Journal of the Operations Research Society of Japan
Vol. 42, No. 4, December 1999

PIECEWISE PSEUDO ESTIMATION FOR NHPP USING DATA IN THE ADJACENT INTERVALS

Kenichi Funaki Hideaki Matoba
Hitachi Ltd.

(Received October 22, 1998; Revised May 10, 1999)

Abstract In modeling a system with NHPP, we generally need to estimate the parameters of its mean function $\Lambda(t)$. In practical situations, however, there exist many cases that the parametric representation of $\Lambda(t)$ is unknown. For such a case, Law and Kelton proposed a nonparametric estimation with a piecewise-constant intensity function, which requires many sample data sets to obtain statistically accurate values. In this paper we propose a piecewise estimation method which covers the case that only one or a few sets of sample data can be obtained. The proposed method gives pseudo estimator by using the sample data of the adjacent intervals, nevertheless it can theoretically gain more statistical accuracy even with one or a few sets of sample data than the method by Law and Kelton. The effectiveness of the proposed method is shown by some experimental results.

1. Introduction

Poisson process is one of the models widely used to describe systems with stochastic events. Many actual systems such as patient arrivals at a hospital, transactions in a database system and occurrence of typhoons have been successfully modeled with Poisson process so far.

The assumptions required for Poisson process are; (1) it is a counting process starting with 0 events at time 0, (2) it has independent increments and (3) the number of events during any time interval of length τ is Poisson distributed with a mean dependent on τ. Especially in terms of (3), if the number of events is Poisson distributed with mean $\lambda\tau$ where λ is a positive real constant, the process has stationary independent increments and therefore is called homogeneous (or stationary) Poisson process (HPP). On the other hand, if the Poisson distribution of the number of events has a mean dependent on both the time point t and the length of the interval τ, it is called nonhomogeneous (or nonstationary) Poisson process (NHPP). Since, in general, occurrence of events depends on time, for example, peak-hour of customer arrivals at a restaurant during a day, NHPP is more applicable to many real cases.

In modeling a system with NHPP, we need to estimate the parameters which characterize the process. Since NHPP is specified by the mean function $\Lambda(t)$, estimation for NHPP is equivalent to identifying the parameters of $\Lambda(t)$. When the parametric representation of $\Lambda(t)$ is known beforehand, the ordinary way of estimation is to maximize likelihood of $\Lambda(t)$ using the collected sample data. The sample data can be collected as either time points of events or number of events in each successive time interval. Maximum likelihood estimation is applicable to whichever type of sample data; Bar-David [3] gave basic results on the likelihood function of $\Lambda(t)$.

In practical situations, however, there exist many cases that the parametric representation of $\Lambda(t)$ is unknown for lack of knowledge about the determinant such as cyclic factors of the event occurrence. For such a case, Law and Kelton (section 6.10.2 in [18]) proposed a

nonparametric estimation with a piecewise-constant intensity function. In their method, dividing the time axis into nonoverlapping time intervals where the intensity function is assumed to be fairly constant, they estimated a single intensity for each interval by calculating the average from many sample data. Although this method is very simple to implement, it still requires many sets of sample data for each interval in order to obtain a statistically reasonable average.

In this paper we propose a piecewise estimation method which is enhanced from that of Law and Kelton [18] in order to cover the case that only one or a few sets of sample data can be obtained. The proposed method gives pseudo estimator by using the sample data of the adjacent intervals. Although it is a pseudo estimation method, we show that it can theoretically gain more statistical accuracy even with one or a few sets of sample data than the method by Law and Kelton. Furthermore we show the effectiveness of the proposed method by some experimental results.

The rest of the paper is organized as follows. Section 2 reviews the related previous works. Section 3 describes the actual case which motivated this study in order to show the actuality of the problem. In section 4 proposed is the pseudo estimator and its effectiveness condition is given. Section 5 describes the pseudo estimation algorithm and its experimental results using the condition given in the section 4. In section 6, we point and clarify the difference between the proposed estimator and the local averaging smoother for more global understanding. Finally section 7 closes the paper with some concluding remarks.

2. Previous Works

Since NHPP is applicable to many systems which include stochastic process with time-varying intensity, many estimation methods have been proposed to date. The oldest and the most basic result in estimation theory for NHPP was given by Bar-David [3], in which the likelihood function of time-varying intensity function is derived in a general form. All the previous works related to the maximum likelihood estimator for the intensity function were done by course of the Bar-David's likelihood function. In performing the maximum likelihood estimation, we need to know the parametric representation of the intensity function beforehand. However there exist many cases that the precise representation of the intensity function is unknown.

For the estimation of the parameters of an unknown intensity function, approximately representative functions have been used being based on knowledge of the cyclic and trend characteristics of the process in question. Lee, Wilson and Crawford [19] used an exponential-polynomial-trigonometric intensity function to model the storm arrival process at an off-shore drilling site, which shows some seasonal cyclic behavior. The exponential-polynomial-trigonometric function is in a very general form and therefore it can model most of the processes with both cyclic behavior and trend. However, it requires very complex calculation to do a statistical treatment because of inclusion of the trigonometric terms. Lewis and Shedler [21] modeled the transactions raised in a database system with daily and weekly cyclic patterns by an exponential-polynomial intensity function. Although the exponential-polynomial function has in itself a tractable structure, if the process is composed of various types of cyclic variations, many high order terms are required to be involved in the function. Kao and Chang [15] studied a method with a piecewise polynomial function to estimate the discontinuous intensity function appeared in the observations of the calls for on-line analysis of electrocardiograms at a hospital, which includes daily and weekly cyclic patterns.

When the process shows some monotonous trend in its event occurrence, Weibull process or power law process is a typical model to be used. Since the minimal repair model of a repairable system in reliability theory has a Weibull process, a lot of literature in reliability

theory studied the estimation of the parameters of the Weibull process (Bain and Engelhardt [2], Jang and Bai [14], Kvaløy and Lindqvist [17], Rigdon and Basu [22][23]). Monotonous trend can be also represented by polynomials. Klein and Roberts [16] modeled the intensity of lunch wagon arrivals at a parking lot by using a combination of three polynomials.

On the other hand, when we do not know any characteristics involved in the process in question, we need to use some nonparametric method. In the forementioned study by Lewis and Shedler [21], they also used a kernel-type intensity function estimation alternatively. However it requires two very restricted functions; a bounded, non-negative and integrable weight function and a positive bandwidth function, so that it is only applicable to very limited cases. Leemis [20] proposed a piecewise linear estimator of the cumulative intensity function. This method seems to be effective even though there is no knowledge on the parametric representation of the intensity function, but it is applicable only when the sample data are collected as time points of event occurrences. When we can only collect sample data by number of events in each time interval, piecewise constant estimator of the intensity function proposed by Law and Kelton [18] is valid. In their method, dividing the time axis into nonoverlapping time intervals, the estimator for each interval is calculated by taking the average rate of the sample number of realized events during the corresponding interval. Although the simple average of the number of events is a sufficient estimator of the mean of the corresponding Poisson distribution, many sample data sets are necessary to obtain statistical accuracy of the estimated value.

If we do not employ the Poisson basis, the estimation of the intensity function can be treated as nonparametric regression of the response function on the univariate t. The typical approaches in the nonparametric regression for scatterplots include local averaging and piecewise polynomials smoothing. The underlying assumption in these approaches is that the response on t should not change much when t does not change much.

The local averaging is a smoothing method which produces the fit at a time point t_i using the average of the data from the neighborhood of t_i. A straightforward method of the local averaging is the running mean. Friedman and Tibshirani [12] gave a monotone smoother based on the running mean which is calculated from the data in the symmetric nearest neighborhood. Friedman and Stuetzle [11] proposed more flexible running mean using the adaptable-span of the neighborhood. As a derivative of the running mean, Hastie and Tibshirani [13] presented the locally-weighted running mean which adopted the weighted average based on the distance of the neighborhood point from t_i. The other method of the local averaging is the kernel smoother. The kernel smoothing is to produce the fit at t_i by calculating the weighted average using the kernel function over a certain span around t_i. Cleveland [6] gave a kernel smoother with the length of the span inversely proportional to the local density of t. Silverman [24] pointed that the spline smoothing is approximately equivalent to the kernel smoothing. For more in-depth review of the local averaging, see Buja, Hastie and Tibshirani [5].

In the local averaging, the selection of the span taken in the average calculation is the key to get better performance. One of the explicit and statistic-based method to select the span is the cross-validation measure (CV). CV is a nearly unbiased estimate of the error rate [8], then the span should be selected so that the corresponding CV is minimized. Although CV is a powerful measure due to its simplicity, it is effective only if the error is supposed to be homoscedastic everywhere.

The piecewise polynomials smoothing is to replace the single response function on t with several polynomials each of which is defined over a subinterval of t axis. The technical significance in the piecewise polynomials smoothing resides in the location of the interval-dividing points on the t axis, which is called knots. Friedman and Silverman [10] proposed

the piecewise polynomials smoothing with the determination of the knots location by the generalized cross-validation measure (GCV) criterion. Agarwal and Studden [1] determined the number of knots based on the degree of the polynomial spline function used in the smoothing.

Both the local averaging and the piecewise polynomials smoothing are more general and flexible estimation in the sense that they do not stand on the Poisson basis. However, it can be also said that they have to search the wider estimation domain due to ignoring the knowledge from the Poisson basis. The proposed estimator in this paper is completely based on the Poisson basis, but it seems to be similar in its procedure to the local averaging. The different points between them are discussed in section 6.

3. Description of the Actual Case Motivating This Study

In performing simulation study of a system including some stochastic elements, it is likely to be obliged to do it with only one or a few sets of sample data provided to prepare the stochastic model. Likewise in our simulation project, which motivated this study, we could obtain only one sample data set of number of orders received every day for a computer product as shown in Figure 1. Figure 1 shows the number of daily orders received during 96 workdays period (exclusive of weekends and holidays), from which we had to generate random numbers for the simulation. As seen in Figure 1, we could not recognize either cyclic or trend characteristics in the process and could not deduce its parametric representation. Nevertheless the potential characteristics of the order reception activity warrants its Poisson assumptions; an order comes indifferently to the past reception of orders and multiple orders do not come at a same time. This characteristics warrants the independent increments and the Poisson distributed number of orders. Therefore it remains to test its homogeneity to verify that it constructs NHPP. One of the most robust and nonparametric test method for homogeneity of random numbers is the runs test (section 7.4.1 in Bendat and Piersol [4]), and the other method but specific to the Poisson variables is the combination of square root transformation with variance test (p.44 in Cox and Lewis [7]).

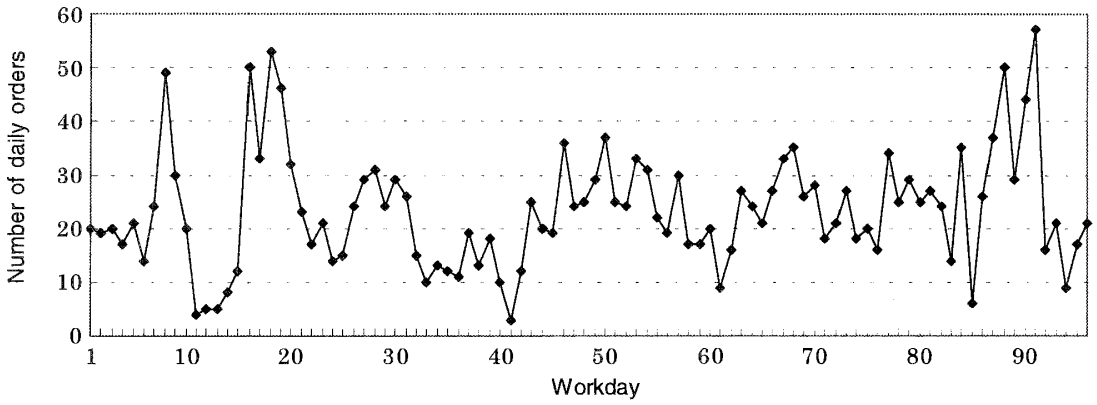

Figure 1 Number of daily orders for a computer product

3.1 Homogeneity test by runs

Here we define a run as a set of consecutive days whose numbers of orders are all consistently larger (or consistently smaller) than the median. For example, since the median over the 96 days in Figure 1 is 21.5, the first 6 days have all smaller number of orders than the median and then they form a run with its length 6. If the numbers of daily orders shown in Figure 1 keep the homogeneity over a certain period, the number of runs within the period falls into the range

given by the distribution of the number-of-runs correspondingly to a given confidence level. The table of the values related to the distribution of the number-of-runs can be found in, for example, Swed and Eisenhart [25]. We observe 25 runs all over the 96 days in Figure 1, which is out of the range [36,62] for the 99% upper- and lower-side confidence level, therefore it cannot be said statistically that it is homogeneous over the 96 days.

Further we performed the runs test for the period of the day 21 through 82 within which numbers of daily orders seem to be rather stationarily distributed. The median of the 62 days period is 23.5 and 16 runs are observed in that period. Since the homogeneity-confident range of the number of runs is [21,43] for the 99% upper- and lower-side confidence level, again we cannot say statistically that it is homogeneous over the 62 days.

3.2 Homogeneity test by square root transformation and variance test

The runs test concludes that there is no statistical evidence of homogeneity in the data of Figure 1. Here we can further perform the other test method postulating Poisson basis to make sure of its nonhomogeneity. The square root transformation given by Cox and Lewis [7] is to transform any Poisson variable into a random variable with a constant variance. Since this gives constant residual variance indifferently to the variance of the original Poisson variable, it is valid in doing statistical treatment based on the assemblies of possibly various Poisson variables. According to Cox and Lewis, the square root of a Poisson variable $N(t)$ plus a fourth, namely $\sqrt{n+1/4}$, has mean approximately equal to $\sqrt{\mu}$ and residual variance of 1/4 if μ is larger than 1, where μ is the mean of the original Poisson variable.

The square root transformation was originally introduced to perform a regression analysis of Poisson distributed values. Therefore the variance of the transformed variable is a residual one associated with approximately normally distributed error. This fact yields the validity to perform a variance test for the transformed variable, which requires normality assumption.

The null hypothesis here is that the transformed sample data have the variance of 1/4. The square-sum of sample deviations for the 96 days in Figure 1 is 124.06, and its ratio to the population variance 1/4 is 496.24. This value is far larger than the upper 99% point of 129.20 given by χ^2 distribution with the degree of freedom 95. Therefore the null hypothesis is statistically rejected, and consequently it means that not all the original Poisson variables have an identical mean. Hence it can be said that the sample data in Figure 1 do not show any legitimate homogeneity. For the period of the day 21 through 82, the square-sum of sample deviations and its ratio to 1/4 are 43.10 and 172.40 respectively, while the upper 99% point given by χ^2 distribution with the degree of freedom 61 is 88.80. Again it can be said that the sample data during the 62 days do not show any legitimate homogeneity.

4. Piecewise Pseudo Estimation Using Adjacent Interval Data

Using the sample data set of Figure 1, one easy but likely-to-be-biased way to generate NHPP random numbers is to adopt each observed number directly as the mean of the Poisson distribution for the corresponding day. However, this method is too much dependent on the only one set of the collected sample data to prevent itself from being a special case. The purpose of the proposed method in this paper is to moderate this kind of specialization to some extent.

4.1 Preliminaries

Given one or a few sets of NHPP sample data as number of realized events in each nonoverlapping successive time interval, we consider the estimation of expectation of the number of events in a target interval. We use the following definitions and notation;

Definitions: (see Figure 2)

sample data set	: a series of numbers of the realized events for every time interval over the time axis
unit interval	: intervals in each of which the number of the realized events are collected
estimation interval	: the target interval of which the estimation is performed
reference interval	: an interval including the estimation interval with its adjacent intervals
supplement interval	: intervals to be added to the estimation interval in order to form a reference interval

Notation:

I	: number of unit intervals where the sample data are collected
J	: number of the collected sample data sets
n_{ij}	: number of events observed in the i th unit interval of the j th sample data set ($i = 1,2,3,\cdots,I$, $j = 1,2,3,\cdots,J$)
K	: number of the estimation intervals
H_{Ek}	: true expectation of the number of events in the k th estimation interval ($k = 1,2,3,\cdots,K$)
H_{Rk}	: true expectation of the number of events in the k th reference interval
X_{Ek}	: sample mean of the number of realized events in the k th estimation interval
X_{Rk}	: sample mean of the number of realized events in the k th reference interval
u	: length of a unit interval
ts	: estimation starting point (a multiple of u)
τ_{Ek}	: length of the k th estimation interval (a multiple of u)
τ_{Rk}	: length of the k th reference interval (a multiple of u)
τ_{Sk}	: length of the k th supplement interval (a multiple of u)
Y_k	: pseudo estimator of H_{Ek}

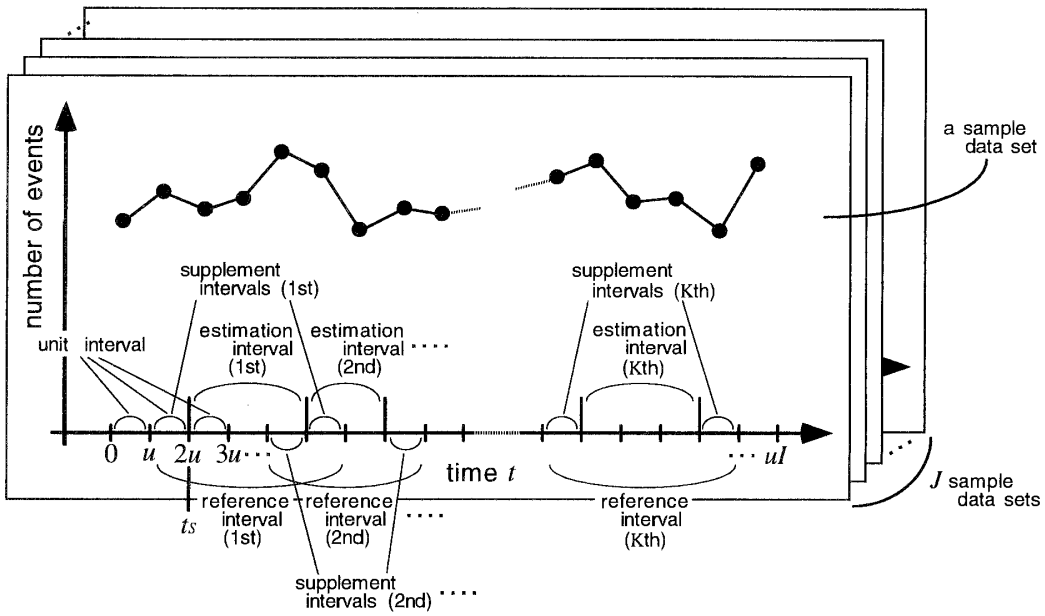

Figure 2 Description of the intervals and sample data sets

Here we have the following assumptions;

Assumptions:

1. The sample data are collected from the time point 0.
2. The estimation intervals are successive but do not overlap each other.

Based on the assumptions, some of the above values naturally have the following relations;

$$\sum_{k=1}^{K} \tau_{Ek} + t_S \le uI \tag{1}$$

$$\tau_{Rk} = \min\left\{uI, \sum_{h=1}^{k} \tau_{Eh} + t_S + \tau_{Sk}\right\} - \max\left\{0, \sum_{h=0}^{k-1} \tau_{Eh} + t_S - \tau_{Sk}\right\}, \qquad \tau_{E0} = 0 \tag{2}$$

$$X_{Ek} = \frac{1}{J}\sum_{j=1}^{J}\sum_{i=i_1}^{i_2} n_{ij}, \qquad i_1 = \sum_{h=0}^{k-1} \frac{\tau_{Eh} + t_S}{u} + 1, \quad i_2 = \sum_{h=1}^{k} \frac{\tau_{Eh} + t_S}{u}, \quad \tau_{E0} = 0 \tag{3}$$

$$X_{Rk} = \frac{1}{J}\sum_{j=1}^{J}\sum_{i=i_3}^{i_4} n_{ij}, \qquad i_3 = \max\left\{1, \sum_{h=0}^{k-1} \frac{\tau_{Eh} + t_S}{u} + 1 - \frac{\tau_{Sk}}{u}\right\}, \quad i_4 = \min\left\{I, \sum_{h=1}^{k} \frac{\tau_{Eh} + t_S}{u} + \frac{\tau_{Sk}}{u}\right\},$$

$$\tau_{E0} = 0. \tag{4}$$

Here i_1 and i_2 are the index of the first and the last unit interval involved in the k th estimation interval respectively and i_3 and i_4 are in the k th reference interval respectively. The max{} and min{} operations to calculate τ_{Rk}, i_3 and i_4 are for adjustment at either end of the period in which the sample data are collected.

4.2 Pseudo estimator

Law and Kelton [18] used the value of X_{Ek} for the estimation of the piecewise constant intensity. X_{Ek} is in itself the sufficient estimator of the mean of the Poisson distribution in the estimation interval. However, unless we obtain many sets of sample data, it does not work well. Here we propose an estimator Y_k instead of X_{Ek}, which is given by

$$Y_k = \frac{\tau_{Ek}}{\tau_{Rk}} X_{Rk}. \tag{5}$$

This estimator is to introduce the sample data in the adjacent supplement intervals into the estimation for the estimation interval. Since it is derived through pseudo enlargement of the sample size, we call it the pseudo estimator. It may violate a principle of statistics that an estimator must be derived from logically related sample data. However, if the Poisson variables in the supplement intervals had similar characteristics to that in the estimation interval (namely if they had mutually close mean values), we could take into account the sample data in the supplement intervals. We emphasize that when we can not obtain sufficient sample data sets, then it might be better to perform an imprecise estimation procedure to get a good value than a precise estimation procedure to get a wrong value. This is a kind of engineering-oriented way of thinking. At the same time, we do not intend to insist stubbornly that the pseudo estimation by (5) is always better than the ordinary one. Actually the pseudo estimator Y_k gives better estimation solely under a certain condition than the Law and Kelton's method. The condition to obtain an effective pseudo estimator will be shown in the next section.

4.3 Condition of the pseudo estimator's effectiveness

In this section we show a condition that Y_k gives more accurate estimated value than the Law and Kelton's. Here the word "accuracy" means the possibility that the estimator falls into a certain domain around the true expectation H_{Ek}.

Firstly, for arbitrary small $\varepsilon > 0$, we have from Chebychev's inequality

$$\Pr\left\{\left|X_{Ek} - H_{Ek}\right| \geq \varepsilon\right\} \leq \frac{H_{Ek}}{J\varepsilon^2}. \tag{6}$$

And rewriting (6) as

$$\Pr\left\{\left|X_{Ek} - H_{Ek}\right| < \varepsilon\right\} \geq 1 - \frac{H_{Ek}}{J\varepsilon^2} = \alpha, \tag{7}$$

then we have

$$\varepsilon = \sqrt{\frac{H_{Ek}}{(1-\alpha)J}}, \qquad \text{for } \alpha < 1. \tag{8}$$

Hence putting (8) back into (7), we have the domain $\mathbf{D}_{Xk}$ around H_{Ek} into which X_{Ek} falls with the probability of greater than or equal to α as follows;

$$\mathbf{D}_{Xk} = \left\{ X \middle| X \in \left(H_{Ek} - \sqrt{\frac{H_{Ek}}{(1-\alpha)J}}, H_{Ek} + \sqrt{\frac{H_{Ek}}{(1-\alpha)J}} \right) \right\}. \tag{9}$$

Similarly, applying Markov's inequality to Y_k around H_{Ek}, we have

$$\Pr\left\{\left|Y_k - H_{Ek}\right| \geq \varepsilon\right\} \leq \frac{1}{\varepsilon^2}\left[\left\{ H_{Ek} - \frac{H_{Rk} \cdot \tau_{Ek}}{\tau_{Rk}} \right\}^2 + \frac{H_{Rk} \cdot {\tau_{Ek}}^2}{J{\tau_{Rk}}^2} \right], \tag{10}$$

where we used the relation

$$\begin{aligned} E\left\{{Y_k}^2\right\} &= Var\left\{Y_k\right\} + E\left\{Y_k\right\}^2 \\ &= \frac{H_{Rk} \cdot {\tau_{Ek}}^2}{J{\tau_{Rk}}^2} + \left(\frac{H_{Rk} \cdot \tau_{Ek}}{\tau_{Rk}} \right)^2. \end{aligned} \tag{11}$$

Rewriting (10) as

$$\Pr\left\{\left|Y_k - H_{Ek}\right| < \varepsilon\right\} \geq 1 - \frac{1}{\varepsilon^2}\left[\left\{ H_{Ek} - \frac{H_{Rk} \cdot \tau_{Ek}}{\tau_{Rk}} \right\}^2 + \frac{H_{Rk} \cdot {\tau_{Ek}}^2}{J{\tau_{Rk}}^2} \right] = \alpha, \tag{12}$$

then we have

$$\varepsilon = \sqrt{\frac{\left\{ H_{Ek} - \frac{H_{Rk} \cdot \tau_{Ek}}{\tau_{Rk}} \right\}^2 + \frac{H_{Rk} \cdot {\tau_{Ek}}^2}{J{\tau_{Rk}}^2}}{1-\alpha}}, \qquad \text{for } \alpha < 1. \tag{13}$$

Therefore we have the domain $\mathbf{D}_{Yk}$ around H_{Ek} into which Y_k falls with the probability of greater than or equal to α as follows;

$$\mathbf{D}_{Yk} = \left\{ Y \middle| Y \in \left(H_{Ek} - \sqrt{\frac{\left\{ H_{Ek} - \frac{H_{Rk} \cdot \tau_{Ek}}{\tau_{Rk}} \right\}^2 + \frac{H_{Rk} \cdot {\tau_{Ek}}^2}{J{\tau_{Rk}}^2}}{1-\alpha}}, \right. \right.$$

$$\left. \left. H_{Ek} + \sqrt{\frac{\left\{ H_{Ek} - \frac{H_{Rk} \cdot \tau_{Ek}}{\tau_{Rk}} \right\}^2 + \frac{H_{Rk} \cdot {\tau_{Ek}}^2}{J{\tau_{Rk}}^2}}{1-\alpha}} \right) \right\}. \tag{14}$$

When $\mathbf{D}_{Yk} \subseteq \mathbf{D}_{Xk}$, Y_k can be close to H_{Ek} with greater probability than X_{Ek}. Therefore, from (9) and (14), if the following inequality (15) is satisfied, the pseudo estimator Y_k gives more accuracy than X_{Ek}.

$$\sqrt{\frac{\left\{H_{Ek}-\frac{H_{Rk}\cdot\tau_{Ek}}{\tau_{Rk}}\right\}^2+\frac{H_{Rk}\cdot\tau_{Ek}^2}{J\tau_{Rk}^2}}{1-\alpha}}\le\sqrt{\frac{H_{Ek}}{(1-\alpha)J}} \tag{15}$$

Eliminating the common terms to both sides in (15), finally we obtain for $H_{Ek}\tau_{Rk}\neq H_{Rk}\tau_{Ek}$

$$J\le\frac{\frac{H_{Ek}}{\tau_{Ek}^2}-\frac{H_{Rk}}{\tau_{Rk}^2}}{\left(\frac{H_{Ek}}{\tau_{Ek}}-\frac{H_{Rk}}{\tau_{Rk}}\right)^2}. \tag{16}$$

The inequality (16) is a sufficient condition that the pseudo estimator Y_k is effective indifferently to α. It means that when the number of sample data sets J is less than or equal to the right side value of (16), we should use Y_k as an estimator for the kth estimation interval.

In course of the derivation of (16) we used Chebychev's (and Markov's) inequality which is often said to be far from sharp (p.234 in Feller [9]). However, in the preceding discussion, no absolute sharpness is required because the inequality (15) is derived by relative comparison between the two domains (9) and (14). This relative treatment warrants the validity of (16) in spite of the dullness of Chebychev's inequality for itself.

Figure 3 shows the feasibility of the condition (16) (see Appendix). The feasible domain is the shaded area surrounded by the slanting parabola (A) and the straight line (B). It shows that a feasible Φ $(=JH_{Rk})$ exists for all Γ $(=JH_{Ek}>0)$ and that the feasible range of Γ extends as Φ increases. Further when $\tau_{Rk}/\tau_{Ek}\to+\infty$, the parabola reduces to the two vertical straight lines and the feasible Φ for $\Gamma>1$ is limited to considerably large values (approaching $+\infty$). On the other hand, when $\tau_{Rk}/\tau_{Ek}\to 1+0$, the parabola reduces to the two straight lines with 45 degrees gradient and the feasible domain reduces to the straight line $\Phi=\Gamma$ $(H_{Ek}=H_{Rk})$.

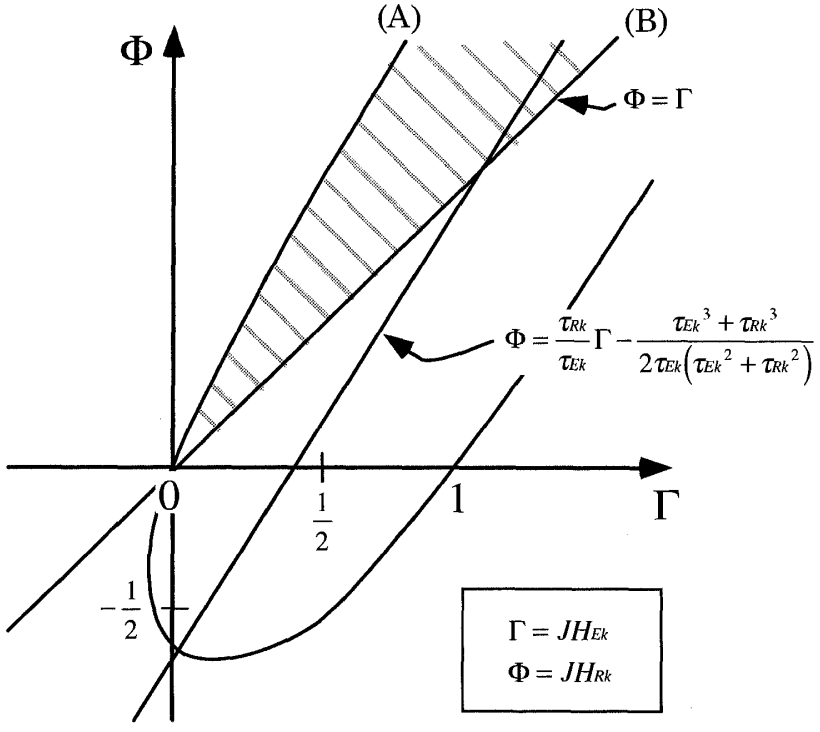

Figure 3 Feasible domain of the condition (16)

From an application point of view, the condition (16) is still insufficient since it includes true expectations H_{Ek} and H_{Rk}, which can not be identified beforehand. Therefore we propose to use the sample means X_{Ek} and X_{Rk} instead of H_{Ek} and H_{Rk} respectively as shown by (17) in order to decide if the pseudo estimator Y_k should be used or not.

$$J \leq \frac{\dfrac{X_{Ek}}{\tau_{Ek}{}^2} - \dfrac{X_{Rk}}{\tau_{Rk}{}^2}}{\left(\dfrac{X_{Ek}}{\tau_{Ek}} - \dfrac{X_{Rk}}{\tau_{Rk}}\right)^2} \tag{17}$$

Although this makeshift treatment can distort the theoretical and universal validity of (16), the experimental results in the next section show that it still holds statistical effectiveness.

5. Estimation Algorithm and Experimental Results

5.1 Estimation algorithm

Based on the theoretical result in the preceding section, we propose the following algorithm to estimate the NHPP piecewise intensities for all the estimation intervals.

Pseudo Estimation Algorithm:

For $k = 1$ to K
  Calculate X_{Ek} by (3)
  $\xi_k \leftarrow X_{Ek}$
  $\tau_{Sk} \leftarrow u$
  Do While $\tau_{Sk} \leq \bar{\tau}_{Sk}$
    Calculate τ_{Rk} by (2)
    Calculate X_{Rk} by (4)
    If the condition (17) is satisfied *Then*
      Calculate Y_k by (5)
      $\xi_k \leftarrow Y_k$
      Exit Do
    End If
    $\tau_{Sk} \leftarrow \tau_{Sk} + u$
  Loop
Next

In the algorithm, ξ_k denotes the resultant estimated value for the k th estimation interval. At the third step, ξ_k is set to the sample mean X_{Ek} as a default value. In the *Do* loop, increasing the length of the corresponding reference interval, ξ_k is replaced by the pseudo estimator Y_k if the condition (17) is satisfied. The *Do* loop is executed until the length of the supplement interval reaches a predetermined limit value $\bar{\tau}_{Sk}$ or the condition (17) is satisfied. This procedure yields the valid pseudo estimator with the possible shorter reference interval (with more adjacent sample data).

5.2 Experiments

In order to confirm the effectiveness of the pseudo estimation algorithm, we carried out the experiments as follows;

Experiment Procedure:

[0] Determine I, K, u, ts and τ_{Ek} as the relation (1) is satisfied.
[1] Determine J and $\bar{\tau}_{Sk}$.

[2] Generate J sets of sample data of NHPP with the mean function of $\Lambda(t)$ by generating a Poisson random variable n_{ij} with mean $\Lambda(ui)-\Lambda(u(i-1))$ for the i th unit interval of the j th sample data set.

[3] Execute the pseudo estimation algorithm on the sample data sets.

[4] Calculate the mean square relative error (MSRE) of ξ_k for all the estimation intervals by

$$\text{MSRE of } \xi_k = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{\xi_k - H_{Ek}}{H_{Ek}}\right)^2, \quad H_{Ek} = \Lambda\left(\sum_{h=1}^{k}\tau_{Eh} + t_S\right) - \Lambda\left(\sum_{h=0}^{k-1}\tau_{Eh} + t_S\right), \quad \tau_{E0} = 0. \tag{18}$$

[5] Calculate the MSRE of X_{Ek} for all the estimation intervals by

$$\text{MSRE of } X_{Ek} = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{X_{Ek} - H_{Ek}}{H_{Ek}}\right)^2. \tag{19}$$

Irregular fluctuation case

In order to examine the effectiveness for irregularly fluctuating NHPP, we consider the mean function as

$$\Lambda(t) = 6t^5 - 45t^4 + 130t^3 - 180t^2 + 210t + \frac{15}{2} - \frac{15}{2}\cos 6t + \frac{45}{8}\sin 8t, \tag{20}$$

which is irregular enough as shown in Figure 4 (the polygonal curve connects the values of H_{Ek}). The experiment was done with $I=60$, $u=0.05$, $t_S=0$ and three patterns of τ_{Ek} as $\tau_{Ek}=0.05(K=60)$, $0.10(K=30)$, $0.15(K=20)$ where the length of the estimation interval is identical for all k. We considered the varieties of J and $\bar{\tau}_{Sk}$ as $J=1,2,\cdots,10$ and $\bar{\tau}_{Sk}=0.05, 0.10, 0.15$ in which the limit length of the supplement interval is identical for all k. Figure 4 is an example of X_{Ek} (Law and Kelton's estimator) and the estimator ξ_k resulting from our algorithm for $\tau_{Ek}=0.05$, $J=1$ and $\bar{\tau}_{Sk}=0.05$. As seen in the figure, ξ_k generally gives a closer value to the true expectation H_{Ek} than X_{Ek}. The algorithm was repeated 100 times for each variety of J and $\bar{\tau}_{Sk}$ so as to obtain a purposive averaged MSRE of ξ_k and X_{Ek}. Table 1 shows the MSREs of ξ_k and X_{Ek} for every J and $\bar{\tau}_{Sk}$.

From Table 1, we see that the MSRE of ξ_k is stably better than that of X_{Ek} for smaller J. However, as J is getting larger, the improvement rate by ξ_k becomes less significant. For $\tau_{Ek}=0.05$ and $\bar{\tau}_{Sk}=0.05$, while more than 20% ($=100-0.17909/0.22518\times 100$) is improved by ξ_k for $J=1$, only 10% ($=100-0.01977/0.02183\times 100$) improvement is done for $J=10$. Further, when J is large, MSRE of X_{Ek} can happen to be better than that of ξ_k. For example, the MSRE of X_{Ek} for $\tau_{Ek}=0.15$ and $J=6$ is better than that of ξ_k for $\bar{\tau}_{Sk}=0.05$. These are because the ordinary sample mean X_{Ek} itself gains its accuracy as J becomes large.

For a combination of τ_{Ek} and J, the MSRE of ξ_k is improved monotonously as $\bar{\tau}_{Sk}$ increases. This is because the opportunity to find a reference interval which satisfies the condition (16) becomes larger as $\bar{\tau}_{Sk}$ increases. However, here we note that the increase of $\bar{\tau}_{Sk}$ means nothing but expansion of the range in which the reference interval to satisfy (16) is searched for. In Figure 3, the increase of the length of the reference interval corresponds to the increase of the gradient of the principal axis of the parabola (A), so that the feasible Γ is limited to the range of small values with an extremely large reference interval. Nevertheless, theoretically speaking, the more $\bar{\tau}_{Sk}$ increases, the better the MSRE of ξ_k becomes. For example, if $\bar{\tau}_{Sk}$ is set to uI, the pseudo estimation algorithm examines all the possible reference intervals then gives the best MSRE.

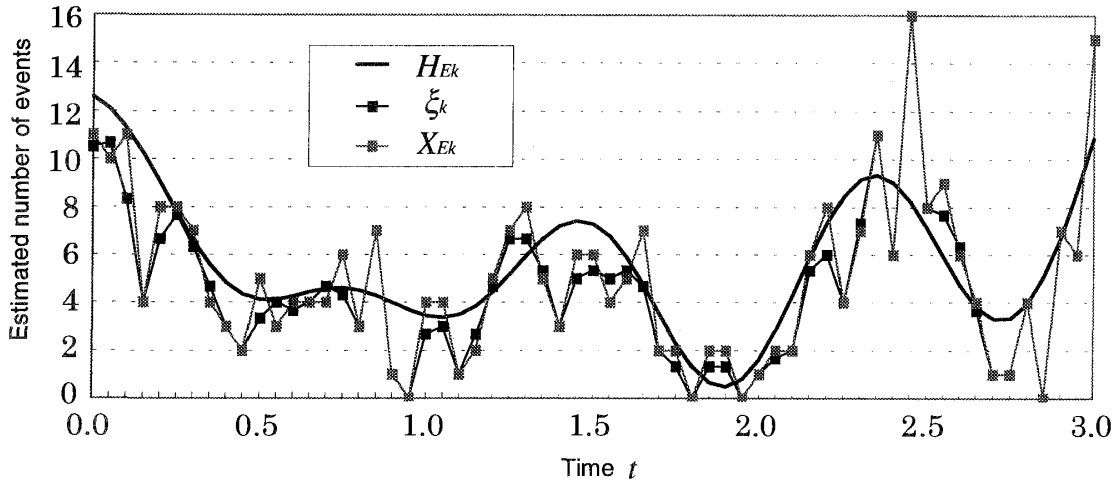

Figure 4 An example of the experiment result for the irregular fluctuation ($\tau_{Ek}=0.05$, $J=1$, $\bar{\tau}_{Sk}=0.05$)

Table 1 Experiment results for the irregular fluctuation

J	$\tau_{Ek}=0.05$				$\tau_{Ek}=0.10$				$\tau_{Ek}=0.15$			
	MSRE of X_{Ek}	MSRE of ξ_k			MSRE of X_{Ek}	MSRE of ξ_k			MSRE of X_{Ek}	MSRE of ξ_k		
		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$
1	0.22518	0.17909	0.17026	0.16711	0.11258	0.09846	0.09605	0.09273	0.07472	0.06788	0.06680	0.06546
2	0.11880	0.09811	0.09250	0.08919	0.05567	0.04913	0.04810	0.04710	0.03533	0.03323	0.03267	0.03219
3	0.07296	0.06251	0.05829	0.05683	0.03364	0.03087	0.03006	0.02958	0.02258	0.02192	0.02154	0.02137
4	0.05601	0.04730	0.04528	0.04434	0.02747	0.02610	0.02498	0.02449	0.01907	0.01849	0.01821	0.01793
5	0.04307	0.03699	0.03572	0.03461	0.02096	0.01990	0.01929	0.01895	0.01382	0.01346	0.01318	0.01305
6	0.03765	0.03270	0.03069	0.02964	0.01807	0.01754	0.01739	0.01708	0.01135	0.01137	0.01121	0.01110
7	0.03421	0.02952	0.02827	0.02755	0.01597	0.01508	0.01482	0.01450	0.00960	0.00956	0.00951	0.00943
8	0.02841	0.02514	0.02423	0.02338	0.01389	0.01346	0.01320	0.01302	0.00885	0.00882	0.00870	0.00863
9	0.02535	0.02279	0.02154	0.02096	0.01236	0.01218	0.01200	0.01181	0.00806	0.00797	0.00789	0.00787
10	0.02183	0.01977	0.01879	0.01829	0.01064	0.01067	0.01051	0.01036	0.00693	0.00686	0.00679	0.00674

Monotonous and cyclic trend case

As supplementary experiments, here we also consider a monotonous and a cyclic trend case. All the experiments here were done using the same parameters and conditions as in the irregular fluctuation case above.

For a monotone case, we chose the mean function as

$$\Lambda(t)=10t^3+50t, \tag{21}$$

which increases monotonously with time. Figure 5 shows its trend and an example of the result for $\tau_{Ek}=0.05$, $J=1$ and $\bar{\tau}_{Sk}=0.05$. Due to the nature of Poisson distribution, its variance increases with time, therefore the plots of X_{Ek} for large t fluctuate much. Nevertheless, the estimator ξ_k gives a closer value to H_{Ek} almost everywhere except where sample values (X_{Ek}) are extremely far from the true values (H_{Ek}) such as seen around $t=2.7$. Table 2 shows the MSREs of ξ_k and X_{Ek} for every J and $\bar{\tau}_{Sk}$. General tendency of the results are just as same as seen in Table 1, but in this case there is no reverse phenomenon between the MSREs

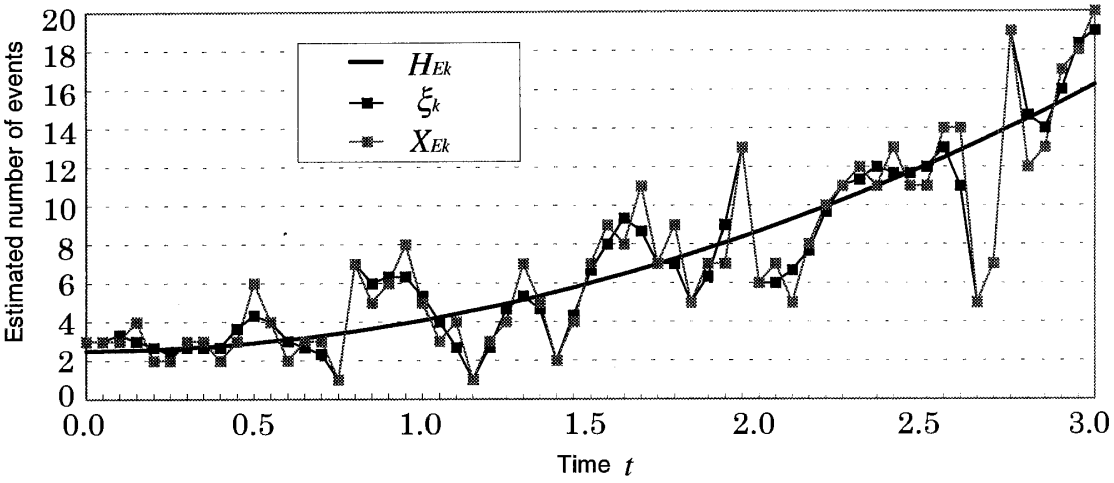

Figure 5 An example of the experiment result for the monotonous trend ($\tau_{Ek}=0.05$, $J=1$, $\bar{\tau}_{Sk}=0.05$)

Table 2 Experiment results for the monotonous trend

J	$\tau_{Ek}=0.05$				$\tau_{Ek}=0.10$				$\tau_{Ek}=0.15$			
	MSRE of X_{Ek}	MSRE of ξ_k			MSRE of X_{Ek}	MSRE of ξ_k			MSRE of X_{Ek}	MSRE of ξ_k		
		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$
1	0.20752	0.17626	0.16979	0.16620	0.10293	0.09310	0.09097	0.08904	0.06879	0.05930	0.05840	0.05753
2	0.09972	0.08480	0.08076	0.07948	0.05224	0.04503	0.04351	0.04267	0.03440	0.03156	0.03089	0.03026
3	0.06521	0.05503	0.05292	0.05181	0.03303	0.02902	0.02784	0.02738	0.02127	0.01970	0.01904	0.01854
4	0.04993	0.04283	0.04130	0.04047	0.02462	0.02162	0.02090	0.02050	0.01668	0.01538	0.01495	0.01466
5	0.03985	0.03451	0.03338	0.03266	0.01989	0.01736	0.01682	0.01655	0.01340	0.01197	0.01172	0.01157
6	0.03272	0.02826	0.02697	0.02636	0.01657	0.01482	0.01432	0.01398	0.01111	0.01027	0.01001	0.00981
7	0.02846	0.02478	0.02371	0.02324	0.01416	0.01270	0.01223	0.01193	0.00995	0.00912	0.00879	0.00872
8	0.02445	0.02103	0.02019	0.01983	0.01202	0.01061	0.01024	0.00996	0.00811	0.00749	0.00718	0.00707
9	0.02280	0.01986	0.01922	0.01873	0.01192	0.01089	0.01044	0.01023	0.00787	0.00729	0.00712	0.00702
10	0.02011	0.01728	0.01653	0.01636	0.01027	0.00918	0.00883	0.00866	0.00693	0.00635	0.00619	0.00612

of ξ_k and X_{Ek} "fortunately". Although it is needless to say that this is only a fortunate occasion, it can be said that the possibility to get better ξ_k is much greater in the monotone case than in the irregular fluctuation case because of the monotonous similarity of the Poisson process in the adjacent intervals.

For a cyclic trend case, consider the mean function as

$$\Lambda(t)=30t^2+100t+5\cos 10t, \tag{22}$$

which gives cyclic fluctuations with a monotonously increasing trend. Conventionally Figure 6 and Table 3 show a result example and the MSREs of ξ_k and X_{Ek} respectively. Here again the estimator ξ_k generally gives a closer value to H_{Ek}, but one can observe some reverse phenomena in this case. For example, the MSRE of X_{Ek} for $\tau_{Ek}=0.15$ and $J=9$ is better than that of ξ_k for $\bar{\tau}_{Sk}=0.05$ and 0.10. The other overall tendency of the results is just as same as in the former two cases.

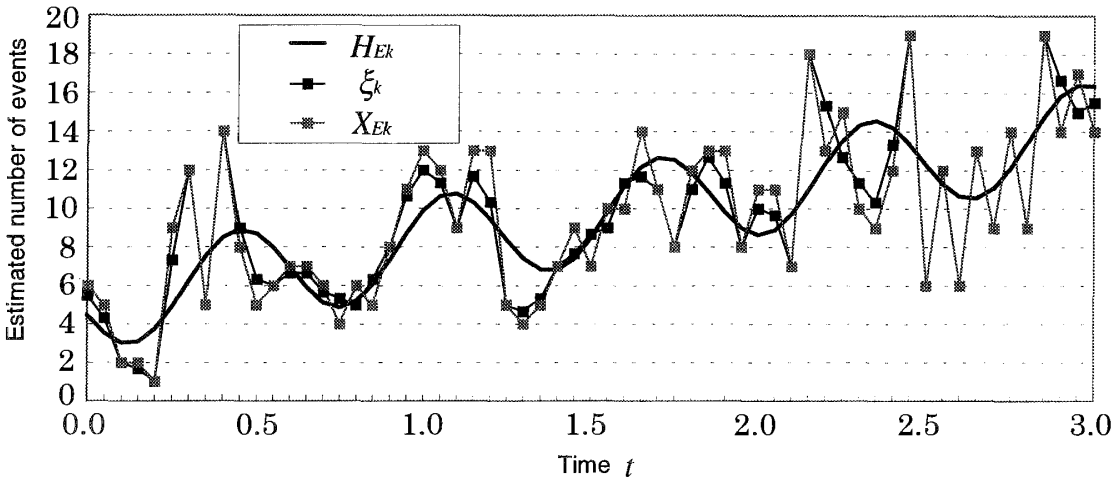

Figure 6 An example of the experiment result for the cyclic trend ($\tau_{Ek}=0.05$, $J=1$, $\bar{\tau}_{Sk}=0.05$)

Table 3 Experiment results for the cyclic trend

J	$\tau_{Ek}=0.05$				$\tau_{Ek}=0.10$				$\tau_{Ek}=0.15$			
	MSRE of X_{Ek}	MSRE of ξ_k			MSRE of X_{Ek}	MSRE of ξ_k			MSRE of X_{Ek}	MSRE of ξ_k		
		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$		$\bar{\tau}_{Sk}=0.05$	$\bar{\tau}_{Sk}=0.10$	$\bar{\tau}_{Sk}=0.15$
1	0.12441	0.10754	0.10315	0.09992	0.06224	0.05627	0.05403	0.05194	0.04026	0.03717	0.03612	0.03564
2	0.06183	0.05292	0.05074	0.04901	0.03138	0.02861	0.02788	0.02692	0.02169	0.02013	0.01975	0.01919
3	0.04105	0.03598	0.03363	0.03219	0.02052	0.01876	0.01792	0.01728	0.01352	0.01299	0.01274	0.01258
4	0.03076	0.02690	0.02561	0.02448	0.01534	0.01427	0.01392	0.01332	0.00988	0.00961	0.00946	0.00939
5	0.02476	0.02187	0.02074	0.01999	0.01201	0.01143	0.01108	0.01072	0.00819	0.00805	0.00778	0.00768
6	0.02048	0.01775	0.01692	0.01611	0.01029	0.00985	0.00973	0.00942	0.00691	0.00678	0.00657	0.00648
7	0.01723	0.01501	0.01436	0.01366	0.00882	0.00829	0.00818	0.00796	0.00581	0.00575	0.00563	0.00559
8	0.01572	0.01392	0.01328	0.01270	0.00792	0.00756	0.00738	0.00711	0.00530	0.00514	0.00508	0.00506
9	0.01338	0.01156	0.01092	0.01050	0.00668	0.00657	0.00647	0.00628	0.00445	0.00452	0.00446	0.00445
10	0.01280	0.01150	0.01078	0.01028	0.00634	0.00615	0.00600	0.00581	0.00437	0.00440	0.00432	0.00426

5.3 Application to the actual case

Admitting that we do not know its true mean function, we applied the pseudo estimation algorithm to the sample data of Figure 1. The case corresponds to $I=K=96$, $u=\tau_{Ek}=1$, $ts=0$ and $J=1$, and the limit of the supplement interval is set as $\bar{\tau}_{Sk}=2$. Figure 7 shows the result of the application of the algorithm. It shows that the estimator ξ_k works as a moderator against fluctuations to some extent. However, if the fluctuations in the adjacent intervals are so large that they do not satisfy the condition (17), ξ_k remains as X_{Ek}. This is because our pseudo estimation method basically stands on the presumption that the data in the adjacent intervals hold the mutually similar characteristics, while once no evidence of mutual similarity in the adjacent intervals is found, it does not take the risk to have a "groundless" pseudo estimator using the dissimilar adjacent sample data. On the other hand, by moderating the acceptable fluctuations, it also avoids the risk to be too much dependent on only one or a few collected sample data sets.

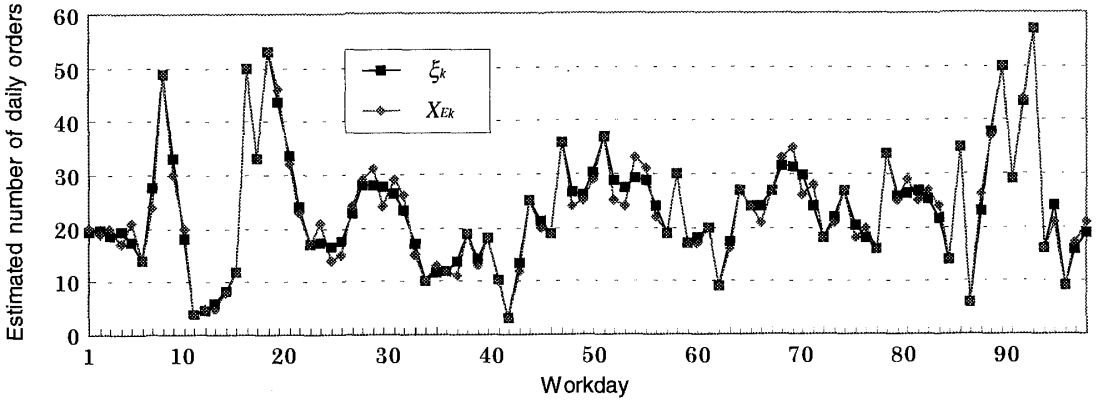

Figure 7 Application of the proposed method to the actual case

6. Difference between the Pseudo Estimation and the Local Averaging

Our idea of taking into account the adjacent sample data in calculating the pseudo estimator is similar to the local averaging of smoothing methods. It is also true that both methods stand on the common assumption that the adjacent data (or response values) should hold the mutually similar characteristics. However, such surface similarity appears to be less important because of the following fundamental difference between them. Note that this is only the clarification of the difference, not the discussion of the pros and cons.

(1) Difference in conception

Smoothing is a method to produce a representative function which summarizes overall scatterplot data, while the pseudo estimator is to estimate the expected number of events in each piecewise time interval.

(2) Difference in objective

Smoothing is generally used in order to capture a fitting function which gives enough smoothness and continuity as well as the minimized residuals. Further it is often required to have good prediction ability for new observations in the future. The piecewise pseudo estimator is to estimate Poisson expectations with better MSRE and the resultant estimates do not form a function.

(3) Difference in supposed data

Most of the smoothing methods are based on the assumption that the observed data consist of true response value and associated white noise which is often supposed to be homoscedastic everywhere. Especially the span for the local averaging is determined by CV, which requires the homoscedastic white noise assumption. However, the observations from NHPP are the realization from the distribution with time-varying mean and variance. The pseudo estimation method can be applied to such time-varying random observations.

(4) Difference in procedure

Once the span that gives the globally minimal CV is determined, it is applied to local averaging over all the observations uniformly. On the other hand, the pseudo estimation needs the local calculation to determine the length of the reference interval for each estimation interval and consequently it gives the varying span which reflects the observations in the adjacent intervals.

7. Concluding Remarks

In this paper we proposed a piecewise pseudo estimation method for NHPP using the sample data in the adjacent intervals and derived the condition to obtain an effective estimated value by the pseudo estimator. The proposed method is useful when the parametric representation of the

NHPP mean function is unknown and only one or a few sets of sample data can be obtained such as Figure 1. Compared to the nonparametric method by Law and Kelton [18], the experimental results showed that it gains more statistical accuracy even with one or a few sample data sets.

As for the pseudo estimation algorithm given in 5.1, it can be said that the algorithm gives the best MSRE with the possible largest limit of the supplement interval $\bar{\tau}_{Sk}$. However, even though the theoretical condition (16) is satisfied, it can give a "groundless" value to calculate the pseudo estimator by using data in the supplement intervals at a great distance from the target estimation interval. This problem is related to the principle of statistics and still remains to be discussed.

Finally, note that all the statistical accuracy of the proposed method in this paper is referred to the mean-based metrics. This implies that although the proposed method is expected to be effective on the average, it can happen to get a worse result than the other method in each individual case.

Appendix

Here we put a brief description of the derivation of Figure 3.

The inequality (16) can be rewritten as

$$\frac{1}{\tau_{Ek}{}^2}(JH_{Ek})^2 - \frac{2}{\tau_{Ek}\tau_{Rk}}(JH_{Ek}JH_{Rk}) + \frac{1}{\tau_{Rk}{}^2}(JH_{Rk})^2 - \frac{1}{\tau_{Ek}{}^2}JH_{Ek} + \frac{1}{\tau_{Rk}{}^2}JH_{Rk} \le 0. \quad \text{(A1)}$$

Letting

$$\Gamma = JH_{Ek}, \Phi = JH_{Rk}, a = \frac{1}{\tau_{Ek}{}^2}, b = \frac{1}{\tau_{Rk}{}^2}, c = 0, f = \frac{1}{2\tau_{Rk}{}^2}, g = -\frac{1}{2\tau_{Ek}{}^2} \text{ and } h = -\frac{1}{\tau_{Ek}\tau_{Rk}},$$

naturally we have

$$\Phi \ge \Gamma, \quad \text{(A2)}$$

and consider

$$a\Gamma^2 + 2h\Gamma\Phi + b\Phi^2 + 2g\Gamma + 2f\Phi + c \le 0. \quad \text{(A3)}$$

Define

$$\mathbf{Q} = \begin{pmatrix} a & h \\ h & b \end{pmatrix} \text{ and } \mathbf{A} = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix},$$

then since $rank\mathbf{Q} = 1$ and $rank\mathbf{A} = 3$, the inequality (A3) is a domain surrounded by a parabola.

The eigen values of $\mathbf{Q}$ are

$$\lambda_1 = 0 \text{ and } \lambda_2 = \frac{1}{\tau_{Ek}{}^2} + \frac{1}{\tau_{Rk}{}^2},$$

and their corresponding normalized eigen vectors are

$$\mathbf{e}_1 = \begin{pmatrix} \dfrac{\tau_{Ek}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} \\ \dfrac{\tau_{Rk}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} \end{pmatrix} \text{ and } \mathbf{e}_2 = \begin{pmatrix} \dfrac{\tau_{Rk}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} \\ \dfrac{-\tau_{Ek}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} \end{pmatrix}.$$

Now there exist a rotating matrix $\mathbf{S}_1$ and a translating matrix $\mathbf{S}_2$ such as

$$\mathbf{S}_1 = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \text{ and } \mathbf{S}_2 = \begin{pmatrix} \mathbf{E} & \mathbf{X}_0 \\ \mathbf{0} & 1 \end{pmatrix}, \text{ where } \mathbf{P} = (\mathbf{e}_1 \quad \mathbf{e}_2) \text{ and } \mathbf{X}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Let $\mathbf{A}'$ denote the matrix of the coefficients of the canonical parabola with coordinate-

transformation, then we have

$$\mathbf{A}' = (\mathbf{S}_1\mathbf{S}_2)^T \mathbf{A}(\mathbf{S}_1\mathbf{S}_2) = \begin{pmatrix} \lambda_1 & 0 & g' \\ 0 & \lambda_2 & \lambda_2 y_0 + f' \\ g' & \lambda_2 y_0 + f' & c - \dfrac{f'^2}{\lambda_2} + 2g'x_0 \end{pmatrix}, \tag{A4}$$

in which $f' = \dfrac{-(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)}{2\tau_{Ek}{}^2\tau_{Rk}{}^2\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}}$ and $g' = \dfrac{\tau_{Ek} - \tau_{Rk}}{2\tau_{Ek}\tau_{Rk}\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}}$.

Since $\mathbf{A}'$ is the matrix of the canonical coefficients, the following equations must be satisfied in (A4).

$$\lambda_2 y_0 + f' = 0$$

$$c - \frac{f'^2}{\lambda_2} + 2g'x_0 = 0$$

Hence they yield

$$x_0 = \frac{(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)^2\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}}{4\tau_{Ek}\tau_{Rk}(\tau_{Ek} - \tau_{Rk})(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} \quad \text{and} \quad y_0 = \frac{\tau_{Ek}{}^3 + \tau_{Rk}{}^3}{2(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}},$$

then we have

$$\mathbf{S}_1\mathbf{S}_2 = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}\begin{pmatrix} \mathbf{E} & \mathbf{X}_0 \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \dfrac{\tau_{Ek}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} & \dfrac{\tau_{Rk}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} & \dfrac{(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)^2}{4\tau_{Rk}(\tau_{Ek} - \tau_{Rk})(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} + \dfrac{\tau_{Rk}(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)}{2(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} \\ \dfrac{\tau_{Rk}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} & \dfrac{-\tau_{Ek}}{\sqrt{\tau_{Ek}{}^2 + \tau_{Rk}{}^2}} & \dfrac{(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)^2}{4\tau_{Ek}(\tau_{Ek} - \tau_{Rk})(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} - \dfrac{\tau_{Ek}(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)}{2(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} \\ 0 & 0 & 1 \end{pmatrix}.$$

From these, the origin of the parabola $\mathbf{O}$ is

$$\mathbf{O} = \begin{pmatrix} \dfrac{(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)^2}{4\tau_{Rk}(\tau_{Ek} - \tau_{Rk})(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} + \dfrac{\tau_{Rk}(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)}{2(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} \\ \dfrac{(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)^2}{4\tau_{Ek}(\tau_{Ek} - \tau_{Rk})(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} - \dfrac{\tau_{Ek}(\tau_{Ek}{}^3 + \tau_{Rk}{}^3)}{2(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)^2} \end{pmatrix}$$

and the gradient of the parabola's principal axis is given by the gradient of $\mathbf{e}_1$ as τ_{Rk} / τ_{Ek}. Therefore the principal axis is obtained as follows;

$$\Phi = \frac{\tau_{Rk}}{\tau_{Ek}}\Gamma - \frac{\tau_{Ek}{}^3 + \tau_{Rk}{}^3}{2\tau_{Ek}(\tau_{Ek}{}^2 + \tau_{Rk}{}^2)}.$$

Consequently the feasible domain of the conditions (A2) and (A3) is as shown in Figure 3.

Acknowledgment

We are pleased to thank Dr. Sakasegawa of Waseda University for his kind advice at the early stage of this study, and two anonymous referees for their helpful comments and constructive suggestions.

References

[1] G.G.Agarwal and W.J.Studden: Asymptotic integrated mean square error using least squares and bias minimizing splines. *The Annals of Statistics*, **8** (1980) 1307-1325.

[2] L.J.Bain and M.Engelhardt: Sequential probability ratio tests for the shape parameter of a nonhomogeneous Poisson process. *IEEE Transactions on Reliability*, **R-31** (1982) 79-83.

[3] I.Bar-David: Communication under the Poisson regime. *IEEE Transactions on Information Theory*, **IT-15** (1969) 31-37.

[4] J.S.Bendat and A.G.Piersol: *Random Data: Analysis and Measurement Procedures* (John Wiley & Sons, 1971).

[5] A.Buja, T.Hastie and R.Tibshirani: Linear smoothers and additive models. *The Annals of Statistics*, **17** (1989) 453-510.

[6] W.S.Cleveland: Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, **74** (1979) 828-836.

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

[8] B.Efron: Estimating the error rate of a prediction rule: improvement on cross-validation. *Journal of the American Statistical Association*, **78** (1983) 316-331.

[9] W.Feller: *An Introduction to Probability Theory and Its Applications, Vol.1: 3rd ed.* (John Wiley & Sons, New York, 1967).

[10]J.H.Friedman and B.W.Silverman: Flexible parsimonious smoothing and additive modeling. *Technometrics*, **31** (1989) 3-21.

[11]J.H.Friedman and W.Stuetzle: Projection pursuit regression. *Journal of the American Statistical Association*, **76** (1981) 817-823.

[12]J.Friedman and R.Tibshirani: The monotone smoothing of scatterplots. *Technometrics*, **26** (1984) 243-250.

[13]T.J.Hastie and R.J.Tibshirani: Nonparametric regression and classification part I-nonparametric regression. *NATO ASI Series. F*, **136** (1994) 62-69.

[14]J.S.Jang and D.S.Bai: Efficient sequential estimation in a nonhomogeneous Poisson process. *IEEE Transactions on Reliability*, **R-36** (1987) 255-258.

[15]E.P.C.Kao and S.Chang: Modeling time-dependent arrivals to service systems: a case in using a piecewise-polynomial rate function in a nonhomogeneous Poisson process. *Management Science*, **34** (1988) 1367-1379.

[16]R.W.Klein and S.D.Roberts: A time-varying Poisson arrival process generator. *Simulation*, **43** (1984) 193-195.

[17]J.T.Kvaløy and B.H.Lindqvist: TTT-based tests for trend in repairable systems data. *Reliability Engineering and System Safety*, **60** (1998) 13-28.

[18]A.M.Law and W.D.Kelton: *Simulation Modeling & Analysis: 2nd ed.* (McGraw-Hill, New York, 1991).

[19]S.Lee, J.R.Wilson and M.M.Crawford: Modeling and simulation of a nonhomogeneous Poisson process having cyclic behavior. *Communications in Statistics, Simulation and Computation*, **20** (1991) 777-809.

[20]L.M.Leemis: Nonparametric estimation of the cumulative intensity function for a nonhomogeneous Poisson process. *Management Science*, **37** (1991) 886-900.

[21]P.A.W.Lewis and G.S.Shedler: Statistical analysis on non-stationary series of events in a data base system. *IBM Journal of Research and Development*, **20** (1976) 465-482.

[22]S.E.Rigdon and A.P.Basu: The power law process: a model for the reliability of repairable systems. *Journal of Quality Technology*, **21** (1989) 251-260.

[23]S.E.Rigdon and A.P.Basu: The effect of assuming a homogeneous Poisson process when

the true process is a power law process. *Journal of Quality Technology*, **22** (1990) 111-117.
[24]B.W.Silverman: Spline smoothing: the equivalent variable kernel method. *The Annals of Statistics*, **12** (1984) 898-916.
[25]F.S.Swed and C.Eisenhart: Tables for testing the randomness of grouping in a sequence of alternatives. *The Annals of Mathematical Statistics*, **14** (1943) 66-87.

Kenichi Funaki
Production Engineering Research Laboratory,
Hitachi Ltd.
292 Yoshida-cho Totsuka-ku, Yokohama,
Kanagawa 244-0817 Japan
E-mail: funaki@perl.hitachi.co.jp