Asymptotic Properties of Mean Cumulative Function Estimators from Window-Observation Recurrence Data

Jianying Zuo*, Huaiqing Wu, William Q. Meeker

Department of Statistics, Iowa State University, Ames, IA 50011

Abstract

A variety of nonparametric and parametric methods have been used to estimate the mean cumulative function (MCF) for the recurrence data collected from the counting process. When the recurrence histories of some units are available in disconnected observation windows with gaps in between, Zuo, Meeker, and Wu (2008) showed that both the nonparametric and parametric methods can be extended to estimate the MCF. In this article, we establish some asymptotic properties of the MCF estimators for the window-observation recurrence data.

Key Words: Asymptotic normality, Consistency, Nonhomogeneous Poisson process, Nonparametric estimation

*Corresponding author

Email address: jiyizu@yahoo.com (Jianying Zuo)
1. Introduction

1.1. Background and Motivation

Window-observation recurrence data arise when the recurrence histories of some units are available in disconnected observation windows with gaps in between. Nelson (2003, page 75) gave an example in which window-observation recurrence data arise, and Zuo, Meeker, and Wu (2008) described two other applications. Nonparametric and parametric methods for analyzing recurrence data are available in many publications. Examples include Nelson (2003), Cook and Lawless (2007), Meeker and Escobar (1998), Lawless and Nadeau (1995), and Rigdon and Basu (2000). Among the quantities to be analyzed with the recurrence data are the locations and counts of recurrences, and statistics of interest include the mean cumulative number of recurrences. Zuo, Meeker, and Wu (2008) showed that both the nonparametric and parametric methods can be extended to estimate the mean cumulative function (MCF) for window-observation recurrence data. When there are time intervals with risk-set-size-zero (RSSZ) (i.e., time intervals in which no unit is under observation), they proposed the local hybrid estimator and the nonhomogeneous Poisson process (NHPP) hybrid estimator. As a continuation of Zuo, Meeker, and Wu (2008), this article establishes some asymptotic properties of the MCF estimators. Note that the hybrid estimators are finite-sample alternatives to the nonparametric (NP) estimator because the latter is downward-biased with the existence of RSSZ intervals. Because there is no information available for estimating the MCF in an RSSZ interval, RSSZ intervals need to go away asymptotically for the MCF estimators to be consistent. Therefore, in presenting the asymptotic properties of the MCF estimators, we only include the NP estimator and the NHPP estimators.

Andersen, Borgan, Gill, and Keiding (1993) provided comprehensive descriptions of the statistical models and methods that can be used to analyze the event history observed in continuous time. They derived the asymptotic (i.e., large-sample) properties for both the nonparametric and parametric estimators described in their book. Building on their conditions and theorems, we show that the NP and the NHPP MCF estimators of Zuo, Meeker, and Wu (2008) have the desirable asymptotic properties under some mild conditions that are generally satisfied in practical analysis.

1.2. Other Previous Work

Peña, Strawderman, and Hollander (2001) described nonparametric methods to estimate the inter-occurrence times with recurrent event data. Ghosh and Lin (2000) also focused on the nonparametric analysis of recurrent event data, possibly with a terminal event, and they estimated the mean frequency function, which is defined the same as the MCF. Both articles established the asymptotic properties of their respective nonparametric estimators, and used simulation studies for the finite-sample properties.
1.3. Overview

The remainder of this article is organized as follows. Section 2 outlines some common notation and assumptions to be used. Sections 3 and 4 establish the asymptotic properties for the NP and NHPP estimators, respectively. Section 5 gives some concluding remarks.

2. Notation and Assumptions

2.1. Notation for the Models

Let $N(t)$ denote the number of events in the time interval $(0, t]$. Then $\mu(t)$, the expectation of $N(t)$, is the MCF. If the MCF is differentiable, then $\nu(t) = d\mu(t)/dt$ is the recurrence rate, and $\nu(t) \times \Delta t$ can be interpreted as the approximate expected number of events to occur during the next short time interval $(t, t + \Delta t]$.

Let $K^{(n)}$ denote the counting process with the multiplicative intensity model $\lambda(t) = \nu(t) \delta(t)$, where $n$ is the total number of units, and $\delta(t)$ is the size of the risk set at time $t$ (i.e., the number of units that are in observation windows at time $t$). Note that $\delta(t) = \sum_{i=1}^{n} \delta_i(t)$, where

$$\delta_i(t) = \begin{cases} 1 & \text{if unit } i \text{ is under observation in a time window at time } t, \\ 0 & \text{otherwise}. \end{cases}$$

Then $M^{(n)}(t) = K^{(n)}(t) - \int_{0}^{t} \lambda(s) ds$ is a local square integrable martingale (see section IV.1.1, Andersen et al. 1993).

2.2. Notation for the Data

For window-observation recurrence data, let $w_i$ be the number of observation windows for unit $i$, and $r_i$ be the number of recurrences recorded in these observation windows for unit $i$. Let $t_{i1}, t_{i2}, ..., t_{ir_i}$ be the corresponding recurrence times, and $(t_{i1L}^{(n)}, t_{i1U}^{(n)}], ..., (t_{i_{w_i}L}^{(n)}, t_{i_{w_i}U}^{(n)}]$ be the corresponding observation windows for unit $i$.

When considering the superposition of all $n$ units, let $t_{\text{max}}$ be the largest end-of-observation time among all units. Then $\delta_i(t)$ is piece-wise constant over $[0, t_{\text{max}}]$. Let $z$ be the number of intervals with constant $\delta_i(t)$, and denote the $z$ intervals as $(t_{1L}^{(n)}, t_{1U}^{(n)}], (t_{2L}^{(n)}, t_{2U}^{(n)}], ..., (t_{zL}^{(n)}, t_{zU}^{(n)}]$, where $t_{1L}^{(n)} = 0$, $t_{zU}^{(n)} = t_{\text{max}}$, and $t_{iL}^{(n)} = t_{(i-1)U}^{(n)}$ for $i = 2, ..., z$. Note that $z$, $t_{\text{max}}$, and these intervals all depend on the data.

2.3. Assumptions

We assume that the set of observation windows for each unit is a simple event from the sample space consisting of a finite number of prespecified sets of non-overlapping observation windows, and that the probability of each simple event is positive. Furthermore, we assume that the overlap of
all these sets of windows leaves no gaps that result in an RSSZ interval. Let \( \tau < \infty \) be the ending time point for all these windows. We have \( t_{max} \leq \tau \) for all \( n \), and \( t_{max} \xrightarrow{a.s.} \tau \) as \( n \to \infty \). Note that \( \xrightarrow{a.s.} \) is used to denote “converges almost surely to.” These assumptions ensure that at any time \( t \in (0, \tau] \), there is a positive probability for a unit to be observed.

Let \( i_0 \) be the number of simple events in the sample space and \( p_1, \ldots, p_{i_0} \) be the corresponding probabilities. Then the expectation of the size of the risk set \( \delta(t) \) with \( n \) units is

\[
E[\delta(t)] = np(t),
\]

where \( p(t) = \sum_{j=1}^{i_0} p_j I(t \text{ is in the } j\text{th set of observation windows}) \) is the probability that a unit is observed at time \( t \). Here \( I(S) \) is the indicator function that equals 1 if the statement \( S \) is true and 0 otherwise. We have

\[
p(t) \geq \big( \min_{1 \leq j \leq i_0} p_j \big) \sum_{j=1}^{i_0} I(t \text{ is in the } j\text{th set of observation windows}) \geq \min_{1 \leq j \leq i_0} p_j \equiv p_0 > 0.
\]

Note that \( \delta_i(t) = \sum_{i=1}^{i_0} \delta_i(t) \) and \( \delta_1(t), \ldots, \delta_n(t) \) are independent and identically distributed (i.i.d.) Bernoulli random variables that take the value of 1 with probability \( p(t) \). Thus \( \delta_i(t) \) is a Binomial \((n, p(t))\) random variable.

Consider the superposition of all \( i_0 \) sets of observation windows in the sample space. Let \( h + 1 \) be the number of unique endpoints of all these observation-window intervals and denote the ordered unique endpoints by \( 0 = t_{1L} < t_{1U} < t_{2U} < \ldots < t_{hU} = \tau \). Let \( t_{iL} = t_{(i-1)U} \) for \( i = 2, \ldots, h \). Then, for \( t \in (t_{iL}, t_{iU}] \), \( i = 1, \ldots, h \),

\[
p(t) = p(t_{iU}) \equiv p_{ci} \quad \text{and} \quad \delta_i(t) = \delta_i(t_{iU}) \equiv \delta_i.
\]

That is, \( p(t) \) is constant and \( \delta_i(t) \) is a random variable not depending on \( t \) for all \( t \in (t_{iL}, t_{iU}] \), \( i = 1, \ldots, h \). Thus we have

\[
\frac{\delta_i(t)}{n} = \frac{\delta_i(t_{iU})}{n} \xrightarrow{a.s.} p_{ci} \quad \text{as } n \to \infty, \quad \text{for } t \in (t_{iL}, t_{iU}], \ i = 1, \ldots, h.
\]

We can link this condition to practical applications. For example, for the extended warranty data in Zuo, Meeker, and Wu (2008), we have, as \( n \to \infty \)

\[
\delta_i(t)/n \xrightarrow{a.s.} \begin{cases} 1 & \text{for } t \in (0, 1], \\ 0.5 & \text{for } t \in (1, 2], \\ 0.4 & \text{for } t \in (2, 3]. \\ \end{cases}
\]
3. Nonparametric MCF Estimator

3.1. Estimation for the Nonparametric Model

Let $m$ denote the number of unique event times, and let $t_1, \ldots, t_m$ be the unique event times. Zuo, Meeker, and Wu (2008) expressed the NP MCF estimator as

$$\hat{MCF}_{NP}(t_j) = \frac{j}{\sum_{k=1}^{m} \delta_i(t_k)} \sum_{k=1}^{m} d_i(t_k) = \sum_{k=1}^{m} \frac{d_i(t_k)}{\delta_i(t_k)}, \quad j = 1, \ldots, m,$$

where $d_i(t_k)$ is the number of events recorded at time $t_k$ for unit $i$.

3.2. Theorems Adapted from Andersen, Borgan, Gill, and Keiding (1993)

For the NP MCF estimator, the following two theorems from Andersen et al. (1993) are simplified and adapted with our notation. Note that “$\overset{P}{\rightharpoonup}$” is used to denote “converges in probability to” and “$\overset{D}{\rightharpoonup}$” is used to denote “converges in distribution to.” Also $\sup_{s \in [0, t]} f(s)$ and $\inf_{s \in [0, t]} f(s)$ stand for the supremum and infimum of $f(s)$ over $[0, t]$, respectively.

**Theorem IV.1.1** of Andersen et al. (1993, page 190). Let $t \in (0, \tau]$ and assume that, as $n \to \infty$,

$$\int_0^t \frac{I[\delta_1(s) > 0]}{\delta_1(s)} \nu(s) ds \overset{P}{\rightharpoonup} 0$$

and

$$\int_0^t \{1 - I[\delta_1(s) > 0]\} \nu(s) ds \overset{P}{\rightharpoonup} 0.$$  \hspace{1cm} (2)

Then, as $n \to \infty$,

$$\sup_{s \in [0, t]} |\hat{\mu}(s) - \mu(s)| \overset{P}{\rightharpoonup} 0,$$

where $\hat{\mu}(t)$ is the NP estimator of $\mu(t)$ defined as

$$\hat{\mu}(t) = \int_0^t \delta_1(s)^{-1} dN(s) = \hat{MCF}_{NP}(t).$$

This theorem can be used to show that the NP estimator of $\mu(t)$, based on window-observation recurrence data, is uniformly consistent on compact intervals.

**Theorem IV.1.2** of Andersen et al. (1993, page 191). Let $t \in (0, \tau]$ and assume that there exists a sequence of positive constants $\{a_n\}$, increasing to infinity as $n \to \infty$, and a positive function $y$ such that $\nu/y$ is integrable over $[0, t]$. Assume further that the following conditions hold.
NP(A) For each $s \in [0, t]$,

$$a_n^2 \int_0^s \frac{I[\delta(u) > 0]}{\delta(u)} \nu(u) du \xrightarrow{P} \sigma^2(s) \quad \text{as} \quad n \to \infty,$$

where

$$\sigma^2(s) = \int_0^s \frac{\nu(u)}{y(u)} du.$$

NP(B) For all $\epsilon > 0$,

$$a_n^2 \int_0^t I[\delta(u) > 0] \frac{I[\delta(u) > 0]}{\delta(u)} \nu(u) I \left\{ \frac{I[\delta(u) > 0]}{\delta(u)} > \epsilon \right\} du \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

NP(C)

$$a_n \int_0^t (1 - I[\delta(u) > 0]) \nu(u) du \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Then

$$a_n (\hat{\mu} - \mu) \xrightarrow{D} U \quad \text{as} \quad n \to \infty$$

on $D[0, t]$, where $U$ is a Gaussian martingale with $U(0) = 0$ and $\text{Cov}(U(s_1), U(s_2)) = \sigma^2(s_1 \wedge s_2)$.

Here $D[0, t]$ is the Skorohod space on $[0, t]$, that is, the space of right-continuous functions with left-hand limits on $[0, t]$; $s_1 \wedge s_2$ is the smaller of $s_1$ and $s_2$.

Theorem IV.1.2 can be used to establish the asymptotic normality of the NP estimator of $\mu(t)$ based on window-observation recurrence data.

3.3. Asymptotic Properties of the Nonparametric MCF Estimator

The following lemma will be useful for verifying the conditions in Section 3.2 for the asymptotic properties of the nonparametric MCF estimator.

**Lemma 1.** Let $\{X_n\}$ be a sequence of nonnegative random variables. If $E X_n \to 0$ as $n \to \infty$, then $X_n \xrightarrow{P} 0$ as $n \to \infty$.

The lemma follows by observing that, for any $\epsilon > 0$,

$$\Pr(X_n \geq \epsilon) \leq \frac{E X_n}{\epsilon} \rightarrow 0 \quad \text{as} \quad n \to \infty.$$
Theorem 1. Assume that $\mu(t) < \infty$ for $t \in (0, \tau]$. Then under the assumptions in Section 2.3, the NP MCF estimator $\hat{MCF}_{NP}(t) = \hat{\mu}(t)$ is uniformly consistent on compact intervals, that is, for any $t \in (0, \tau]$,

$$\sup_{s \in [0, t]} |\hat{\mu}(s) - \mu(s)| \xrightarrow{P} 0;$$

and $\hat{MCF}_{NP}(t)$ has the asymptotic distribution given in Theorem IV.1.2 of Andersen et al. (1993).

Proof. Assume that $\mu(t) < \infty$ for $t \in (0, \tau]$. By the assumptions in Section 2.3, for any $t \in (0, \tau]$ and $s \in [0, t]$, we have

$$E \left[ \frac{I[\delta(s) > 0]}{\delta(s)} \right] = \sum_{x=1}^{n} \frac{1}{x!(n-x)!} \frac{n!}{x!(n-1)!} p(s)^x \left[ 1 - p(s) \right]^{n-x}$$

$$= \sum_{x=1}^{n} \frac{(n+1)!}{(x+1)!(n+1) - (x+1)!} p(s)^{x+1} \left[ 1 - p(s) \right]^{(n+1)-(x+1)} \left( \frac{1}{n+1} \frac{x+1}{x} \frac{1}{p(s)} \right)$$

$$\leq \frac{1}{n+1} \frac{2}{p_0} \sum_{u=0}^{n+1} \frac{(n+1)!}{u!(n+1)-u!} p(s)^u \left[ 1 - p(s) \right]^{(n+1)-u}$$

$$\left( \text{because } \frac{x+1}{x} \leq 2 \text{ for } x \geq 1 \text{ and } \frac{1}{p(s)} \leq \frac{1}{p_0} \right)$$

$$= \frac{2}{p_0(n+1)}.$$

Thus

$$E \left[ \int_0^t \frac{I[\delta(s) > 0]}{\delta(s)} \nu(s) ds \right] = \int_0^t E \left[ \frac{I[\delta(s) > 0]}{\delta(s)} \right] \nu(s) ds$$

$$\leq \frac{2}{p_0(n+1)} \int_0^t \nu(s) ds$$

$$= \frac{2\mu(t)}{p_0} \frac{1}{n+1}$$

$$\rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Thus by Lemma 1, (1) is satisfied.

Similarly, to verify (2), we have
\[
\begin{align*}
E \left[ \int_0^t \{1 - I[\delta(s) > 0]\} \nu(s) ds \right] &= \int_0^t \{1 - \Pr[\delta(s) > 0]\} \nu(s) ds \\
&= \int_0^t \Pr[\delta(s) = 0] \nu(s) ds \\
&= \int_0^t [1 - p(s)]^n \nu(s) ds \\
&\leq \int_0^t [1 - p_0]^n \nu(s) ds \\
&= \mu(t)[1 - p_0]^n \\
&\to 0 \quad \text{as } n \to \infty
\end{align*}
\]

because \(0 < p_0 < 1\). Thus by Lemma 1, (2) is satisfied. Therefore, the uniform consistency of the
NP estimator of \(\mu(t)\) is established by Theorem IV.1.1 of Andersen et al. (1993).

To establish the asymptotic normality of the NP estimator of \(\mu(t)\), we need to show that conditions NP(A) to NP(C) are satisfied.

Let \(a_n = \sqrt{n}\) and \(y(t) = p(t)\), the probability that a unit is being observed at time \(t\). From the verification of (2), we have

\[
E \left[ \sqrt{n} \int_0^t \{1 - I[\delta(u) > 0]\} \nu(u) du \right] \leq \sqrt{n} \mu(t)[1 - p_0]^n \to 0 \quad \text{as } n \to \infty.
\]

Thus by Lemma 1, condition NP(C) is satisfied.

To verify conditions NP(A) and NP(B), first note that as \(n \to \infty\),

\[
\frac{\delta_i}{n} \to p_{ci} \quad \text{and} \quad I(\delta_i > 0) \to 1,
\]

where \(\delta_i = \delta_i(t_0)\), for \(i = 1, \ldots, h\). Thus

\[
\left| n \int_0^s \frac{I[\delta(u) > 0]}{\delta(u)} \nu(u) du - \sigma^2(s) \right| = \left| \int_0^s \left[ \frac{I[\delta(u) > 0]}{\delta(u)/n} - \frac{1}{p(u)} \right] \nu(u) du \right| \\
\leq \int_0^s \left[ \frac{I[\delta(u) > 0]}{\delta(u)/n} - \frac{1}{p(u)} \right] \nu(u) du \\
\leq \int_0^s \nu(u) du \max_{1 \leq i \leq h} \left| \frac{I[\delta_i > 0]}{\delta_i/n} - \frac{1}{p_{ci}} \right| \\
\to 0 \quad \text{as } n \to \infty.
\]

Therefore, condition NP(A) is satisfied.

Furthermore,
\[ n \int_0^t \frac{I[\delta_i > 0]}{\delta_i(u)} \nu(u) I \left\{ \sqrt{n} \frac{I[\delta_i(u) > 0]}{\delta_i(u)} > \epsilon \right\} du \]

\[ \leq \int_0^t \nu(u) du \max_{1 \leq i \leq h} \frac{I[\delta_i > 0]}{\delta_i/n} I \left\{ \max_{1 \leq i \leq h} \frac{I[\delta_i > 0]}{\delta_i/n} \frac{1}{\sqrt{n}} > \epsilon \right\} \]

\[ \xrightarrow{P} 0 \quad \text{as } n \to \infty. \]

Thus, condition NP(B) is satisfied. This concludes the proof of the asymptotic properties of the NP estimator of \( \mu(t) \).

4. Nonhomogeneous Poisson Process (NHPP) MCF Estimators

4.1. Estimation for the NHPP Model

For a parametric model, let \( \theta \) denote the model parameter, which is a \( q \)-dimensional real vector. Also let \( \theta_0 \) be the true value of \( \theta \), and \( \Theta_0 \) be a neighborhood of \( \theta_0 \).

A particular NHPP model is specified by its recurrence rate function \( \nu(t; \theta) \), and the expected number of events in the time range \( (a, b] \) is \( \mu(a, b; \theta) = \int_a^b \nu(t; \theta) dt \).

Given the maximum likelihood estimator (MLE), \( \hat{\theta} \), of \( \theta \), the MLE of the NHPP MCF \( \mu(t; \theta) \) is

\[ \hat{\text{MCF}}_{\text{NHPP}}(t) = \int_0^t \nu(s; \hat{\theta}) ds. \]  

(3)

The most commonly used NHPP recurrence rate functions are:

- **Constant recurrence rate**, also known as homogeneous Poisson process (HPP):

  \[ \nu(t) = c. \]

- **Power recurrence rate**, also known as the power law process or the Weibull process:

  \[ \nu(t; \beta, \eta) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1}, \quad \beta > 0, \eta > 0. \]

  Here, \( \theta = (\beta, \eta)' \). The power law NHPP model is flexible enough to model a counting process with either an increasing or a decreasing recurrence rate function. When the shape parameter \( \beta > 1 \), the recurrence function is increasing, and when \( \beta < 1 \), the recurrence rate function is decreasing. When \( \beta = 1 \), the power law NHPP model simplifies to the HPP model with constant recurrence rate \( 1/\eta \).

- **Loglinear recurrence rate**:

  \[ \nu(t; \gamma_0, \gamma_1) = \exp(\gamma_0 + \gamma_1 t). \]
Here, $\theta = (\gamma_0, \gamma_1)'$. The loglinear NHPP model assumes that the log recurrence rate function is a linear function of time, where $\gamma_0$ is the intercept and $\gamma_1$ is the slope. A positive value of $\gamma_1$ indicates an increasing recurrence rate, while a negative value of $\gamma_1$ indicates a decreasing recurrence rate. When $\gamma_1 = 0$, the loglinear NHPP model simplifies to the HPP model with recurrence rate $\exp(\gamma_0)$.

### 4.2. Conditions and Theorems Adapted from Andersen, Borgan, Gill, and Keiding (1993)

To establish the asymptotic properties of the NHPP MCF estimators, the following conditions and theorems from Andersen et al. (1993) are simplified and adapted with our notation.

From page 402 of Andersen et al. (1993), the log-partial-likelihood function for the parametric model is

$$
\ell_{\tau}(\theta) = \int_{\tau_0}^{\tau} \log \lambda(t; \theta)dN(t) - \int_{\tau_0}^{\tau} \lambda(t; \theta)dt,
$$

and, assuming that we may interchange the order of differentiation and integration, the vector $U_{\tau}(\theta)$ of score statistics $U_{j\tau}(\theta)$, $j = 1, ..., q$, is given by

$$
U_{j\tau}(\theta) = \int_{\tau_0}^{\tau} \frac{\partial}{\partial \theta_j} \log \lambda(t; \theta)dN(t) - \int_{\tau_0}^{\tau} \frac{\partial}{\partial \theta_j} \lambda(t; \theta)dt.
$$

Note that under the assumptions of Section 2.3, the partial likelihood is equivalent to the likelihood (up to a constant scalar not depending on $\theta$). Thus the partial likelihood equation $U_{\tau}(\theta) = (0, ..., 0)' \equiv 0$ is also the likelihood equation.

Condition VI.1.1 of Andersen et al. (1993, pages 420 and 421) can be simplified below for our model.

(A) There exists a neighborhood $\Theta_0$ of $\theta_0$ such that for all $n$ and $\theta \in \Theta_0$, and almost all $t \in (0, \tau]$, the partial derivatives of $\lambda(t; \theta)$ and log $\lambda(t; \theta)$ of the first, second, and third order with respect to $\theta$ exist and are continuous in $\theta$ for $\theta \in \Theta_0$. Moreover, the log-likelihood function (4) may be differentiated three times with respect to $\theta \in \Theta_0$ by interchanging the order of integration and differentiation.

(B) There exist a sequence $\{a_n\}$ of nonnegative constants increasing to infinity as $n \to \infty$ and finite functions $\sigma_{jl}(\theta)$ defined on $\Theta_0$ such that for all $j, l$

$$
a_n^{-2} \int_{\tau_0}^{\tau} \left\{ \frac{\partial}{\partial \theta_j} \log \lambda(t; \theta_0) \right\} \left\{ \frac{\partial}{\partial \theta_l} \log \lambda(t; \theta_0) \right\} \lambda(t; \theta_0)dt \xrightarrow{P} \sigma_{jl}(\theta_0),
$$

as $n \to \infty$. 
(C) For all $j$ and all $\epsilon > 0$, we have that

$$a_n^{-2} \int_0^\tau \left\{ \frac{\partial}{\partial \theta_j} \log \lambda(t; \theta_0) \right\}^2 I \left( \left| a_n^{-1} \frac{\partial}{\partial \theta_j} \log \lambda(t; \theta_0) \right| > \epsilon \right) \lambda(t; \theta_0) dt \overset{P}{\to} 0$$

as $n \to \infty$.

(D) The matrix $\Sigma = \{\sigma_{jl}(\theta_0)\}$ with $\sigma_{jl}(\theta_0)$ defined in condition (B) is positive definite.

(E) For any $n$ there exist predictable processes $G_n$ and $H_n$ not depending on $\theta$ such that for all $t \in (0, \tau]$

$$\sup_{\theta \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} \lambda(t; \theta) \right| \leq G_n(t),$$

and

$$\sup_{\theta \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} \log \lambda(t; \theta) \right| \leq H_n(t),$$

for all $j, l, m$. Moreover

$$a_n^{-2} \int_0^\tau G_n(t) dt,$$

$$a_n^{-2} \int_0^\tau H_n(t) \lambda(t; \theta_0) dt$$

as well as (for all $j, l$)

$$a_n^{-2} \int_0^\tau \left\{ \frac{\partial^2}{\partial \theta_j \partial \theta_l} \log \lambda(t; \theta_0) \right\}^2 \lambda(t; \theta_0) dt$$

all converge in probability to finite quantities as $n \to \infty$, and for all $\epsilon > 0$,

$$a_n^{-2} \int_0^\tau H_n(t)I \left\{ a_n^{-1} [H_n(t)]^{1/2} > \epsilon \right\} \lambda(t; \theta_0) dt \overset{P}{\to} 0.$$
\[ a_n(\hat{\theta} - \theta_0) \overset{p}{\rightarrow} N(0, \Sigma^{-1}), \]

where \( \Sigma = \{\sigma_{jl}(\theta_0)\} \) defined in condition (D) may be estimated consistently by \( a_n^{-2} J_\tau(\hat{\theta}) \). Here \(- J_\tau(\hat{\theta})\) is the matrix of the second-order partial derivatives of the log-likelihood function (4). That is, the \((j,l)\)th element of the matrix \( J_\tau(\theta) \) can be written as

\[
J_{jl}^{(2)}(\theta) = \int_0^\tau \frac{\partial^2}{\partial \theta_j \partial \theta_l} \lambda(t; \theta) dt - \int_0^\tau \frac{\partial^2}{\partial \theta_j \partial \theta_l} \log \lambda(t; \theta) dN(t).
\]

Theorem VI.1.1 establishes the existence of a consistent MLE of \( \theta \), and Theorem VI.1.2 establishes the asymptotic normality of the MLE.

4.3. Asymptotic Properties of the NHPP Estimators

**Theorem 2.** Under the following conditions (i) to (iv), the NHPP MCF estimators have the asymptotic properties as stated in Theorem VI.1.1 and Theorem VI.1.2 of Andersen et al. (1993).

(i) There exists a neighborhood \( \Theta_0 \) of \( \theta_0 \) such that for all \( n \) and \( \theta \in \Theta_0 \), and almost all \( t \in (0, \tau] \), the partial derivatives of \( \nu(t; \theta) \) and \( \log \nu(t; \theta) \) of the first, second, and third order with respect to \( \theta \) exist and are continuous in \( \theta \) for \( \theta \in \Theta_0 \). Moreover, the log-likelihood function (4) may be differentiated three times with respect to \( \theta \in \Theta_0 \) by interchanging the order of integration and differentiation.

(ii) \[
\int_0^\tau \left( \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right)^2 \nu(t; \theta_0) dt < \infty \quad \text{for all } j.
\]

(iii) \[
\int_0^\tau \left[ \sum_{j=1}^q c_j \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right]^2 \nu(t; \theta_0) dt > 0
\]

for any \( c = (c_1, ..., c_q)' \neq 0 \).

(iv) For all \( j,l \),

\[
\int_0^\tau \left( \frac{\partial^2}{\partial \theta_j \partial \theta_l} \log \nu(t; \theta_0) \right)^2 \nu(t; \theta_0) dt,
\]

\[
\int_0^\tau \Gamma(t; \Theta_0) dt, \tag{9}
\]
and

\[ \int_0^T \Delta(t, \Theta_0) \nu(t; \theta_0) dt \]

are finite quantities, where \( \Gamma(t, \Theta_0) \) and \( \Delta(t, \Theta_0) \) are defined as follows:

\[ \Gamma(t, \Theta_0) = \max_{1 \leq j, l, m \leq q} \sup_{\theta \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} \nu(t; \theta) \right|, \quad (10) \]

and

\[ \Delta(t, \Theta_0) = \max_{1 \leq j, l, m \leq q} \sup_{\theta \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} \log \nu(t; \theta) \right|. \quad (11) \]

**Proof.** First note that the consistency of \( \hat{M}_{CFNHP}(t) \) in (3) follows from that of \( \hat{\theta} \), provided that \( \mu(t; \theta) \) is continuous in \( \theta \) for \( \theta \in \Theta_0 \). Furthermore, the asymptotic normality of \( \hat{M}_{CFNHP}(t) \) follows from that of \( \hat{\theta} \) by the \( \delta \)-method, provided that \( \mu(t; \theta) \) is differentiable with respect to \( \theta \) at \( \theta_0 \) (Billingsley 1986, page 402). To establish the consistency and asymptotic normality of \( \hat{\theta} \) for the NHPP models, we need to show that conditions (A) to (E) are satisfied.

By the assumptions in Section 2.3, \( \delta(t) = \delta(t_iU) \equiv \delta_i \) for \( t \in (t_{iL}, t_{iU}] \equiv J_i \), for \( i = 1, \ldots, h \), and \( \Pr[\delta(t) > 0] \to 1 \) as \( n \to \infty \). Thus it suffices to consider \( \delta(t) > 0 \). Note that for the multiplicative intensity function, we have \( \lambda(t; \theta) = \nu(t; \theta) \delta(t) \). Thus condition (A) can be restated as condition (i).

For conditions (B) to (E), let \( a_n = \sqrt{n} \). Then the left side of (5) can be written as

\[
 n^{-1} \int_0^T \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\} \left\{ \frac{\partial}{\partial \theta_l} \log \nu(t; \theta_0) \right\} \nu(t; \theta_0) \delta(t) dt
 = \sum_{i=1}^h \frac{\delta_i}{n} \int_{J_i} \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\} \left\{ \frac{\partial}{\partial \theta_l} \log \nu(t; \theta_0) \right\} \nu(t; \theta_0) dt
 \xrightarrow{P} \sum_{i=1}^h p_{ci} \sigma_{ijl} \equiv \sigma_{ijl}(\theta_0) \quad \text{as } n \to \infty,
\]

where

\[ \sigma_{ijl} = \int_{J_i} \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\} \left\{ \frac{\partial}{\partial \theta_l} \log \nu(t; \theta_0) \right\} \nu(t; \theta_0) dt. \]

That is, condition (B) follows from condition (ii).

Similarly, the left side of (6) can be written as

\[
 \sum_{i=1}^h \frac{\delta_i}{n} \int_{J_i} \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\}^2 I \left( \left| \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right| > \epsilon \right) \nu(t; \theta_0) dt
 \equiv \sigma_{ijl}(\theta_0) \quad \text{as } n \to \infty,
\]

where

\[ \sigma_{ijl} = \int_{J_i} \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\}^2 I \left( \left| \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right| > \epsilon \right) \nu(t; \theta_0) dt. \]

That is, condition (D) follows from condition (iii).
and (12) \( P \to 0 \) as \( n \to \infty \) provided that condition (ii) is true, because

\[
\frac{\delta_i}{n} P \to p_{ci}
\]

and

\[
\int_{J_i} \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\}^2 I \left( \left| \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right| > \epsilon \right) \nu(t; \theta_0) dt \\
\leq \int_0^\tau \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\}^2 I \left( \left\{ \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right\}^2 > n \epsilon^2 \right) \nu(t; \theta_0) dt \\
\to 0 \text{ by (ii).}
\]

That is, condition (ii) also implies condition (C).

To verify condition (D), it suffices to show that, for any constant vector \( c = (c_1, ..., c_q)' \neq 0 \), \( c' \Sigma c > 0 \). We have

\[
c' \Sigma c = \sum_{i=1}^h p_{ci} \sum_{j,l=1}^q c_j c_l \sigma_{ijl} \\
= \sum_{i=1}^h p_{ci} \int_{J_i} \left[ \sum_{j=1}^q c_j \frac{\partial}{\partial \theta_j} \log \nu(t; \theta_0) \right]^2 \nu(t; \theta_0) dt \\
> 0,
\]

which holds because of condition (iii).

To verify condition (E), first note that (7) can be written as

\[
\sum_{i=1}^h \frac{\delta_i}{n} \int_{J_i} \left\{ \frac{\partial^2}{\partial \theta_j \partial \theta_l} \log \nu(t; \theta_0) \right\}^2 \nu(t; \theta_0) dt \to \sum_{i=1}^h p_{ci} \int_{J_i} \left\{ \frac{\partial^2}{\partial \theta_j \partial \theta_l} \log \nu(t; \theta_0) \right\}^2 \nu(t; \theta_0) dt
\]
as \( n \to \infty \). Thus (7) can be restated as

\[
\int_0^\tau \left\{ \frac{\partial^2}{\partial \theta_j \partial \theta_l} \log \nu(t; \theta_0) \right\}^2 \nu(t; \theta_0) dt < \infty.
\]

Now let \( \Gamma(t, \Theta_0) \) and \( \Delta(t, \Theta_0) \) be as defined in (10) and (11), respectively. Then

\[
\sup_{\theta \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} \lambda(t; \theta) \right| \leq \Gamma(t, \Theta_0) \delta(t) \equiv G_n(t)
\]
and
\[
\sup_{\theta \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} \log \lambda(t; \theta) \right| \leq \Delta(t, \Theta_0) \equiv H_n(t).
\]

Thus
\[
a_n^{-2} \int_0^T G_n(t) dt = \sum_{i=1}^h \frac{\delta_i}{n} \int_{J_i} \Gamma(t, \Theta_0) dt \quad \overset{P}{\to} \quad \sum_{i=1}^h p_{ci} \int_{J_i} \Gamma(t, \Theta_0) dt < \infty \quad \text{as } n \to \infty,
\]

provided that
\[
\int_0^T \Gamma(t, \Theta_0) dt < \infty.
\]

Also
\[
a_n^{-2} \int_0^T H_n(t) \lambda(t; \theta_0) dt = \sum_{i=1}^h \frac{\delta_i}{n} \int_{J_i} \Delta(t, \Theta_0) \nu(t; \theta_0) dt \quad \overset{P}{\to} \quad \sum_{i=1}^h p_{ci} \int_{J_i} \Delta(t, \Theta_0) \nu(t; \theta_0) dt < \infty \quad \text{as } n \to \infty,
\]

provided that
\[
\int_0^T \Delta(t, \Theta_0) \nu(t; \theta_0) dt < \infty.
\]

Furthermore, for all \( \epsilon > 0 \), the left side of (8) can be written as
\[
\sum_{i=1}^h \frac{\delta_i}{n} \int_{J_i} \Delta(t, \Theta_0) I \left\{ \frac{1}{\sqrt{n}} \left| \Delta(t, \Theta_0) \right|^{1/2} > \epsilon \right\} \nu(t; \theta_0) dt,
\]

and (13) \( \overset{P}{\to} 0 \) as \( n \to \infty \), because \( \delta_i/n \overset{P}{\to} p_{ci} \) and
\[
\int_{J_i} \Delta(t, \Theta_0) I \left\{ \Delta(t, \Theta_0) > n\epsilon^2 \right\} \nu(t; \theta_0) dt \to 0
\]
as \( n \to \infty \), provided that \( \int_0^T \Delta(t, \Theta_0) \nu(t; \theta_0) dt < \infty \).

Thus condition (E) follows from condition (iv).
4.3.1. Asymptotic Properties of the Loglinear NHPP Estimator

Corollary 1. The loglinear NHPP estimator has the asymptotic properties stated in Theorem 2.

Proof. For the loglinear NHPP model, \( \nu(t; \gamma_0, \gamma_1) = \exp(\gamma_0 + \gamma_1 t) \), and log \( \nu(t; \gamma_0, \gamma_1) = \gamma_0 + \gamma_1 t \).

Let \( \theta_0 = \left( \gamma_0^{(0)}, \gamma_1^{(0)} \right)' \) be the true value of \( \theta = (\gamma_0, \gamma_1)' \). Note that for any nonnegative integers \( \alpha \) and \( \beta \),

\[
\frac{\partial^{\alpha+\beta}}{\partial \gamma_0^\alpha \partial \gamma_1^\beta} \nu(t; \gamma_0, \gamma_1) = t^\beta \nu(t; \gamma_0, \gamma_1),
\]

\[
\frac{\partial}{\partial \gamma_0} \log \nu(t; \gamma_0, \gamma_1) = 1,
\]

\[
\frac{\partial}{\partial \gamma_1} \log \nu(t; \gamma_0, \gamma_1) = t,
\]

and

\[
\frac{\partial^{\alpha+\beta}}{\partial \gamma_0^\alpha \partial \gamma_1^\beta} \log \nu(t; \gamma_0, \gamma_1) = 0 \quad \text{for } \alpha + \beta > 1. \tag{14}
\]

These are all continuous functions in both \( (\gamma_0, \gamma_1)' \) and in \( t \in [0, \tau] \). Thus, conditions (i) and (ii) are satisfied.

To verify condition (iii), note that for any \( c = (c_0, c_1)' \neq 0 \),

\[
\int_0^\tau \left[ c_0 \frac{\partial}{\partial \gamma_0} \log \nu(t; \theta_0) + c_1 \frac{\partial}{\partial \gamma_1} \log \nu(t; \theta_0) \right]^2 \nu(t; \theta_0) dt = \int_0^\tau (c_0 + c_1 t)^2 \exp \left( \gamma_0^{(0)} + \gamma_1^{(0)} t \right) dt > 0.
\]

Thus condition (iii) is satisfied.

To verify condition (iv), because of (14), it suffices to show that (9) < \( \infty \). This is again straightforward because

\[
\Gamma(t, \Theta_0) \leq (1 + t)^3 \exp \left[ \left( \frac{\gamma_0^{(0)}}{\gamma_0} + 1 \right) + \left( \frac{\gamma_1^{(0)}}{\gamma_1} + 1 \right) t \right],
\]

where \( \Theta_0 = \{(\gamma_0, \gamma_1)' : \left| \gamma_0 - \gamma_0^{(0)} \right| < 1, \left| \gamma_1 - \gamma_1^{(0)} \right| < 1 \}. \)
4.3.2. Asymptotic Properties of the Power Law NHPP Estimator

**Corollary 2.** The power law NHPP estimator has the asymptotic properties stated in Theorem 2.

**Proof.** To verify conditions (i), (ii), (iii), and (iv) for the power law NHPP model, we note that

\[ \nu(t; \beta, \eta) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta - 1} \]

and

\[ \log \nu(t; \beta, \eta) = \log(\beta) - \beta \log(\eta) + (\beta - 1) \log(t) \]

are smooth functions of \((t, \beta, \eta)\) for \(t, \beta, \eta > 0\) and their partial derivatives of any order exist and are continuous. Thus condition (i) is satisfied. Furthermore, the partial derivatives with respect to \(\beta\) and \(\eta\) are polynomial functions of \(t^{\beta - 1}\) and \(\log t\) that are linear in \(t^{\beta - 1}\). Thus to verify conditions (ii), (iii), and (iv), we only need to show that for any \(0 < a < 1\), and \(j \geq 0\),

\[ \int_0^a t^{\beta - 1} |\log t|^j dt < \infty, \]

which is indeed true because \(\int_0^a t^{\beta - 1} |\log t|^j dt = \int_{\log a}^{\infty} u^j \exp(-\beta u) du < \infty. \)

5. Concluding Remarks

In this article, for the one-dimension counting process, we have established the asymptotic properties for both the nonparametric and the NHPP estimators of the MCF from window-observation recurrence data. With the assumptions in Section 2.3, the NP estimator of the MCF is shown to be uniformly consistent on compact intervals, and the NHPP estimators of the MCF have consistent solutions to the likelihood equations as long as conditions (i), (ii), (iii), and (iv) in Section 4.3 are satisfied. One advantage of these restated conditions is that they are on the recurrence rate function \(\nu(t; \theta)\) of the NHPP model instead of the multiplicative intensity model \(\lambda(\theta; t) = \nu(\theta; t) \delta(t)\), and illustrations with the loglinear and the power law NHPP models are given in Sections 4.3.1 and 4.3.2. The asymptotic normality for the NP and the NHPP MCF estimators is also established when the assumptions and the conditions (needed for the NHPP MCF estimators) are satisfied.

The results in this article for univariate MCF estimators, with those conditions and theorems from Andersen et al. (1993), can be extended to the multiple-dimension counting process. Examples of multiple-dimension counting processes include numbers of recurrences for multiple failure modes from automobile repair history.
REFERENCES


