Analysis of Results

Comparing the results in Table II for Case 0 to the results for the Base Case, we see that in Case 0 the theoretical rate and mean-value functions are both estimated with greater accuracy. For example, in the Base Case the average maximum absolute errors $\bar{\delta}^*$ and $\bar{\Delta}^*$ in estimating the rate and mean-value functions are 58.3 ± 0.29 and 31.7 ± 2.09 , respectively; by contrast in Case 0 the corresponding statistics are 23.7 ± 1.42 and 25.1 ± 1.86 , respectively. These results provide some

TABLE III Goodness-of-Fit Statistics for Estimating $N(t)$, $t \in [0, 12]$, with \pm the Estimated Standard Errors for Selected Statistics.

Case	$\mu(S)$	\overline{SS}_E	V_{SS_E}	\overline{MS}_E	V_{MS_E}	\overline{D}	\overline{D}^*	Q_D	Q_{D^*}	H_D	H_{D^*}
Base	586	95.1 ± 0.57	0.060	0.162 ± 0.00097	0.060	7.94	24.6	0.026	0.082	0.0002	0.0007
0	586	36.6 ± 0.13	0.035	0.062 ± 0.00022	0.035	5.40	17.3	0.018	0.058	0.0001	0.0006
1	588	32.3 ± 0.21	0.066	0.055 ± 0.00036	0.065	5.37	17.2	0.018	0.057	0.0001	0.0006
2	1126	46.0 ± 0.02	0.004	0.041 ± 0.00004	0.009	6.11	20.8	0.013	0.045	0.0001	0.0004
3	967	26.9 ± 0.04	0.016	0.028 ± 0.00004	0.014	4.65	16.8	0.013	0.049	0.0001	0.0005
4	396	34.7 ± 1.39	0.402	0.088 ± 0.00355	0.403	4.92	15.5	0.026	0.083	0.0007	0.0019
5	599	33.8 ± 0.09	0.027	0.056 ± 0.00014	0.025	5.52	17.8	0.018	0.058	0.0002	0.0005
6	714	37.8 ± 0.04	0.011	0.053 ± 0.00062	0.116	5.96	19.9	0.016	0.052	0.0003	0.0007

indication of the potential problems associated with using a homogeneous Poisson process to model a time-dependent arrival process.

Whereas the performance measures in Table II that describe the estimation errors in fitting the underlying rate function are generally larger (worse) for the OLS estimation procedure than the corresponding results reported by Kuhl, Wilson, and Johnson (1997) for maximum likelihood estimation, the performance measures that describe the errors in fitting the underlying mean-value function are approximately the same for the two estimation methods. The results for Case 3 exemplify these phenomena, where we see that the statistics $\overline{\delta}$ and $\overline{\delta}^*$ for OLS estimation are 12.2 ± 0.33 and 80.3 ± 1.61 , respectively, according to Table II; and the corresponding statistics for maximum likelihood estimation are 6.3 and 45.4, respectively, according to Table 2 of Kuhl, Wilson, and Johnson (1997). On the other hand, the statistics $\overline{\Delta}$ and $\overline{\Delta}^*$ for OLS estimation in Case 3 are 12.9 ± 0.76 and 33.3 ± 1.73 , respectively; and the corresponding statistics for maximum likelihood estimation in Case 3 are 13.1 and 33.4, respectively. To put these goodness-of-fit statistics in perspective, we remark that on a personal computer with a 266 MHz Pentium II processor, the maximum likelihood estimation procedure `mp3m1e` (Kuhl and Wilson, 1996) required 380 minutes to fit the 100 replications of Case 3; by contrast, the OLS estimation procedure `mp31s` (Kuhl and Wilson, 1998) required 70 minutes to fit the same 100 data sets. For all the other experimental cases, we have observed similar speed-ups in using `mp31s` versus `mp3m1e`.

The larger rate-function estimation errors that were obtained with the OLS procedure may be

partially explained as follows. Essentially the objective function in (12) for the OLS procedure is based on the discrepancies between the square root of the fitted mean-value function and the square root of the empirical mean-value function; and there is no guarantee that the corresponding fitted rate function will closely approximate the underlying true rate function. The difference between the quality of the fits for the two estimation methods is also evident in the plots of the rate and mean-value functions.

The results in Table III provide some evidence that the OLS procedure generally yields an accurate estimate of the target empirical mean-value function in each data set to which the procedure is applied. For example, in Case 3 we see that $H_D = 0.0001$ and $Q_{D^*} = 0.0005$ so that on a “typical” replication of Case 3, the average and maximum absolute errors in estimating the empirical mean-value function are respectively 0.01% and 0.05% of the time-weighted average level of the empirical mean-value function computed over that replication of Case 3. Moreover, the residual mean square $\overline{\text{MS}}_E$ in Case 3 is 0.028 ± 0.00004 ; and this provides another perspective on the accuracy with which the empirical mean-value function is estimated on each realization of Case 3. More generally, the performance measures $\overline{\text{MS}}_E$, Q_D , Q_{D^*} , H_D , and H_{D^*} in Table III indicate that the OLS estimation procedure yields consistently accurate fits to the empirical mean-value function across all seven cases.

In Section 2 and Appendix B, we have shown that if the fitted mean-value function coincides exactly with the underlying mean-value function, then the expected value of $\overline{\text{MS}}_E$ should approach 0.25 as $n \rightarrow \infty$; however, the observed values of $\overline{\text{MS}}_E$ in Table III range from 0.028 to 0.088 across the seven cases. This may be an indication that the fitted mean-value function is tracking random variability associated with the empirical mean-value function and thus is deviating from the underlying mean-value function.

Table IV indicates the ability of the fitting procedure to determine the degree of the exponential-polynomial trend present in the underlying NHPP rate function. These results indicate that the approximate likelihood ratio test (13) works reasonably well in general for rate functions having up to $p = 4$ periodic components—provided at least a “moderate” sample size ($n \geq 500$) is available. For example, we experimented with a variant of Case 4 in which the intercept α_0 was increased from 3.6269 to 4.5197 so that the expected sample size increased from 396 to 968; and in this situation the observed frequencies of fitted polynomial trends of degree 0, 1, 2, 3, and 4 were 3, 0, 1, 95, and 1, respectively.

The plots of the 90% tolerance bands about the theoretical rate functions indicate that the OLS estimation procedure is consistently able to fit a reasonable EPTMP-type rate function to the underlying NHPP. Similar to the results reported by Kuhl, Wilson, and Johnson (1997) for maximum likelihood estimation, the plots of the tolerance bands for OLS estimation are widest at the peaks and valleys of the arrival rate. In addition, the tolerance bands tend to be wider as the number of periodic components increases.

The plots of the 90% tolerance bands about the theoretical mean-value functions also indicate

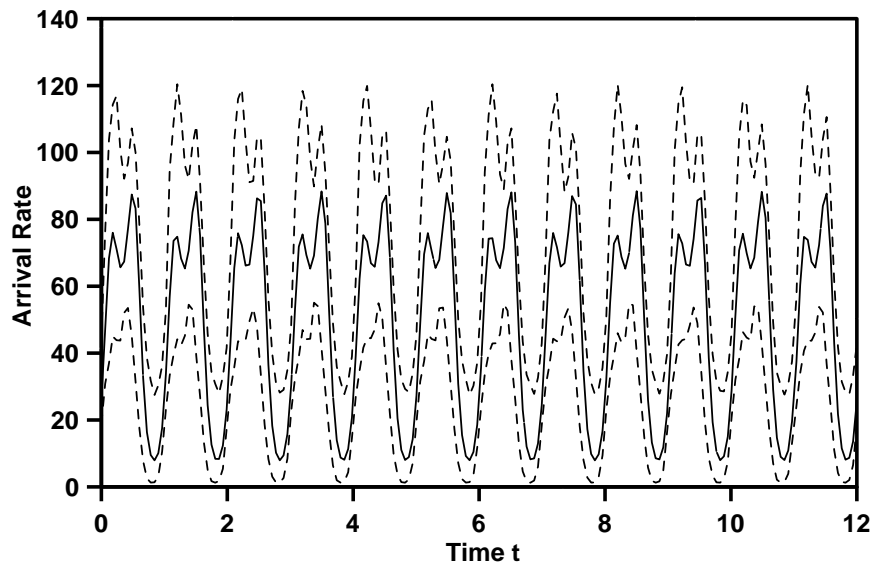


FIGURE 1 90% Tolerance Intervals for $\lambda(t)$, $t \in [0, 12]$, in Case 1.

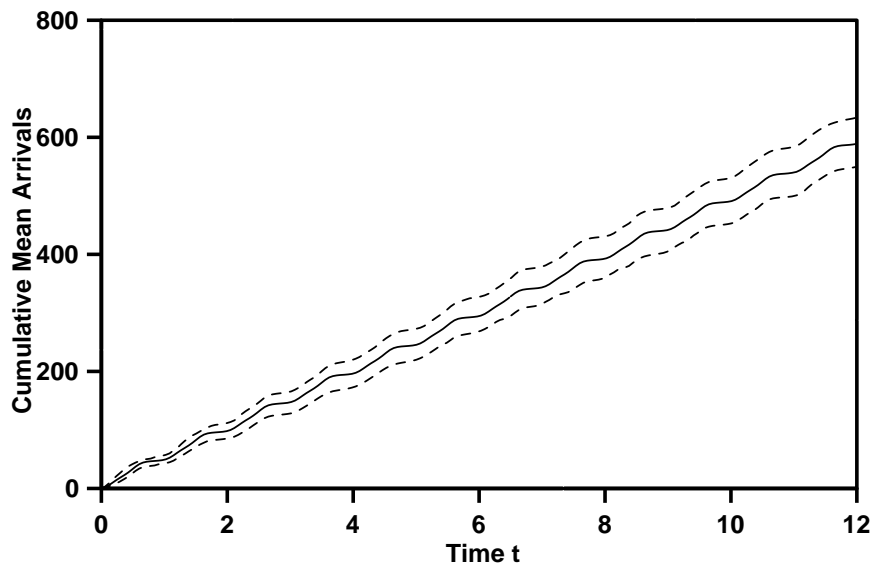


FIGURE 2 90% Tolerance Intervals for $\mu(t)$, $t \in [0, 12]$, in Case 1.

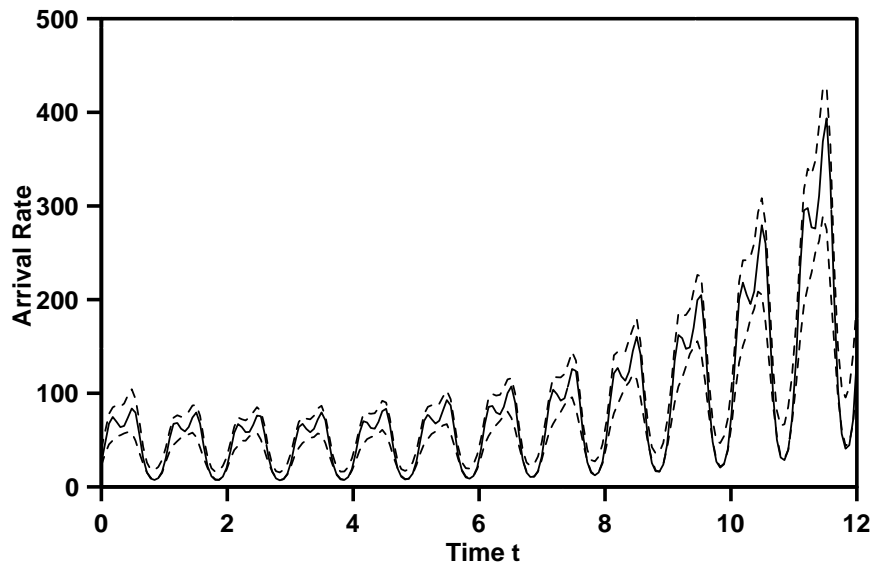


FIGURE 3 90% Tolerance Intervals for $\lambda(t)$, $t \in [0, 12]$, in Case 3.

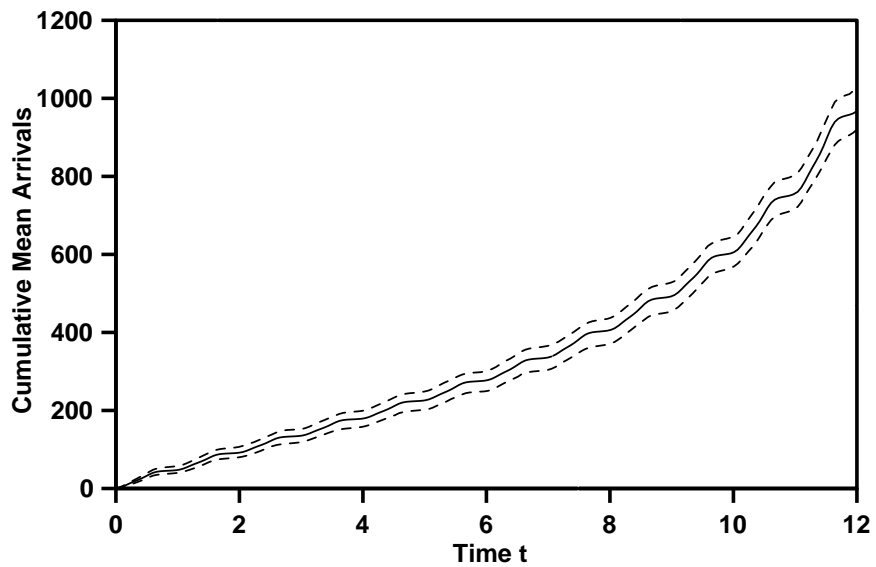


FIGURE 4 90% Tolerance Intervals for $\mu(t)$, $t \in [0, 12]$, in Case 3.

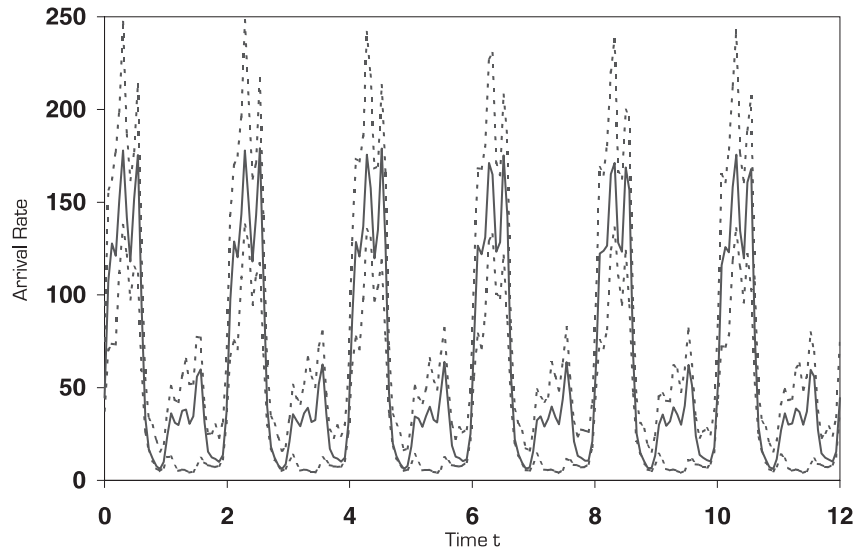


FIGURE 5 90% Tolerance Intervals for $\lambda(t)$, $t \in [0, 12]$, in Case 6.

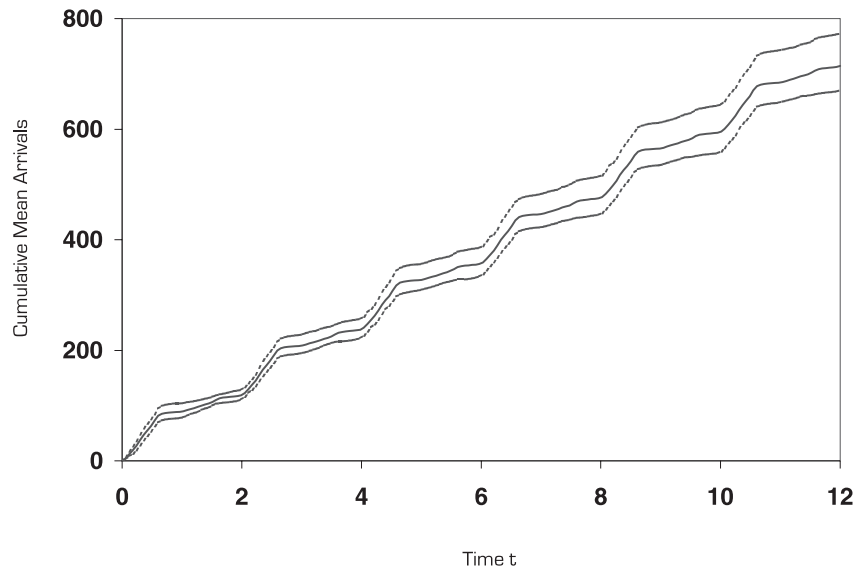


FIGURE 6 90% Tolerance Intervals for $\mu(t)$, $t \in [0, 12]$, in Case 6.

TABLE IV Frequency of Fitted Polynomial Degree for $K = 100$ Realizations.

Case	True Degree	Fitted Degree						
		0	1	2	3	4	5	6
0	0	93	7	0	0	0	0	0
1	0	87	13	0	0	0	0	0
2	1	0	94	6	0	0	0	0
3	2	0	0	87	13	0	0	0
4	3	26	10	14	48	2	0	0
5	0	98	2	0	0	0	0	0
6	0	99	1	0	0	0	0	0

that the OLS estimation procedure consistently provides reasonable estimates of the underlying NHPP. Also from the plots of the tolerance bands, we observe that the widths of the tolerance bands increase over time. This behavior is expected. Because the error is cumulative over time, the estimation error increases as the mean-value function increases.

Beyond the results displayed in Tables II through IV and in Figures 1 through 6, our practical experience in applying the OLS estimation procedure to time-dependent arrival processes exhibiting a “moderate” to “large” number of cyclic effects (that is, in processes with $4 < p \leq 20$) indicates the procedure’s robustness in handling increasingly complex estimation problems. By contrast, in many of these applications involving more than $p = 4$ cyclic effects, the maximum likelihood estimation procedure implemented in `mp3m1e` (Kuhl and Wilson 1998) simply failed to converge to a final answer.

5. CONCLUSIONS AND RECOMMENDATIONS

In this paper we developed and evaluated WLS and OLS methods for estimating the parameters of an NHPP having an EPTMP-type rate function. Anomalous performance of the WLS procedure is explained by an analysis of the associated residuals. The OLS estimation procedure is specifically designed to handle situations in which a large number of arrival processes must be fitted to historical records of “moderate” to “large” length (that is, sample sizes of at least 500). A central limit theorem for the OLS residuals together with an extensive experimental performance evaluation of the OLS estimation procedure suggest that this procedure can yield an adequate approximation to an arrival process with much less computational effort than would be required by maximum likelihood estimation, provided a sufficient volume of sample data is available. Moreover, several practical applications of the OLS estimation procedure to the development of stochastic input models for large-scale simulation experiments have demonstrated the robustness of the OLS estimation procedure in handling increasingly complex estimation problems.

Further improvement in the performance of the OLS estimation procedure may require the

development of an alternative to the approximate likelihood ratio test (13) for determining the degree of the polynomial trend component of the rate function. It would also be highly desirable to develop an automated procedure to determine the number of trigonometric rate components. These issues are the subject of ongoing work.

REFERENCES

- Abramowitz, M. and Stegun, I. A. (1972) *Handbook of Mathematical Functions*. New York: Dover Publications.
- Akaike, H. (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, **AC-19**, 716–723.
- Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978) *Statistics for Experimenters*, New York: John Wiley & Sons.
- Bratley, P., Fox, B. L., and Schrage, L. E. (1987) *A Guide to Simulation*, 2nd Edn., New York: Springer-Verlag.
- Cherry, T. M. (1946) Expansions in Terms of Parabolic Cylinder Functions. *Proceedings of the Edinburgh Mathematical Society*, **7**, 50–65.
- Çinlar, E. (1975) *Introduction to Stochastic Processes*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Davidon, W. C. (1959) Variable Metric Method for Minimization. AEC Research and Development Report ANL-5990 (revised).
- Dwight, H. B. (1961) *Tables of Integrals and Other Mathematical Data*, 2nd Edn., New York: Macmillan.
- Flanigan, M. A. (1993) A Flexible, Interactive, Graphical Approach to Modeling Stochastic Input Processes, *Ph.D. Dissertation*, School of Industrial Engineering, Purdue University, West Lafayette, Indiana, U.S.A.
- Fletcher, R., and Powell, M. 1963. A Rapidly Convergent Descent Method for Minimization. *Computer Journal*, **6**, 163–168.
- Härtler, G. (1989) The Nonhomogeneous Poisson Process—A Model for the Reliability of Complex Repairable Systems. *Microelectronics and Reliability*, **29**, 381–386.
- Hogg, R. V. and Craig, A. T. (1995) *Introduction to Mathematical Statistics*, 5th Edn., Upper Saddle River, New Jersey: Prentice-Hall.
- Kennedy, W. J., Jr., and Gentle, J. E. (1980) *Statistical Computing*, New York: Marcel Dekker.
- Kuhl, M. E. (1997) Modeling, Estimation and Simulation of Nonstationary Point Processes, *Ph.D. Dissertation*, Department of Industrial Engineering, North Carolina State University, Raleigh, North Carolina, U.S.A.
- Kuhl, M. E. and Wilson, J. R. (1996) User’s Manual for `mp3mle` and `mp3sim`: Software for Estimating and Simulating Nonhomogeneous Poisson Processes. Technical Report, Department of Industrial Engineering, North Carolina State University, Raleigh, North Carolina, U.S.A. Self-

- extracting executable files containing the software and documentation are available on the Web via [ftp://ftp.eos.ncsu.edu/pub/jwilson/mp3mlese.exe] and [ftp://ftp.eos.ncsu.edu/pub/jwilson/mp3simse.exe].
- Kuhl, M. E. and Wilson, J. R. (1998) User's Manual for **mp31s**: Software for Least Squares Estimation of Nonhomogeneous Poisson Processes. Technical Report, Department of Industrial Engineering, North Carolina State University, Raleigh, North Carolina, U.S.A. A self-extracting executable file containing the software and documentation is available via [ftp://ftp.eos.ncsu.edu/pub/jwilson/mp31s.exe] on the Web.
- Kuhl, M. E., Wilson, J. R., and Johnson, M. A. (1997) Estimating and Simulating Poisson Processes Having Trends or Multiple Periodicities. *IIE Transactions*, **29**, 201–211.
- Kuhl, M. E., Damerджи, H., and Wilson, J. R. (1998) Least-Squares Estimation of Nonhomogeneous Poisson Processes. *Proceedings of the 1998 Winter Simulation Conference* (ed. D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan), Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, U.S.A., 637–645.
- Lee, S., Wilson, J. R., and Crawford, M. M. (1991) Modeling and Simulation of a Nonhomogeneous Poisson Process Having Cyclic Behavior. *Communications in Statistics—Simulation and Computation*, **20**, 777–809.
- Lewis, P. A. W. (1970) Remarks on the Theory, Computation and Application of the Spectral Analysis of Series of Events. *Journal of Sound and Vibration*, **12**, 353–375.
- Nash, S. G. (1984) Newton-Type Minimizations via the Lanczos Method. *Siam Journal on Numerical Analysis*, **21**, 770–788.
- Nelder, J. A. and Mead, R. (1964) A Simplex Method for Function Minimization. *Computer Journal*, **7**, 308–313.
- Olsson, D. M. (1974) A Sequential Simplex Program for Solving Minimization Problems. *Journal of Quality Technology*, **6**, 53–57.
- Olsson, D. M., and Nelson, L. S. (1975) The Nelder-Mead Simplex Procedure for Function Minimization. *Technometrics*, **17**, 45–51.
- Powell, M. J. D. (1964) An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives. *Computer Journal*, **7**, 155–162.
- Pritsker, A. A. B. (1998) Life & Death Decisions: Organ Transplantation Allocation Policy Analysis. *OR/MS Today*, **25**, 22–28.
- Pritsker, A. A. B., Martin, D. L., Reust, J. S., Wagner, M. A., Wilson, J. R., Kuhl, M. E., Allen, M. D., Daily, O. P., Harper, A. M., Edwards, E. B., Bennett, L. E., Roberts, J. P., and Burdick, J. F. (1995) Organ Transplantation Policy Evaluation. *Proceedings of the 1995 Winter Simulation Conference* (ed. C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman), Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, U.S.A., 1314–1323.
- Seber, G. A. F., and Wild, C. J. (1989) *Nonlinear Regression*, New York: John Wiley & Sons.
- Swain, J. J., Venkatraman, S., and Wilson, J. R. (1988) Least Squares Estimation of Distribution

Functions in Johnson's Translation System. *Journal of Statistical Computation and Simulation*, **29**, 271–297.

Venables, W. N., and Ripley, B. D. (1994) *Modern Applied Statistics with S-Plus*, New York: Springer-Verlag.

Wagner, M. A. F., and Wilson, J. R. (1996) Using Univariate Bézier Distributions to Model Simulation Input Processes. *IIE Transactions*, **28**, 699–712.

APPENDIX A: PROPERTIES OF RESIDUALS ARISING FROM WLS ESTIMATION OF NHPPs

If the sample size n is fixed in the WLS procedure of Subsection 2.2, then it follows from (8) that the i th transformed residual for $i = 1, 2, \dots, n - 1$ is

$$u_i(\hat{\Theta}) = \sqrt{\frac{i+1}{i}} \left\{ \mu(\tau_i; \hat{\Theta}) - E[\mu(\tau_i; \hat{\Theta})] \right\} - \sqrt{\frac{i}{i+1}} \left\{ \mu(\tau_{i+1}; \hat{\Theta}) - E[\mu(\tau_{i+1}; \hat{\Theta})] \right\} \quad (\text{A.1})$$

$$\approx \sqrt{\frac{i+1}{i}} \left\{ \mu(\tau_i; \hat{\Theta}) - N(\tau_i) \right\} - \sqrt{\frac{i}{i+1}} \left\{ \mu(\tau_{i+1}; \hat{\Theta}) - N(\tau_{i+1}) \right\} \quad (\text{A.2})$$

$$= \mu(\tau_i; \hat{\Theta}) \sqrt{\frac{i+1}{i}} - \mu(\tau_{i+1}; \hat{\Theta}) \sqrt{\frac{i}{i+1}}, \quad (\text{A.3})$$

since in (A.1), the expectation $E[\mu(\tau_i; \hat{\Theta})]$ is unknown and in practice we take $E[\mu(\tau_i; \hat{\Theta})] \approx E[\mu(\tau_i; \Theta)] = i = N(\tau_i)$ to obtain the i th approximated residual (A.2); and similarly the last transformed residual is

$$u_n(\hat{\Theta}) = \frac{1}{\sqrt{n}} \left\{ \mu(\tau_n; \hat{\Theta}) - E[\mu(\tau_n; \hat{\Theta})] \right\} \approx \frac{1}{\sqrt{n}} \left[\mu(\tau_n; \hat{\Theta}) - N(\tau_n) \right]. \quad (\text{A.4})$$

Notice the exact cancellation occurring in (A.3):

$$\sqrt{\frac{i+1}{i}} \left\{ -N(\tau_i) \right\} - \sqrt{\frac{i}{i+1}} \left\{ -N(\tau_{i+1}) \right\} = -\sqrt{i(i+1)} + \sqrt{i(i+1)} = 0,$$

so that in practice the information about the discrepancy between the empirical mean-value function $N(t)$ and the fitted mean-value function $\mu(t; \hat{\Theta})$ has been completely eliminated from all but the last element of $\mathbf{u}(\hat{\Theta})$; and thus the resulting WLS estimator (9) of the parameter vector Θ cannot be expected to yield acceptable fits to historical data.

In the rest of this appendix, we establish the asymptotic mean and variance of the ideal WLS residuals

$$u_i(\Theta) = \begin{cases} \sqrt{\frac{i+1}{i}} \tau_i^* - \sqrt{\frac{i}{i+1}} \tau_{i+1}^*, & i = 1, \dots, n - 1, \\ \frac{1}{\sqrt{n}} (\tau_n^* - n), & i = n, \end{cases} \quad (\text{A.5})$$

where for each positive integer j , the random variable $\tau_j^* = \sum_{\ell=1}^j X_\ell^*$; and the $\{X_\ell^*\}$ are i.i.d. exponential variates with mean 1.

Definition 1 A standard reversed exponential variate W has p.d.f.

$$f_W(w) = \begin{cases} e^{-(1-w)}, & w \leq 1, \\ 0, & w > 1, \end{cases}$$

so that W has mean 0, variance 1, and moment generating function $M_W(z) \equiv \mathbb{E}[e^{zW}] = e^z/(1+z)$ for all $z > -1$.

Proposition 1 The ideal WLS residuals in (A.5) converge in distribution to a standard reversed exponential variate W as $i \rightarrow \infty$ provided $i \neq n$; and the last ideal WLS residual in (A.5) converges in distribution to a standard normal variate as $n \rightarrow \infty$ so that

$$u_i(\Theta) \underset{i \neq n}{\xrightarrow{i \rightarrow \infty} \mathcal{D}} W \quad \text{and} \quad u_n(\Theta) \xrightarrow{n \rightarrow \infty} \mathcal{D} N(0, 1). \quad (\text{A.6})$$

Proof. The second part of (A.6) follows immediately from (A.5) for the case $i = n$ together with the central limit theorem. For the case that $i \neq n$ in (A.5), direct calculation of the moment generating function $M_{u_i(\Theta)}(z) \equiv \mathbb{E}\{\exp[zu_i(\Theta)]\}$ shows that $\ln[M_{u_i(\Theta)}(z)] = -i \ln\{1 - z/[i(i+1)]^{1/2}\} - \ln\{1 + z[i/(i+1)]^{1/2}\}$ for all $z \in (-1, 1)$; and combining this result with the expansion $\ln(1-y) = -y + O(y^2)$ for all y with $|y| < 1$ (Dwight 1961, Equation 601.1), we see that $\lim_{i \rightarrow \infty, i \neq n} M_{u_i(\Theta)}(z) = e^z/(1+z) = M_W(z)$ for all $z \in (-1, 1)$. Thus the first part of (A.6) follows immediately from the continuity theorem for moment generating functions (Hogg and Craig, 1995, Section 5.3).

APPENDIX B: PROPERTIES OF THE SQUARE-ROOT-TRANSFORMED OLS PROCEDURE

In this appendix we establish the asymptotic mean and variance as well as the asymptotic distribution of the ideal OLS residuals (10).

Proposition 2 The ideal OLS residuals in (10) satisfy

$$\lim_{i \rightarrow \infty} \mathbb{E}[\eta_i(\Theta)] = 0 \quad \text{and} \quad \lim_{i \rightarrow \infty} \text{Var}[\eta_i(\Theta)] = \frac{1}{4}. \quad (\text{B.1})$$

Proof. For $i = 1, 2, \dots$, we have

$$\mathbb{E}[\eta_i(\Theta)] = \int_0^\infty \sqrt{y} \frac{1}{\Gamma(i)} y^{i-1} e^{-y} dy - \sqrt{i - \frac{1}{4}} = \sqrt{i - \frac{1}{4}} \left[\frac{\Gamma(i + \frac{1}{2})}{\Gamma(i)\sqrt{i}} \cdot \frac{1}{\sqrt{1 - 1/(4i)}} - 1 \right]. \quad (\text{B.2})$$

Next we show that

$$Q(x) \equiv \frac{\Gamma(x + \frac{1}{2})}{\Gamma(x)\sqrt{x}} \quad \text{for all } x > 0 \quad \text{has derivative } Q'(x) > 0 \quad \text{for all sufficiently large } x. \quad (\text{B.3})$$

Observe that $Q'(x) = Q(x) [\psi(x + \frac{1}{2}) - \psi(x) - 1/(2x)]$, where $\psi(x) \equiv \Gamma'(x)/\Gamma(x)$ is the psi (digamma) function. Equation (6.3.18) of Abramowitz and Stegun (1972) yields the asymptotic expansion $\psi(x) = \ln(x) - 1/(2x) - 1/(12x^2) + O(x^{-4})$ as $x \rightarrow \infty$; and combining this with the expansion $\ln(1 + y) = y - \frac{1}{2}y^2 + O(y^3)$ for all y with $|y| < 1$ (Dwight 1961, Equation 601), we obtain the following asymptotic expansion for $Q'(x)$:

$$Q'(x) = \frac{2x - 1}{16x^3 + 8x^2} + O(x^{-3}) > 0 \text{ as } x \rightarrow \infty. \quad (\text{B.4})$$

Using (B.4) and Stirling's approximation, we see that for sufficiently large x , the function $Q(x)$ increases monotonically to the limit

$$\lim_{x \rightarrow \infty} Q(x) = \lim_{x \rightarrow \infty} \frac{e^{-x-1/2}(x + \frac{1}{2})^x}{e^{-x}x^x} = \lim_{x \rightarrow \infty} \frac{(1 + \frac{1}{2x})^x}{\sqrt{e}} = 1. \quad (\text{B.5})$$

It follows from (B.2) through (B.5) that as $i \rightarrow \infty$,

$$|E[\eta_i(\Theta)]| = \left| \sqrt{i - \frac{1}{4}} \left[\frac{Q(i)}{\sqrt{1 - 1/(4i)}} - 1 \right] \right| \leq \left| \sqrt{i - \frac{1}{4}} \left[\frac{1}{\sqrt{1 - 1/(4i)}} - 1 \right] \right| = \frac{1}{2} (\sqrt{4i} - \sqrt{4i - 1}) \rightarrow 0.$$

The variance of $\eta_i(\Theta)$ is given by

$$\text{Var}[\eta_i(\Theta)] = i - \frac{\Gamma^2(i + \frac{1}{2})}{\Gamma^2(i)} = i[1 - Q^2(i)] \text{ for } i = 1, 2, \dots \quad (\text{B.6})$$

Applying L'Hospital's rule to (B.6), and using (B.4) and (B.5) to simplify the resulting asymptotic expression, we obtain

$$\lim_{i \rightarrow \infty} \text{Var}[\eta_i(\Theta)] = \lim_{x \rightarrow \infty} \frac{2Q(x)Q'(x)}{x^{-2}} = \lim_{x \rightarrow \infty} \left[\frac{4x^3 - 2x^2}{16x^3 + 8x^2} + O(x^{-1}) \right] = \frac{1}{4}.$$

Proposition 3 *The ideal OLS residuals in (10) converge in distribution to a normal variate with mean 0 and variance $\frac{1}{4}$ as $i \rightarrow \infty$,*

$$\eta_i(\Theta) \xrightarrow[i \rightarrow \infty]{\mathcal{D}} N(0, \frac{1}{4}). \quad (\text{B.7})$$

Proof. From (10) we see that the moment generating function of $\eta_i(\Theta)$ is

$$M_{\eta_i(\Theta)}(z) \equiv E\{\exp[z\eta_i(\Theta)]\} = \exp\left(-z\sqrt{i - \frac{1}{4}}\right) \int_0^\infty \frac{\exp(z\sqrt{y})y^{i-1}e^{-y}dy}{\Gamma(i)} \quad (\text{B.8})$$

for $-\infty < z < \infty$. Making the change of variables $\frac{1}{2}x^2 = y$ in (B.8), we see that

$$M_{\eta_i(\Theta)}(z) = \frac{\exp\left(-z\sqrt{i - \frac{1}{4}}\right) \Gamma(2i)}{2^{i-1} \Gamma(i)} \int_0^\infty \frac{\exp[-(x^2/2 - zx/\sqrt{2})] x^{2i-1}}{\Gamma(2i)} dx \quad (\text{B.9})$$

for $-\infty < z < \infty$. Display (1) of Cherry (1946) establishes that for $-\infty < a < \infty$, the parabolic cylinder function satisfies

$$D_{-\nu-1}(a) \equiv \frac{\exp\left(-\frac{1}{4}a^2\right)}{\Gamma(\nu+1)} \int_0^\infty \exp\left[-(x^2/2 + ax)\right] x^\nu dx \sim \frac{\exp\left(\frac{1}{2}\nu - a\sqrt{\nu}\right)}{\sqrt{2\nu}^{(\nu+1)/2}} \text{ as } \nu \rightarrow \infty. \quad (\text{B.10})$$

If we take $a \equiv -z/\sqrt{2}$ and $\nu \equiv 2i - 1$ in (B.10), then we see that (B.9) has the form

$$M_{\eta_i(\Theta)}(z) = \left[\frac{\exp\left(-z\sqrt{i-\frac{1}{4}}\right)}{2^{i-1}} \cdot \frac{\Gamma(2i)}{\Gamma(i)} \cdot D_{-2i}(-z/\sqrt{2}) \right] \exp\left(\frac{1}{8}z^2\right) \text{ for } -\infty < z < \infty. \quad (\text{B.11})$$

Letting $\omega_i(z)$ denote the term in square brackets on the right-hand side of (B.11), we show that $\lim_{i \rightarrow \infty} \omega_i(z) = 1$ for all real z . It follows from (B.10) that for each $z \in (-\infty, \infty)$, we have

$$\omega_i(z) \sim \frac{\exp\left(-z\sqrt{i-\frac{1}{4}}\right)}{2^{i-1}} \cdot \frac{\Gamma(2i)}{\Gamma(i)} \cdot \frac{\exp\left(i-\frac{1}{2}+z\sqrt{i-\frac{1}{2}}\right)}{\sqrt{2}(2i-1)^i} \text{ as } i \rightarrow \infty. \quad (\text{B.12})$$

Applying Stirling's Formula to (B.12) and rearranging terms, we see that for each $z \in (-\infty, \infty)$,

$$\omega_i(z) \sim \left[\frac{\exp\left(z\sqrt{i-\frac{1}{2}}\right)}{\exp\left(z\sqrt{i-\frac{1}{4}}\right)} \right] \left[\frac{(2i)^{2i-1/2}}{(2i)^{i-1/2}(2i-1)^i} \right] \left[\frac{\exp\left(i-\frac{1}{2}\right)}{e^i} \right] \text{ as } i \rightarrow \infty. \quad (\text{B.13})$$

We have

$$\lim_{i \rightarrow \infty} \frac{\exp\left(z\sqrt{i-\frac{1}{2}}\right)}{\exp\left(z\sqrt{i-\frac{1}{4}}\right)} = \lim_{i \rightarrow \infty} \exp\left[z\left(\sqrt{i-\frac{1}{2}} - \sqrt{i-\frac{1}{4}}\right)\right] = 1 \text{ for } -\infty < z < \infty, \quad (\text{B.14})$$

and

$$\lim_{i \rightarrow \infty} \left[\frac{(2i)^{2i-1/2}}{(2i)^{i-1/2}(2i-1)^i} \right] \left[\frac{\exp\left(i-\frac{1}{2}\right)}{e^i} \right] = \lim_{i \rightarrow \infty} \frac{e^{-1/2}}{\left(1-\frac{1}{2i}\right)^i} = 1. \quad (\text{B.15})$$

It follows from (B.11) and (B.13) through (B.15) that $\lim_{i \rightarrow \infty} M_{\eta_i(\Theta)}(z) = \exp\left[\frac{1}{2}\left(z^2/4\right)\right]$ for $-\infty < z < \infty$; thus the conclusion (B.7) follows from the continuity theorem for moment generating functions (Hogg and Craig, 1995).