ASYMPTOTIC PROPERTIES OF MAXIMUM LIKELIHOOD ESTIMATORS FOR SPATIAL NONHOMOGENEOUS POISSON PROCESS MODELS

Han Wu\(^1\)\(^8\), Mark S. Kaiser\(^2\), Dewi Rahardja\(^3\), Yan D. Zhao\(^4\)

\(^1\)Department of Mathematics and Statistics
Minnesota State University
Mankato, MN 56001, USA
e-mail: han.wu@mnsu.edu

\(^2\)Department of Statistics
Iowa State University
Ames, IA 50011, USA
e-mail: mskaiser@iastate.edu

\(^3\), \(^4\)Department of Clinical Sciences and Simmons Cancer Center
UT Southwestern Medical Center
Dallas, TX 75390-8822, USA
\(^3\)e-mail: rahardja@gmail.com
\(^4\)e-mail: yandzhao@gmail.com

Abstract: Techniques have been developed for estimating the parameters of spatial point processes, given data at either the aggregate or point levels. However, it remains unclear how to model aggregate data with a subset of point data (i.e., exact locations of some events). For a sample region \(A \subset \mathbb{R}^d\), Wu and Kaiser [26] propose an aggregate-point combined model for a mixture of an aggregate and point data to accommodate both aggregate level and point level information. In this paper, we show that the maximum likelihood estimator is consistent, asymptotically normal, and asymptotically efficient as \(A \uparrow \mathbb{R}^d\). These results extend the findings of Rathbun and Cressie [20], where they study the asymptotic properties based on point data.

AMS Subject Classification: 62M09, 60G55, 62E20, 62F12

Key Words: aggregate-point combined model, asymptotic properties, maximum likelihood estimator, nonhomogeneous Poisson process

Received: March 10, 2011 © 2011 Academic Publications

\(^8\)Correspondence author
In the context of spatial observations, there are two types of commonly used asymptotic structures. In the first case, the observed region of the spatial process of interest is bounded, and more and more observations are taken from within this given region. The minimum distance between the data sites approaches zero as the sample size goes to infinity. Following Cressie [4], we call this “infill asymptotics”. This structure is common in mining, soil science and other geostatistical applications where information is sampled from a fixed, bounded region. No general results are available for maximum likelihood estimators under infill asymptotics. Lahiri [15] shows that under some fairly mild conditions, the least squares estimators of a spatial regression parameter vector and the method of moments estimator of the variogram of \( x(\cdot) \) at a point converge to non degenerate limiting random vectors, and are therefore inconsistent for the underlying parameters. In the second asymptotic setting, we observe a stochastic process at an increasing number of sites with fixed minimum distance, while the region becomes unbounded as the sample size goes to infinity. This setting is known as expanding domain asymptotics.

Ibragimov and Has’minskii [13] have verified that under appropriate regularity conditions, for any family of statistical experiments, the maximum likelihood estimator is consistent, asymptotically normal and asymptotically efficient. Kutoyants [14] shows that for one dimension \([0, T] \subset \mathbb{R}\) if several conditions are satisfied, the point model allows the same results as the more general case of Ibragimov and Has’minskii [13]. Rathbun and Cressie [20] show that the results of Kutoyants [14] are also true for spatial domain \( \mathbb{R}^d \), where \( d \geq 1 \).

However, for our aggregate-point combined model, conditions similar to those used by Kutoyants [14] are not readily available. In our case, we don’t know all the locations, besides the underlying point process \( N \), there is a stochastic generating mechanism \( M \) involved that is location related. \( M \leq N \) and \( M \) is a counting measure.

The goal of this paper is to study asymptotic properties of maximum likelihood estimator for the proposed combined model. We show that the maximum likelihood estimator is consistent, asymptotically normal, and asymptotically efficient as the sample region \( A \uparrow R^d \). We use the similar notations and problem setting used by Rathbun and Cressie [20].
2. Preliminaries

Let $|u| = (u'u)^{1/2}$ denote the norm of the vector $u$ and let $|H| = \sup\{|Hu| : |u| = 1\}$ denote the norm of the matrix $H$. Graybill [8] has provided the detailed properties of matrix norm.

For a deterministic function $f$, mapping $A$ onto $\mathbb{R}$, and Lebesgue measure $\nu$, we define its norm in the linear space $L_2(A)$ (We say that $f \in L_2(A)$ if $\|f\|_A < \infty$. as,

$$\|f\|_A = \left(\int_A (f(s))^2 \nu(ds)\right)^{1/2}.$$  

A function $f(s; y\theta)$ is absolutely continuous with respect to $y \in [a, b]$ at the point $s \in \mathbb{R}^d$, if for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$\sum_{i=1}^m |f(s; \alpha_i \theta) - f(s; \beta_i \theta)| < \epsilon,$$

for any finite collection $\{(\alpha_i, \beta_i) : i = 1, \ldots, m\}$ of non-overlapping intervals in $[a, b]$ with

$$\sum_{i=1}^m |\alpha_i \theta - \beta_i \theta| < \delta.$$

For a nonhomogeneous Poisson process $N$ with intensity function $\lambda(\cdot; \theta)$, define the random signed-measure,

$$Q_\theta(A) = N(A) - \int_A \lambda(s; \theta)\nu(ds); A \in \mathcal{A}.$$  

Let $\xi$ be any deterministic Lebesgue measurable function such that,

$$\int_A \xi(s)\lambda(s; \theta)\nu(ds) < \infty,$$

(1)

for any bounded set $A \in \mathcal{A}$. For each realization of $Q_\theta$, define the stochastic integral,

$$J_\theta(\xi; A) = \int_A \xi(s)Q_\theta(ds); A \in \mathcal{A},$$

where $\xi$ satisfies (1). We will see in Section 4 that $J_\theta(\xi; A)$ is well-defined and measurable.
3. Local Asymptotic Normality and Asymptotic Properties of Maximum Likelihood Estimators

In a number of interesting papers of Hájek ([9], [10] and [11]), LeCam([16], [17] and [18]), and other authors, it was provided that many important properties of statistical estimators follow from the asymptotic normality of the logarithm of the likelihood ratio for neighborhood hypothesis (for values of parameters close to each other) regardless of the relation between the observations which produced the given likelihood function. The main purpose of this section is to provide the asymptotic properties of maximum likelihood estimator adapted from Ibragimov and Has’minskii [13]. Ibragimov and Has’minskii [13] consider a family of statistical experiments $\mathcal{E}_\epsilon = (\mathcal{X}(\epsilon), \mathcal{U}(\epsilon), \{P^\epsilon_\theta : \theta \in \Theta\})$ generated by observations $X^\epsilon$. They study the asymptotic behavior of maximum likelihood estimator of $\theta$, as $\epsilon \to 0$. In our case, the family of experiments is a sequence $\{(M_j, N_j) : j = 1, 2, \ldots\}$, where $N_j = N(A_j)$ and $M_j = M(A_j)$ for $j \geq 1$. $N$ is a nonhomogeneous Poisson process and $M$ is a count generating mechanism with support $[0,N]$. We study the behavior of the maximum likelihood estimator of $\theta$ as $j \to \infty$.

The point process $N$ and $N$-related $M$ together induce a probability measure on the space of locally-finite counting measures $(\Phi, \mathcal{N})$ on $\mathbb{R}^d$, which belongs to $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$. Here $\mathcal{N}$ is the smallest $\sigma$-algebra generated by sets of the form $\{(M, N) \in \Phi : N(A) = n \& M(A) \leq n\}$, for all $A \in \mathcal{A}$ and all $n \in \{1,2,\ldots\}$, where $\mathcal{A}$ denote the Borel sets in $\mathbb{R}^d$. Consider a nested sequence of bounded Borel sets $\{A_j \in \mathcal{A} : j = 1,2,\ldots\}$ such that $\nu(A_j) \to \infty$. Let $\Phi_j$ be a collection of locally-finite counting measures on $A_j$. Define $A_j = A_j \cap \mathcal{A}$ to be the trace $\sigma$-algebra of $\mathcal{A}$ on $A_j$, and define $P^\epsilon_{\theta_j} = P_{\theta_j(M_j,N_j)}$ to be the restriction of $P_{\theta_j}$ to $N_j$, where $N_j = \Phi_j \cap \mathcal{N}$ is the trace $\sigma$-algebra of $\mathcal{N}$ on $\Phi_n$. So $N_j$ is a nonhomogeneous Poisson process on $A_j$ with intensity $\lambda(s;\theta)$ and $M_j$ is a restricted count generating mechanism on $A_j$. Note that $\mathcal{N}_1 \subset \mathcal{N}_2 \subset \cdots \subset \mathcal{N}$, i.e., $\{N_j : j = 1,2,\ldots\}$ is a nested sequence of $\sigma$-algebras. Let $E^{(j)}_{\theta_j}$ denote the mathematical expectation operator with respect to $P^\epsilon_{\theta_j}$.

We now study the asymptotic properties of an estimator $\hat{\theta}_{A_j}$ of $\theta$, as $j \to \infty$. The family of estimators $\{\hat{\theta}_{A_j} : j = 1,2,\ldots\}$ is consistent for $\theta$ if $\forall \epsilon > 0$,

$$\lim_{j \to \infty} P^\epsilon_{\theta_j}\{|\hat{\theta}_{A_j} - \theta| > \epsilon\} = 0.$$ 

For a typical family of probability measures $\{P^\epsilon_{\theta_j} : \theta \in \Theta\}$, Ibragimov and Has’minskii [13] give conditions under which for all $\theta \in K$, a compact set,
maximum likelihood estimator $\hat{\theta}_j$ is consistent and asymptotically normal, as $j \to \infty$. The conditions for the asymptotic properties of the point model include the requirement that the family of probability measures $\{P^A_j : \theta \in \Theta\}$ is locally asymptotically normal as $j \to \infty$, uniformly for all $\theta \in K$; this may be obtained if the log likelihood ratio can be approximated by a Gaussian random variable (Wald [24]). For any $\theta_1, \theta_2 \in \Theta$, the likelihood ratio is defined as

$$\Lambda_{A_j}(\theta_1, \theta_2) = \frac{L_{A_j}(\theta_1; N_j, M_j)}{L_{A_j}(\theta_2; N_j, M_j)}.$$ 

where $L_{A_j}$ is the likelihood function. Now, we list the results of Ibragimov and Has’minskii [13] as follows.

**Definition 1.** A family $\{P^A_j : \theta \in \Theta\}$ is called locally asymptotically normal (LAN) at point $\theta \in \Theta \subset \mathbb{R}^k$, as $j \to \infty$, if for some nonsingular $k \times k$ normalizing matrix $\phi_{A_j}(\theta)$ and any $u \in \mathbb{R}^k$, the likelihood ratio, $Z_{A_j}(u) = \Lambda_{A_j}(\theta + \phi_{A_j}(\theta)u, \theta)$ can be written as,

$$Z_{A_j}(u) = \exp\{u'\Delta_{A_j}(\theta) - \frac{1}{2}u'u + f_{A_j}(u, \theta)\}, \quad (2)$$

where

$$\Delta_{A_j}(\theta) \xrightarrow{L} N(0, I_k), \text{ as } j \to \infty, \quad (3)$$

and

$$f_{A_j}(u, \theta) \xrightarrow{p} 0, \text{ as } j \to \infty. \quad (4)$$

Note that $\xrightarrow{L}$ denotes convergence in distribution, $\xrightarrow{p}$ denotes convergence in probability, and $I_k$ denotes $k \times k$ identity matrix. Typically, $\phi_{A_j}(\theta) = (I_{A_j}(\theta))^{-\frac{1}{2}}$, where $I_{A_j}(\theta)$ is the Fisher information matrix.

**Definition 2.** A family $\{P^A_j : \theta \in \Theta\}$ is called uniformly asymptotically normal (UAN) as $j \to \infty$, if (2)–(4) are satisfied uniformly for all $\theta \in K$, where $K$ is an arbitrary compact subset of $\Theta$. This is simply a uniform version of LAN.

In what follows we assume that $\Theta$ is an open subset of $\mathbb{R}^k$. Let $K$ be an arbitrary compact subset of $\Theta$, and define $U_j = \{u : \theta + \phi_{A_j}(\theta)u \in \Theta\}$. We state the conditions under which the properties of estimators will be investigated below,

(i) The family $\{P^A_j : \theta \in \Theta\}$ is uniformly asymptotically normal (UAN) as $j \to \infty$ for all $\theta$ in $K \subset \Theta$. (Please note this $K$ is an arbitrary one.)

(ii) For any compact set $K \subset \Theta$, $\limsup_{j \to \infty} \sup_{\theta \in K} \left| \phi_{A_j}(\theta)\phi_{A_j}(\theta)' \right| = 0.$
(iii) For any compact set \( K \subset \Theta \), \( \exists \) some \( \beta > 0 \), \( m > 0 \), \( B = B(K) \), and \( a = a(K) \), and for all \( R > 0 \), \( \sup \{ E^{(j)}_\theta | Z_{A_j}(u) |^{1/m} - Z_{A_j}(v) |^{1/m} \}^m : \theta \in K; \ u, v \in U_j; \ |u| < R; \ |v| < R \} < B(1 + R^a) |u - v|^\beta \).

(iv) For any compact set \( K \subset \Theta \) and any integer \( j > 0 \), there exists \( j_0 = j_0(j, K) \) such that, \( \sup_{\theta \in K} \sup_{j > j_0} u \in U_j \) \( \sup \{ u^{\beta} E^{(j)}_\theta \} \{(Z_{A_j}(u))^{1/2} \} < \infty \).

The asymptotic behavior of the maximum likelihood estimator is provided by Ibragimov and Has’minskii \[14\] in the following theorem,

**Theorem 1.** (Ibragimov and Has’minskii \[13\]) Let \( \Theta \subset \mathbb{R}^k \) and let the conditions (i) – (iv) be satisfied with \( \beta > k \) in condition (iii). Let \( K \) be an arbitrary compact set in \( \Theta \). Then for all \( \theta \in K \):

1. the maximum likelihood estimator \( \hat{\theta}_{A_j} \) is consistent for \( \theta \), as \( j \to \infty \);
2. \( (\phi_{A_j}(\theta))^{-1}(\hat{\theta}_{A_j} - \theta) \xrightarrow{\mathcal{L}} N(0, I_k) \), as \( j \to \infty \), where \( I_k \) is the \( k \times k \) identity matrix;
3. \( \lim_{j \to \infty} E^{(j)}_\theta \left| (\phi_{A_j}(\theta))^{-1}(\hat{\theta}_{A_j} - \theta) \right|^r = E |\xi|^r \); \( r = 1, 2, \ldots \), where \( \xi \sim N(0, I_k) \) and \( I_k \) is the \( k \times k \) identity matrix.

Kutoyants \[14\] shows that the condition (iv) can be modified as:

(iv)' For any compact set \( K \subset \Theta \), there exist constants \( \gamma > 0 \) and \( c > 0 \) such that \( \sup_{\theta \in K} P^j_\theta \{ Z_{A_j}(u) > \exp(-c |u|^\gamma) \} \leq \exp(-c |u|^\gamma) \).

Asymptotic efficiency is defined as follows,

**Definition 3.** (see Ibragimov and Has’minskii \[13\], Chapter I) A family of estimators \( \hat{\theta}_{A_j} \) is called \( w \)-asymptotically efficient for \( \theta \) in \( K \subset \Theta \) (asymptotically efficient with respect to the family of loss functions \( w \)) if for any nonempty open set \( U \subset K \), the relation \( \lim_{j \to \infty} \left[ \inf_{T_j} \sup_{\theta \in U} E^{(j)}_\theta \{ w\{(\phi_{A_j}(\theta))^{-1}(T_j - \theta)\} \} \right] \) is satisfied.

The asymptotic efficiency of an estimator \( \hat{\theta}_{A_j} \) of \( \theta \) can be shown by the following theorem provided by Hájek \[11\].

**Theorem 2.** Let the family \( P^j_\theta \) satisfy the LAN conditions at the point \( \theta_0 \) with the normalizing matrix \( \phi_{A_j}(\theta) \) and let \( \lim_{j \to \infty} \sup_{\theta \in K} |\phi_{A_j}(\theta)\phi_{A_j}(\theta)'| = 0 \).
Suppose that \( w \in W_\epsilon \), where \( W_\epsilon \) denotes the class of loss functions such that for any \( w \in W_\epsilon \), the growth of \( w(x) \), as \( |x| \to \infty \), is slower than the growth of \( \exp\{ \epsilon |x|^2 \} \) as \( |x| \to \infty \), for all \( \epsilon > 0 \). Then for any family of estimators \( \theta_\epsilon \) of \( \theta \) and any \( \epsilon > 0 \),

\[
\liminf_{j \to \infty} \sup_{|\theta - \theta_0| < \delta} E^{(j)}_{\theta} \left[ w\{ (\phi_{A_j}(\theta))^{-1}(\theta^*_A - \theta) \} \right] \geq E\{ w(\xi) \},
\]

where \( \xi \sim N(0, I_k) \) and \( I_k \) is the \( k \times k \) identity matrix.

**Proof.** See Hájek [11] and Ibragimov and Has’minskii [12].

Theorem 1 shows that the moment of \( (\phi_{A_j}(\theta))^{-1}(\theta^*_A - \theta) \) converges to the moment of \( \xi \) as \( j \to \infty \), where \( \xi \sim N(0, I_k) \). So Theorem 2 shows that \( \hat{\theta}_A \) are asymptotically efficient for \( \theta \) with respective to loss functions of form \( w(x) = |x|^r; \ r = 1, 2, \cdots \).


In this section, we follow Ibragimov and Has’minskii’s theorem to obtain proofs of asymptotic properties of maximum likelihood estimators for spatial nonhomogeneous Poisson point process based on combined model. Suppose that \( N_j \) is a nonhomogeneous point process on \( A_j \) with intensity \( \lambda(s; \theta) \), where \( \theta \in \Theta \), an open convex subset of \( \mathbb{R}^k \). \( A_j \subset A \) is a nested sequence of Borel sets such that \( \int_{A_j} \lambda(s; \theta) \nu(ds) < \infty \) for all \( j \geq 1 \) and \( \nu(A_j) \uparrow \infty \) as \( j \to \infty \), where \( \nu(\cdot) \) is the Lebesgue measure. \( M_j \) is a count generating mechanism with support \( [0, N_j] \) and it represents the known exact locations among total \( N_j \). Let \( K \) be any compact subset of \( \Theta \) and let \( \psi(s; \theta) = 2(\lambda(s; \theta))^{1/2} \). This section considers the asymptotic properties of the maximum likelihood estimator \( \hat{\theta}_j \), as \( j \to \infty \).

Let’s consider the following assumptions first:

1. For all \( \theta \in \Theta \) and all \( s \in \mathbb{R}^d \), \( \lambda(s; \theta) > 0 \). The function \( \lambda(s; \theta + yu) \) is absolutely continuous with respective to \( y \in [0, 1] \), for all \( \theta \in \Theta \), all \( u \in \mathbb{R}^k \), and for almost all \( s \in \mathbb{R}^d \). The derivative \( \psi(s; \theta) = \hat{\lambda}(s; \theta)/(\lambda(s; \theta))^{1/2} \in L_2(A) \), for all \( \theta \in \Theta \) and all \( A \subset \mathbb{R}^d \).

2. For every \( \theta \in \Theta \), \( \lim_{j \to \infty} \left| \phi_{A_j}(\theta)\phi_{A_j}(\theta)^t \right| = 0 \), and there exists a constant \( c_0 \) such that

\[
\sup_{\theta_1, \theta_2 \in \Theta} \left| \phi_{A_j}(\theta_1) \left( \int_{A_j} \frac{\hat{\lambda}(s; \theta_2)/(\lambda(s; \theta_2))}{\lambda(s; \theta_1)} \nu(ds) \right)^{1/2} \right| \leq c_0;
\]

\[ j = 1, 2, \cdots \], where \( | \cdot | \) denotes the matrix norm defined in Section 2.
There exists a measure $\mu$ on $\mathbb{R}^d$ such that $\lim_{j \to \infty} \mu(A_j) = \infty$, and

$$\lim_{j \to \infty} \sup \{ \left\| \phi_{A_j}(\theta)'(\psi(\theta_u) - \psi(\theta)) \right\|_{A_j} : \theta \in K, u \in U_j, |u| < \mu(A_j) \} = 0,$$

where $\theta_u = \theta + \phi_{A_j}(\theta)u$ and $U_j = \{ u : \theta_u \in \Theta \}$.

For the measure $\mu$,

$$\lim_{j \to \infty} \sup \{ \int_{A_j} \left| \phi_{A_j}(\theta)' \frac{\lambda(s;\theta_1)}{\lambda(s;\theta_2)} \right|^4 \lambda(s;\theta_v) : \theta \in K, u, v \in U_j, |u| + |v| < \mu(A_j) \} = 0.$$

For some $\alpha > 0$,

$$\lim_{j \to \infty} \inf \{ |\phi_{A_j}(\theta)|^\alpha |\psi(\theta_u) - \psi(\theta)|_{A_j} : \theta \in K, u \in U_j, |u| > \mu(A_j) \} > 0.$$

For some $m > k/2$,

$$\sup \{ \int_{A_j} \left| \phi_{A_j}(\theta)' \frac{\lambda(s;\theta_1)}{\lambda(s;\theta_2)} \right|^{2m} \lambda(s;\theta_2)v(ds) : \theta_1, \theta_2 \in K, j \geq 1 \} < \infty.$$

For some $\delta > 0$,

$$\sup_{\theta^* \in (\theta - \delta, \theta + \delta)} \{ \int_{A_j} |\lambda(s;\theta^*)| v(ds) : j \geq 1 \} < \infty,$$

thus, with condition $(C_2)$, we have

$$\lim_{j \to \infty} \int_{A_j} (\lambda(s;\theta + \phi_{A_j}(\theta)u) - \lambda(s;\theta)) v(ds) = 0.$$

Conditions $(C_1)$, $(C_3)$, $(C_4)$, $(C_5)$ and $(C_6)$ are the same as Rathbun and Cressie’s [20]. $(C_2)$ is stronger than Rathbun and Cressie’s [20]. $(C_7)$ is an extra condition needed for our combined case. Condition $(C_2)$ means that $\phi_{A_j}(\theta)$ tends to zero at a uniform rate for all $\theta \in \Theta$, conditions $(C_3)$, $(C_4)$, and $(C_7)$ together address the smoothness of $\lambda(\cdot;\theta)$ and rate of increase of the derivative $\lambda(\cdot;\theta)$ (which means the properties of the intensity function with respect to parameter $\theta$), conditions $(C_5)$ shows sufficient separability of the intensities for adjacent values of $\theta$, conditions $(C_6)$ is required to show Lemmas 7 and 8.

The conditions constrain not only the family of intensity functions $\{ \lambda(\cdot;\theta) : \theta \in \Theta \}$, but also the nested sequence of sets that result from domain expansion $\{ A_j : j = 1, 2, \ldots \}$, which actually depend on the family of intensity functions. Proofs of asymptotic properties of estimators may be obtained by showing that conditions (i)-(iv) of Ibragimov and Has’minskii’s theorem are satisfied when the intensity function and the sequence of sets $A_j$ satisfy conditions $(C_1)$-$(C_7)$.

The main result of the asymptotic behavior of the maximum likelihood estimator is given by the following theorem.
Theorem 3. Let $\Theta \subset \mathbb{R}^k$ and let the conditions $(C_1) - (C_7)$ be satisfied. Let $K$ be an arbitrary compact set in $\Theta$. Then for all $\theta \in K$:

1. the maximum likelihood estimator $\hat{\theta}_{A_j}$ is consistent for $\theta$, as $j \to \infty$;
2. $\left(\phi_{A_j}(\theta)\right)^{-1}(\hat{\theta}_{A_j} - \theta) \xrightarrow{p} N(0, I_k)$, as $j \to \infty$, where $I_k$ is the $k \times k$ identity matrix;
3. $\lim_{j \to \infty} E_{\theta}^{(j)} \left|\left| \left(\phi_{A_j}(\theta)\right)^{-1}(\hat{\theta}_{A_j} - \theta) \right| \right|^r = E|\xi|^r, r = 1, 2, \ldots$, where $\xi \sim N(0, I_k)$ and $I_k$ is the $k \times k$ identity matrix.

Proof. Condition (ii) of Theorem 1 follows directly from condition $(C_2)$, we will show that conditions (i), (iii), and (iv) used in proving Theorem 1 are satisfied by proving the following Lemmas 6-8.

Theorem 4. Let the loss function be $w \in \mathcal{W}_c$, where $\mathcal{W}_c$ is the same as the one defined in Theorem 2. Suppose conditions $(C_1) - (C_4)$ are satisfied. Then for any family of estimators $\hat{\theta}_{A_j}$ of $\theta$ and any $\theta_0 \in \Theta$,

$$\lim \inf_{j \to \infty} \sup_{|\theta - \theta_0| < \delta} E_{\theta}^{(j)}[w((\phi_{A_j}(\theta))^{-1}(\theta^*_{A_j} - \theta))] \geq E\{w(\xi)\},$$

where $\xi \sim N(0, I_k)$ and $I_k$ is the $k \times k$ identity matrix.

Proof. The proof consists of showing the family $\{P^{A_j}_{\theta} : \theta \in \Theta\}$ is uniformly asymptotically normal (UAN) as $j \to \infty$ for all $\theta$ in $K \subset \Theta$ (see following Lemma 8) and then apply Theorem 2.

Assume that $\lambda(s; \theta) > 0$ for all $s \in \mathbb{R}^d$ and all $\theta \in \Theta$, so that $\{P^{A_j}_{\theta} : \theta \in \Theta\}$, $j \geq 1$ is an equivalent family of measures; that is, $P^{A_j}_{\theta_1}$ is absolutely continuous with respect to $P^{A_j}_{\theta_2}$ for all $\theta_1, \theta_2 \in \Theta$. Equivalence of the measure $\{P^{A_j}_{\theta} : \theta \in \Theta\}$ implies the existence of likelihood ratio (Liptser and Shirayaev [19]).

Assume that the conditional density function of $M|N$ is $g(\cdot; N)$. And we know that given total number of points $N$ and known(observed) number of points $M$, for a bounded set $A \subset \mathbb{R}^d$, the known locations $(s_1, \ldots, s_M)$ are randomly independently distributed with probability density proportional to $\lambda(\cdot; \theta)$. So our proposed combined model, the likelihood (joint likelihood of $(s_1, \ldots, s_M), N$ and $M$) is as follows. We know that $M \leq N$, and assume that the density function of $M|N$ has nothing to do with $\theta$. For example, we may assume $M|N \sim \text{binomial}(N, p)$.

$$L_A(\theta; (s_1, \ldots, s_M), N, M) = [(s_1, \ldots, s_M)|M, N][M|N][N]$$
\[ \prod_{i=1}^{M} \lambda(s_i; \theta) \left( \int_{A} \lambda(s; \theta) \nu(ds) \right)^M \times g(\cdot; N) \times \exp \left( - \int_{A} \lambda(s; \theta) \nu(ds) \right) \left( \int_{A} \lambda(s; \theta) \nu(ds) \right)^N/N! \]

where \([\star]\) is the density function of \([\star]\).

So the joint likelihood is as follows,

\[ L_A(\theta; (s_1, \cdots, s_M), N, M) = g(\cdot; N) \exp \left( \int_{A} \log \lambda(s; \theta) M(ds) - \int_{A} \lambda(s; \theta) \nu(ds) \right) \left( \int_{A} \lambda(s; \theta) \nu(ds) \right)^{N-M}/N!. \]

For any \(\theta_1, \theta_2 \in \Theta\), the likelihood ratio of our proposed combined model is

\[ \Lambda_{A_j}(\theta_1, \theta_2; N, M) = \frac{L_{A_j}(\theta_1; N, M)}{L_{A_j}(\theta_2; N, M)} \]

\[ = \exp \left( \int_{A_j} \log \frac{\lambda(s; \theta_1)}{\lambda(s; \theta_2)} M(ds) - \int_{A_j} (\lambda(s; \theta_1) - \lambda(s; \theta_2)) \nu(ds) \right) \times \left( \frac{\int_{A_j} \lambda(s; \theta_1) \nu(ds)}{\int_{A_j} \lambda(s; \theta_2) \nu(ds)} \right)^{N-M}. \]

Proofs of the asymptotic behavior of the maximum likelihood estimator need the following moments of random variables,

\[ \int_{A} \xi(s) \{N(ds) - \lambda(s; \theta) \nu(ds)\}, \]

and

\[ \exp \left( \int_{A} \log \xi(s) N(ds) \right), \]

where \(\xi\) is any (possibly complex-valued) deterministic Lebesgue measurable function such that \(\int_{A} \xi(s) \lambda(s; \theta) \nu(ds) < \infty\), for any bounded set \(A \in A\).

The following lemma provides the first two moments of \(\int_{A} \xi(s) \{N(ds) - \lambda(s; \theta) \nu(ds)\}\).

**Lemma 1.** If \(\int_{A} |\xi(s)| \lambda(s; \theta) \nu(ds) < \infty\), then

\[ E_{\theta} \left( \int_{A} \xi(s) Q_{\theta}(ds) \right) = 0. \]
Furthermore, if \( \int_A (\xi(s))^2 \lambda(s; \theta) \nu(ds) < \infty \), then
\[
E_\theta \left( \left( \int_A \xi(s)Q_\theta(ds) \right)^2 \right) = \int_A (\xi(s))^2 \lambda(s; \theta) \nu(ds).
\]  
(6)

Proof. The proof is trivial.

In order to calculate the moment of \( \exp \left( \int_A \log \xi(s)N(ds) \right) \), we need to define the probability generating functional of the point process \( N \) as follows,
\[
G_N(\xi) = E_\theta \left( \exp \left( \int_{\mathbb{R}^d} \log \xi(s)N(ds) \right) \right),
\]
where \( \xi \) is any deterministic Lebesgue measurable function on \( \mathbb{R}^d \). It is well defined if \( \int_{\mathbb{R}^d} (1 - \xi(s))\lambda(s; \theta) \nu(ds) < \infty \), for any point process \( N \) (Vere-Jones [23], Westcott [25]). For a nonhomogeneous Poisson process \( N \) on \( \mathbb{R}^d \), the probability generating functional is given by the following lemma.

**Lemma 2.** Let \( N \) be a nonhomogeneous Poisson process on \( \mathbb{R}^d \) with intensity \( \lambda(\cdot; \theta) \). Then the probability generating functional of \( N \) is
\[
G_N(\xi) = \exp \left( -\int_{\mathbb{R}^d} (1 - \xi(s))\lambda(s; \theta) \nu(ds) \right).
\]

Proof. See Daley and Vere-Jones [5].

**Lemma 3.** Suppose that \( N \) is a nonhomogeneous point process on \( A \) with intensity \( \lambda(\cdot; \theta) \), where \( \theta \in \Theta \), an open convex subset of \( \mathbb{R}^k \). \( A \subset A \) is a Borel set such that \( \int_A \lambda(s; \theta) \nu(ds) < \infty \), where \( \nu(\cdot) \) is Lebesgue measure. \( M \) is a count generating mechanism with support \([0, N]\) and it represents the known exact locations among total \( N \). Then, the elements on the diagonal of the Fisher information matrix has the following property,
\[
\frac{EM}{EN} \int_A \frac{\left( \hat{\lambda}(s; \theta) \right)^2}{\lambda(s; \theta)} \nu(ds) + (EN - EM) \left( \frac{\int_A \hat{\lambda}(s; \theta) \nu(ds)}{\int_A \lambda(s; \theta) \nu(ds)} \right)^2 \leq \int_A \frac{\left( \hat{\lambda}(s; \theta) \right)^2}{\lambda(s; \theta)} \nu(ds),
\]


**Lemma 4.** Let \( N \) be a nonhomogeneous Poisson process on a bounded Borel set \( A \subset \mathbb{R}^d \) with intensity \( \lambda(\cdot; \theta) \). Let \( \xi : A \rightarrow \mathbb{R} \) be any Lebesgue
measurable function. Then for some constant $c_{2m}$ such that

$$E_\theta \left( \int_A \xi(s) Q_\theta(ds) \right)^{2m} \leq c_{2m} \max \left( \int_A (\xi(s))^{2m} \lambda(s; \theta) \nu(ds), \left( \int_A (\xi(s))^{2} \lambda(s; \theta) \nu(ds) \right)^m \right).$$

**Proof.** See Rathbun and Cressie [20].

Lemmas 2-4 are needed for the proofs of Lemmas 5-8.

A distribution is not necessarily determined by its moments, but if the moment sequence is unique (there is only one probability distribution with this sequence of moments) then the moment generating function determines the distribution. A sufficient condition for the uniqueness of moment sequence is Carleman’s Condition [3]. If $X \sim F_X$ and we denote $EX^r = \mu'_r$, then the moment sequence is unique if

$$\sum_{r=1}^{\infty} \frac{1}{(\mu'_2)^{1/(2r)}} = +\infty.$$  

This condition is, in general, not easy to verify. So using the mgf to determine the distribution is hard. A better way is to use the characteristic functions. See Billingsley [1] or Resnick [21] for the detail of the characteristic functions. They necessitate understanding complex analysis though the characterization of a distribution is simplified. The characteristic function of $X$ is defined as

$$\Phi_X(t) = Ee^{itX},$$

where $i$ is the complex number $\sqrt{-1}$. The characteristic function always exists and it completely determines the distribution. That is, each CDF has a unique characteristic function. So we may state the following results.

**Lemma 5.** Assume conditions $(C_4)$ is satisfied and define

$$\Delta_{A_j}(\theta) = \phi_{A_j}(\theta) \int_{A_j} \frac{\hat{\lambda}(s; \theta)}{\lambda(s; \theta)} (N(ds) - \lambda(s; \theta) \nu(ds)), $$

where

$$(\phi_{A_j}(\theta))^{-2} = I_{A_j}(\theta) = \frac{EM}{EN} \int_{A_j} \frac{\hat{\lambda}(s; \theta)}{\lambda(s; \theta)} \left( \frac{\hat{\lambda}(s; \theta)}{\lambda(s; \theta)} \right)' \nu(ds).$$
\[(EN - EM) \left( \frac{\int_{A_j} \dot{\lambda}(s; \theta) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right) \left( \frac{\int_{A_j} \dot{\lambda}(s; \theta) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right)'.\]

Then, for all \(\theta\) in any compact set \(K \subset \Theta\),
\[
\Delta_{A_j}(\theta) \xrightarrow{L} N(0, I_k), \quad \text{as } j \to \infty. \tag{7}
\]

**Lemma 6.** Assume conditions \((C_1) - (C_7)\) are satisfied. The family \(\{P^A_j : \theta \in \Theta\}\) is uniformly asymptotically normal (UAN) as \(j \to \infty\).

**Lemma 7.** Assume conditions \((C_1), (C_2)\) and \((C_6)\) are satisfied. Then for any compact set \(K \subset \Theta\),
\[
\sup_{\theta \in K} E^A_{\theta} \left| (Z_{A_j}(u_1))^{1/2m} - Z_{A_j}(u_0))^{1/2m} \right|^{2m} \leq c |u_1 - u_0|^{2m},
\]
where \(c\) does not depend on \(u_0, u_1\) and \(A_j\), and \(Z_{A_j}(u)\) is the likelihood ratio defined in Definition 1 as \(Z_{A_j}(u) = \Lambda_{A_j}(\theta + \phi_{A_j}(\theta)u, \theta)\).

**Lemma 8.** Assume conditions \((C_1) - (C_7)\) are satisfied. Then for any compact set \(K \subset \Theta\), there exist constants \(c_1 > 0\) and \(j_0 > 0\) such that for all \(j > j_0\),
\[
\sup_{\theta \in K} P^A_{\theta} \{ Z_{A_j}(u) > \exp(-c_1 g(u)) \} \leq \exp(-c_1 g(u)),
\]
where \(g(u) = \min(|u|^2, |u|^{2\alpha})\), and \(\alpha\) is defined in \((C_5)\).

The proofs of Lemmas 5-8 are given in Appendix.

5. Conclusions and Discussions

We have investigated the asymptotic properties of maximum likelihood estimator of the parameters of the combined model. It can be verified that under appropriate regularity conditions, for our aggregate-point combined model, the maximum likelihood estimator is consistent, asymptotically normal and asymptotically efficient. A Monte Carlo Simulation study by Wu and Kaiser [26] suggests that the estimates of asymptotic standard deviations give a good approximation of the actual standard deviations as represented by the Monte Carlo estimation values, which indicates the applicability of the asymptotic theory. To have the fact \(\lim_{j \to \infty} \int_{A_j} (\lambda(s; \theta + \phi_{A_j}(\theta)u) - \lambda(s; \theta)) \nu(ds) = 0\), our assumption \((C_7)\) might be too strong. In fact, for the case in Wu and Kaiser [26]’s paper, the intensity function does not satisfy this assumption. But the simulation study still verifies the applicability of the asymptotic properties.
References


Appendix

This appendix gives the proofs of Lemmas 5, 6, 7, and 8.

Proof of Lemma 5. Let
\[ f(s) = \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)}; \ s \in \mathbb{R}^d. \]
Then the characteristic function of $\Delta A_j(\theta)$ is,

$$
\eta_{A_j}(t) = E^{(j)}_{\theta} \left( e^{it \Delta A_j(\theta)} \right)
$$

$$
= E^{(j)}_{\theta} \left( \exp \left( it' \phi A_j(\theta)' \int_{A_j} \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)} \left( N(ds) - \lambda(s; \theta)\nu(ds) \right) \right) \right)
$$

$$
= E^{(j)}_{\theta} \left( \exp \left( it' \phi A_j(\theta)' \int_{A_j} f(s) \left( N(ds) - \lambda(s; \theta)\nu(ds) \right) \right) \right)
$$

$$
= E^{(j)}_{\theta} \left( \exp \left( it' \phi A_j(\theta)' \int_{A_j} f(s)N(ds) \right) \right)
$$

$$
\times \exp \left( -it' \phi A_j(\theta)' \int_{A_j} f(s)\lambda(s; \theta)\nu(ds) \right).
$$

Applying Lemma 2, we know that,

$$
E^{(j)}_{\theta} \left( \exp \left( it' \phi A_j(\theta)' \int_{A_j} f(s)N(ds) \right) \right)
$$

$$
= \exp \left( - \int_{A_j} \left( 1 - \exp \left( it' \phi A_j(\theta)' f(s) \right) \right) \lambda(s; \theta)\nu(ds) \right).
$$

Therefore,

$$
\eta_{A_j}(t) = \exp \left( - \int_{A_j} \left( 1 - \exp \left( it' \phi A_j(\theta)' f(s) \right) \right) \lambda(s; \theta)\nu(ds) \right)
$$

$$
- it' \phi A_j(\theta)' \int_{A_j} f(s)\lambda(s; \theta)\nu(ds)
$$

$$
= \exp \left( \int_{A_j} \left( \exp \left( it' \phi A_j(\theta)' f(s) \right) - 1 - it' \phi A_j(\theta)' f(s) \right) \lambda(s; \theta)\nu(ds) \right).
$$

From Lemma 3, we know that, for scalar parameter case of the Fisher information matrix,

$$
I_{A_j}(\theta) \leq \int_{A_j} \frac{(\dot{\lambda}(s; \theta))^2}{\lambda(s; \theta)}\nu(ds)
$$

$$
\Rightarrow (\phi A_j(\theta))^{-2} = I_{A_j}(\theta) \leq \int_{A_j} \frac{(\dot{\lambda}(s; \theta))^2}{\lambda(s; \theta)}\nu(ds) = \int_{A_j} (f(s))^2\lambda(s; \theta)\nu(ds).
$$
Then,

\[
\log \eta_{A_j}(t) + \frac{1}{2}|t|^2 = \left| \int_{A_j} \left( \exp \left( it' \phi_{A_j}(\theta)' f(s) \right) - 1 - it' \phi_{A_j}(\theta)' f(s) \right) \lambda(s; \theta) \nu(ds) + \frac{1}{2}|t|^2 \right|
\]

\[
\leq \int_{A_j} |\exp \left( it' \phi_{A_j}(\theta)' f(s) \right) - 1 - it' \phi_{A_j}(\theta)' f(s) + \frac{1}{2}t' \phi_{A_j}(\theta)' f(s) f(s)'(\phi_{A_j}(\theta))'| \lambda(s; \theta) \nu(ds)
\]

\[
\leq \int_{A_j} \frac{1}{6} |t' \phi_{A_j}(\theta)' f(s)|^3 \lambda(s; \theta) \nu(ds),
\]

since \(|e^{ix} - 1 - ix + \frac{1}{2}x^2| \leq \min \left( \frac{1}{6}|x|^3, |x|^2 \right) \leq \frac{1}{6}|x|^3\) (see Durrett [6]). Therefore,

\[
\log \eta_{A_j}(t) + \frac{1}{2}|t|^2 \leq \frac{1}{6} \int_{A_j} |t' \phi_{A_j}(\theta)' f(s)|^3 \lambda(s; \theta) \nu(ds)
\]

\[
\leq \left( \int_{A_j} |t' \phi_{A_j}(\theta)' f(s)|^4 \lambda(s; \theta) \nu(ds) \int_{A_j} |t' \phi_{A_j}(\theta)' f(s)|^2 \lambda(s; \theta) \nu(ds) \right)^{1/2},
\]

by Cauchy-Schwarz inequality. By assumption \((C_4)\), the first integral of the right-hand side in the above expression approaches to zero uniformly for all \(\theta \in K\), as \(j \to \infty\) and the second integral is bounded by assumption \((C_2)\). Therefore, \(\eta_{A_j}(t) \to \exp \left( - \frac{1}{2}|t|^2 \right) ; t \in \mathbb{R}^k\), uniformly for all \(\theta \in K\), as \(j \to \infty\). From Feller [7], (7) on page 163 is proved.

**Proof of Lemma 6.** By condition \((C_1)\), \(\{P^A_j : \theta \in \Theta\}, j \geq 1\) is an equivalent family of measures. Since equivalence of the measure \(\{P^A_j : \theta \in \Theta\}\) implies the existence of likelihood ratio (Liptser and Shiryayev [19]), from (5), the log likelihood ratio of our proposed combined model is

\[
\log Z_{A_j}(u) = \log \left( A_{A_j}(\theta + \phi_{A_j}(\theta) u, \theta) \right) = u' \Delta_{A_j}(\theta) - \frac{1}{2}u'u + g_{A_j}(u, \theta)
\]

\[
+ \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta) u)}{\lambda(s; \theta)} M(ds) - \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta) u)}{\lambda(s; \theta)} N(ds)
\]
\[ + (N - M) \log \left( \frac{\int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta)u) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right) + h_{A_j}(u, \theta), \]

where
\[
g = \int_{A_j} \left( \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} - u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)} \right) (N(ds) - \lambda(s; \theta)) \nu(ds),
\]
and
\[
h = \frac{1}{2} u'u - \int_{A_j} \left( \lambda(s; \theta + \phi_{A_j}(\theta)u) - \lambda(s; \theta) - \lambda(s; \theta) \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} \right) \nu(ds).
\]

Part (II) will be proved to converge to zero in probability. Part (I) has the similar form as the corresponding part in Rathbun and Cressie [20]. We will deal it in the similar way. First of all, part (I) may be written as
\[
I = u'\Delta_{A_j}(\theta) - \frac{1}{2} u'u + g_{A_j}(u, \theta),
\]
by Lemma 5,
\[
\Delta_{A_j}(\theta) \xrightarrow{\mathcal{L}} N(0, I_k), \text{ as } j \to \infty.
\]
It remains to show that,
\[
g_{A_j}(u, \theta) \xrightarrow{p} 0, \text{ as } j \to \infty,
\]
uniformly for all \( \theta \in K \), as \( j \to \infty \).

Consider the function \( g_{A_j}(u, \theta) \) first. By Chebyshev’s inequality and Lemma 1, for any \( \epsilon > 0 \),
\[
P_{\theta}^{A_j}\{|g_{A_j}(u, \theta)| > \epsilon\} \leq \epsilon^{-2} E_{\theta}^{(j)} \left( (g_{A_j}(u, \theta))^2 \right)
\]
\[
= \epsilon^{-2} \int_{A_j} \left( \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} - u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)} \right)^2 \lambda(s; \theta) \nu(ds). \quad (8)
\]

Let \( \theta_{yu} = \theta + y\phi_{A_j}(\theta)u; \ y \in [0, 1] \). By \( (C_1) \), we know that the functions \( \lambda(s; \theta_{yu}) \) and \( \log \lambda(s; \theta_{yu}) \) are absolutely continuous with respective to \( y \) for almost all \( s \in \mathbb{R}^{d} \). (which means for all \( s \in \mathbb{R}^{d} \), except on a set of Lebesgue measure zero.) Then,
\[
\int_{A_j} \left( \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} - u'\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)} \right)^2 \lambda(s; \theta) \nu(ds)
\]
\[ \int_{A_j} \left( \int_0^1 u' \phi_{A_j}(\theta)' \left( \psi(s; \theta_{yu}) \left( \frac{\lambda(s; \theta)}{\lambda(s; \theta_{yu})} \right)^{1/2} - \dot{\psi}(s; \theta) \right) dy \right)^2 \nu(ds) \]
\[ \leq \int_0^1 \int_{A_j} \left( u' \phi_{A_j}(\theta)' \left( \psi(s; \theta_{yu}) \left( \frac{\lambda(s; \theta)}{\lambda(s; \theta_{yu})} \right)^{1/2} - \dot{\psi}(s; \theta) \right) \right)^2 \nu(ds)dy, \]

by Jensen’s inequality and Fubini’s Theorem. Assumption \((C_3)\) tells us that
\[ \lim \mu(A_j) = \infty; \]
so for \( j \) large enough, \( \{\theta_{yu} : y \in [0,1]\} \subset \{\theta_v : |v| \leq \mu(A_j)\}. \)

By adding and subtracting \( \dot{\psi}(s; \theta_{yu}) \) inside the above norm, knowing the inequality \((a + b)^2 \leq 2a^2 + 2b^2\) and taking the supremum over the interval \([0,1]\), for \( j \) large enough, we have the following,
\[ \int_0^1 \int_{A_j} \left( u' \phi_{A_j}(\theta)' \left( \psi(s; \theta_{yu}) \left( \frac{\lambda(s; \theta)}{\lambda(s; \theta_{yu})} \right)^{1/2} - \dot{\psi}(s; \theta) \right) \right)^2 \nu(ds)dy \]
\[ = \int_0^1 \int_{A_j} \left( u' \phi_{A_j}(\theta)' \left( \psi(s; \theta_{yu}) \left( \frac{\lambda(s; \theta)}{\lambda(s; \theta_{yu})} \right)^{1/2} - \dot{\psi}(s; \theta) + \dot{\psi}(s; \theta_{yu}) \right) - \dot{\psi}(s; \theta_{yu}) \right)^2 \nu(ds)dy \]
\[ \leq 2|u|^2 \sup_{y \in [0,1]} \left\| \phi_{A_j}(\theta)' \left( \psi(s; \theta_{yu}) \left( \frac{\lambda(s; \theta)}{\lambda(s; \theta_{yu})} \right)^{1/2} - \dot{\psi}(s; \theta_{yu}) \right) \right\|^2_{A_j} \]
\[ + 2|u|^2 \sup_{|v| < \mu(A_j)} \left\| \phi_{A_j}(\theta)' \left( \psi(s; \theta_v) - \dot{\psi}(s; \theta) \right) \right\|^2_{A_j}. \]

By assumption \((C_3)\), the above second term converges to zero uniformly for all \( \theta \in K, \) as \( j \to \infty. \) Again, since the functions \( \lambda(s; \theta_{zu}) \) is absolutely continuous with respect to \( z, \) the above first term may be written as,
\[ 2|u|^2 \sup_{y \in [0,1]} \left\| \phi_{A_j}(\theta)' \left( \psi(s; \theta_{yu}) \left( \frac{\lambda(s; \theta)}{\lambda(s; \theta_{yu})} \right)^{1/2} - \dot{\psi}(s; \theta_{yu}) \right) \right\|^2_{A_j} \]
\[ = 2|u|^2 \sup_{y \in [0,1]} \left( \int_{A_j} \left| \phi_{A_j}(\theta)' \frac{\lambda(s; \theta_{yu})}{\lambda(s; \theta_{yu})} \right|^2 \right) \times \left( \int_0^y u' \phi_{A_j}(\theta)' \frac{\lambda(s; \theta_{yu})}{2(\lambda(s; \theta_{yu}))^{1/2}} dz \right)^2 \nu(ds). \]

We follow the procedure we did at the beginning of this proof, by Jensen’s inequality, Fubini’s Theorem and taking the supremum over the interval \([0, y],\)
the above expression is as follows,

\[
2|u|^2 \sup_{y \in [0, 1]} \int_{A_j} \left| \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{yu})}{\lambda(s; \theta_{yu})} \right|^2 \left( \int_0^y u' \phi_{A_j}(\theta)' \frac{\dot{\lambda}(s; \theta_{zu})}{2(\lambda(s; \theta_{zu}))^{1/2}}\nu(ds) \right)^2 \nu(ds) \\
\leq \frac{1}{2} |u|^4 \sup_{y \in [0, 1]} \sup_{z \in [0, y]} \int_{A_j} \left| \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{yu})}{\lambda(s; \theta_{yu})} \right|^2 \left( \phi_{A_j}(\theta)' \frac{\dot{\lambda}(s; \theta_{zu})}{\lambda(s; \theta_{zu})} \right)^2 \lambda(s; \theta_{zu}) \nu(ds)
\]

Again, since \( \lim_{j \to \infty} \mu(A_j) = \infty \), for \( j \) large enough, \( \{ (\theta_{yu}, \theta_{zu}) : y \in [0, 1], z \in [0, y] \} \subset \{ \theta_v, \theta_w \} : |v| + |w| \leq \mu(A_j) \}. Thus for \( j \) large enough, we have,

\[
\frac{1}{2} |u|^4 \sup_{y \in [0, 1]} \sup_{z \in [0, y]} \int_{A_j} \left| \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{yu})}{\lambda(s; \theta_{yu})} \right|^2 \left( \phi_{A_j}(\theta)' \frac{\dot{\lambda}(s; \theta_{zu})}{\lambda(s; \theta_{zu})} \right)^2 \lambda(s; \theta_{zu}) \nu(ds) \\
\leq \frac{1}{2} |u|^4 \sup_{|u| + |v| < \mu(A_j)} \int_{A_j} \left| \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{yu})}{\lambda(s; \theta_{yu})} \right|^2 \left( \phi_{A_j}(\theta)' \frac{\dot{\lambda}(s; \theta_{zu})}{\lambda(s; \theta_{zu})} \right)^2 \lambda(s; \theta_{zu}) \nu(ds) \\
\leq \frac{1}{2} |u|^4 \sup_{|u| + |v| < \mu(A_j)} \left( \int_{A_j} \left| \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{yu})}{\lambda(s; \theta_{yu})} \right|^4 \lambda(s; \theta_{zu}) \nu(ds) \right)^{1/2} \\
\times \int_{A_j} \left| \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{zu})}{\lambda(s; \theta_{zu})} \right|^4 \lambda(s; \theta_{zu}) \nu(ds) \right)^{1/2}
\]

by Cauchy-Schwarz inequality. By assumption \((C_4)\), the right-hand side of the above inequality converges to zero uniformly for all \( \theta \in K \), as \( j \to \infty \). So the right-hand side of (8) converges to zero uniformly for all \( \theta \in K \), as \( j \to \infty \). Therefore, \( g_{A_j}(u, \theta) \to 0 \), as \( j \to \infty \), uniformly for all \( \theta \in K \). So far, we are done with part (I).

Now we work on part (II). From now on, for simplicity, we assume scalar parameter for the rest of the proof of this lemma. First of all, from Lemma 5, we have,

\[
\frac{1}{2} u^2 = \frac{1}{2} u^2 (\phi_{A_j}(\theta))^2 \left( \frac{EM}{EN} \int_{A_j} \left( \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \nu(ds) \right) \\
+(EN - EM) \left( \frac{\int_{A_j} \dot{\lambda}(s; \theta) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right)^2 = B_1 + B_2,
\]
where \( B_1 = \frac{1}{2} u^2 (\phi_{A_j}(\theta))^2 \int_{A_j} \left( \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \nu(ds) \), and
\[
B_2 = \frac{1}{2} u^2 (\phi_{A_j}(\theta))^2 (EN - EM) \left( \left( \frac{\int_{A_j} \dot{\lambda}(s; \theta) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right)^2 - \frac{1}{EN} \int_{A_j} \left( \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \nu(ds) \right).
\]

So,
\[
h_{A_j}(u, \theta) = B_1 + B_2 - B_3,
\]
where,
\[
B_3 = \int_{A_j} \left( \lambda(s; \theta + \phi_{A_j}(\theta)u) - \lambda(s; \theta) - \lambda(s; \theta) \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} \right) \nu(ds).
\]

For \( B_1 \) and \( B_3 \), we will prove they converge to zero uniformly as follows. For \( B_2 \), we will prove it with the rest of part (II) together converges to zero in probability after we are done with \( B_1 \) and \( B_3 \).

It is very obvious that,
\[
B_1 = \frac{1}{2} u^2 (\phi_{A_j}(\theta))^2 \int_{A_j} \left( \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \nu(ds)
= \int_{A_j} \int_0^1 \int_0^y \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \nu(ds) dz dy.
\]

Since the function \( \lambda(s; \theta) \) is absolutely continuous with respect to \( z \), for almost all \( s \in \mathbb{R}^d \),
\[
B_3 = \int_{A_j} \left( \lambda(s; \theta + \phi_{A_j}(\theta)u) - \lambda(s; \theta) - \lambda(s; \theta) \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} \right) \nu(ds)
= \int_0^1 \int_0^y \int_{A_j} u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta\nu)}{\lambda(s; \theta\nu)} u\phi_{A_j}(\theta) \lambda(s; \theta\nu) \nu(ds) dz dy.
\]

Therefore, we have,
\[
|B_1 - B_3| = \left| \int_{A_j} \int_0^1 \int_0^y \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \nu(ds) dz dy \right|
\]
\[-\int_{A_j} \int_0^1 \int_0^y u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_yu)}{\lambda(s; \theta_yu)} u\phi_{A_j}(\theta) \dot{\lambda}(s; \theta_{zu}) dz dy \nu(ds)\]

\[\leq \sup_{0 \leq z \leq y \leq 1} \int_{A_j} \left| u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_yu)}{\lambda(s; \theta_yu)} u\phi_{A_j}(\theta) \dot{\lambda}(s; \theta_{zu}) - \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \right| \nu(ds)\]

By using the triangle inequality, we may deal the right-hand side of the above as follows,

\[\sup_{0 \leq z \leq y \leq 1} \int_{A_j} \left| u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_yu)}{\lambda(s; \theta_yu)} u\phi_{A_j}(\theta) \dot{\lambda}(s; \theta_{zu}) - \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \right| \nu(ds) \]

\[\leq \sup_{0 \leq z \leq y \leq 1} \int_{A_j} \left| u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_yu)}{\lambda(s; \theta_yu)} u\phi_{A_j}(\theta) \dot{\lambda}(s; \theta_{zu}) \right.\]

\[- \left. \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{zu})}{(\lambda(s; \theta_{zu}))^{1/2}} \right)^2 \right| \nu(ds)\]

\[+ \sup_{|v| < \mu(A_j)} \int_{A_j} \left| u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_v)}{(\lambda(s; \theta_v))^{1/2}} \right.\]

\[- \left. \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \right| \nu(ds).\]

By Cauchy-Schwarz inequality and triangle inequality, the second term of above is,

\[\sup_{|v| < \mu(A_j)} \int_{A_j} \left| \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_v)}{(\lambda(s; \theta_v))^{1/2}} \right)^2 \right.\]

\[- \left. \left( u\phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \right| \nu(ds)\]

\[\leq \sup_{|v| < \mu(A_j)} \left| u\phi_{A_j}(\theta)(\dot{\psi}(s; \theta_v) - \dot{\psi}(s; \theta)) \right|_{A_j} \left| u\phi_{A_j}(\theta)\dot{\psi}(s; \theta_v) \right|_{A_j}\]

\[+ \sup_{|v| < \mu(A_j)} \left| u\phi_{A_j}(\theta)(\dot{\psi}(s; \theta_v) - \dot{\psi}(s; \theta)) \right|_{A_j} \left| u\phi_{A_j}(\theta)\dot{\psi}(s; \theta) \right|_{A_j}\]

\[= \sup_{|v| < \mu(A_j)} \left| u\phi_{A_j}(\theta)(\dot{\psi}(s; \theta_v) - \dot{\psi}(s; \theta)) \right|_{A_j} \left| u\phi_{A_j}(\theta)\dot{\psi}(s; \theta_v) \right|_{A_j}\]

\[\times \left( \left| u\phi_{A_j}(\theta)\dot{\psi}(s; \theta_v) \right|_{A_j} + \left| u\phi_{A_j}(\theta)\dot{\psi}(s; \theta) \right|_{A_j} \right).\]
By assumptions \((C_2)\) and \((C_3)\), the above converges to zero uniformly for all \(\theta \in K\), as \(j \to \infty\). By Cauchy-Schwarz inequality, the first term of the right-hand side of (10) is,

\[
\sup_{0 \leq z \leq y \leq 1} \left| \int_{A_j} \left( u \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_{yu})}{\lambda(s; \theta_{yu})} u \phi_{A_j}(\theta) \lambda(s; \theta_{zu}) \right) ds \right|^2 \nu(ds) = \sup_{0 \leq z \leq y \leq 1} \left| \int_{A_j} \left( u \phi_{A_j}(\theta) \dot{\psi}(s; \theta_{zu}) \right) ds \right| \nu(ds) \leq \sup_{0 \leq z \leq y \leq 1} \left| u \phi_{A_j}(\theta) \dot{\psi}(s; \theta_{zu}) \right|_{A_j} \nu(ds) \leq c_0 |u|.
\]

By assumption \((C_2)\),

\[
\left| u \phi_{A_j}(\theta) \dot{\psi}(s; \theta_{zu}) \right|_{A_j} = \left( \int_{A_j} \left( u \phi_{A_j}(\theta) \dot{\psi}(s; \theta_{zu}) \right)^2 ds \right)^{1/2} \leq \sup_{0 \leq z \leq y \leq 1} \phi_{A_j}(\theta) \left( \frac{\dot{\lambda}(s; \theta_{yu})^2}{\lambda(s; \theta_{yz})} \right)^{1/2} \nu(ds) \leq c_0 |u|.
\]

Since \((a + b)^2 \leq 2a^2 + 2b^2\) and by subtracting and adding \(\dot{\psi}(s; \theta_{yu})\) inside the norm sign, we have,

\[
\left| u \phi_{A_j}(\theta) \left( \dot{\psi}(s; \theta_{yu}) \frac{\lambda(s; \theta_{zu})}{\lambda(s; \theta_{yu})} \right)^{1/2} \right|_{A_j} \leq \left( \int_{A_j} \left( \frac{\dot{\lambda}(s; \theta_{yu})^2}{\lambda(s; \theta_{yu})} \right)^{1/2} \right)^{1/2} \leq c_0 |u|.
\]
\[
\leq 2 \left\| u \phi_{A_j}(\theta) \left( \frac{\dot{\psi}(s; \theta_{yu})}{\lambda(s; \theta_{yu})^{1/2}} - \psi(s; \theta_{yu}) \right) \right\|_{A_j}^2 \\
+ 2 \left\| u \phi_{A_j}(\theta) \left( \dot{\psi}(s; \theta_{yu}) - \psi(s; \theta_{zu}) \right) \right\|_{A_j}^2.
\]

Using the similar arguments we used for the first term of (9), the first term of the right-hand side of (10) converges to zero uniformly for all \( \theta \in K \), as \( j \to \infty \).

So it then finally follows that \( B_1 - B_3 \to 0 \), uniformly for all \( \theta \in K \), as \( j \to \infty \).

So far, we have done with \( B_1 - B_3 \).

Now, we prove the rest of part (II) converges to zero in probability. Since the rest of part (II) can be written as following,

\[
II' = B_2 - \sum_{i=1}^{N-M} \log \frac{\lambda(s_i; \theta + \phi_{A_j}(\theta)u)}{\lambda(s_i; \theta)} + (N-M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta)u)\nu(ds)}{\int_{A_j} \lambda(s; \theta)\nu(ds)},
\]

by Chebyshev’s inequality, we have,

\[
P(|II'| \geq \epsilon) \leq \epsilon^{-2} E(II')^2.
\]

Now, we deal with \( E(II')^2 \) as follows,

\[
E(II')^2 = D_1 + D_2 + B_2^2.
\]

(11)

\( D_1 \) and \( D_2 \) are defined as follows,

\[
D_1 = E \left( (N-M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta)u)\nu(ds)}{\int_{A_j} \lambda(s; \theta)\nu(ds)} \right. \\
- \left. \sum_{i=1}^{N-M} \log \frac{\lambda(s_i; \theta + \phi_{A_j}(\theta)u)}{\lambda(s_i; \theta)} \right)^2 = E(N-M)^2 \frac{E(N-M)^2}{(\int_{A_j} \lambda(s; \theta)\nu(ds))^2} G_1^2 + \frac{E(N-M)}{\int_{A_j} \lambda(s; \theta)\nu(ds)} G_2;
\]

where

\[
G_1 = \int_{A_j} \left( \log \frac{\int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta)u)\nu(ds)}{\int_{A_j} \lambda(s; \theta)\nu(ds)} \lambda(s; \theta) \\
- \lambda(s; \theta) \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} \right) \nu(ds);
\]

and
\[ G_2 = \int_{A_j} \lambda(s; \theta) \left( \log \frac{\lambda(s; \theta + \phi A_j(\theta)u)}{\lambda(s; \theta)} \right)^2 \nu(ds) \]

\[ \quad - \frac{\left( \int_{A_j} \lambda(s; \theta) \log \frac{\lambda(s; \theta + \phi A_j(\theta)u)}{\lambda(s; \theta)} \nu(ds) \right)^2}{\int_{A_j} \lambda(s; \theta) \nu(ds)}. \]

\[ D_2 = 2B_2 E \left( (N - M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi A_j(\theta)u) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right. \]

\[ \quad - \sum_{i=1}^{N-M} \log \frac{\lambda(s_i; \theta + \phi A_j(\theta)u)}{\lambda(s_i; \theta)} \bigg)^2 \nu(ds) \bigg) \]

\[ = 2B_2 E(N - M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi A_j(\theta)u) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \]

\[ - 2B_2 \frac{E(N - M)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \int_{A_j} \lambda(s; \theta) \log \frac{\lambda(s; \theta + \phi A_j(\theta)u)}{\lambda(s; \theta)} \nu(ds). \]

Since

\[ B_2 = \frac{1}{2} u^2 (\phi A_j(\theta))^2 (EN - EM) \]

\[ \times \left( \frac{\int_{A_j} \lambda(s; \theta) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right)^2 - \frac{1}{EN} \int_{A_j} \left( \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)^{1/2}} \right)^2 \nu(ds) \bigg) \]

\[ = \frac{1}{2} u^2 (\phi A_j(\theta))^2 \left( \frac{EN - EM}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right) \left( \frac{\left( \int_{A_j} \lambda(s; \theta) \nu(ds) \right)^2}{\int_{A_j} \lambda(s; \theta) \nu(ds)} - \int_{A_j} \frac{\dot{\lambda}(s; \theta)^2}{\lambda(s; \theta)} \nu(ds) \right), \]

\[ D_2 = 2B_2 E(N - M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi A_j(\theta)u) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \]

\[ - 2B_2 \frac{E(N - M)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \int_{A_j} \lambda(s; \theta) \log \frac{\lambda(s; \theta + \phi A_j(\theta)u)}{\lambda(s; \theta)} \nu(ds) \]

\[ = u^2 (\phi A_j(\theta))^2 \left( \frac{(EN - EM)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right)^2 H_1 G_1. \]
So, from (11), we have
\[
E(II')^2 = D_1 + D_2 + B_2^2
= \frac{Var(N - M)}{(\int_{A_j} \lambda(s; \theta) \nu(ds))^2} G_1^2 + \frac{E(N - M)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} G_2
+ \left( \frac{E(N - M)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right)^2 \left( G_1 + \frac{1}{2} u^2(\phi_{A_j}(\theta))^2 H_1 \right)^2.
\] (12)

By Cauchy-Schwarz inequality, we know that \( G_2 \geq 0 \). From the proof of Lemma 3, we know that \( H_1 \leq 0 \). It seems that \( G_1, G_2 \) and \( H_1 \) are not zero obviously, we only know they are when the intensity function is a constant, e.g., \( \lambda(s; \theta) = \exp(\theta) \), which is not what we expected. Actually, by \((C_2)\), \( H_1 \) is bounded, for \( G_1 \) and \( G_2 \), since they are zero for constant intensity case, we may conclude they are bounded for general case, otherwise there is a contradiction. By the continuity of \( \lambda(\cdot) \) with respect to \( y \), Taylor expansion, and \((C_7)\), for large \( j \), we may easily verify that both \( G_1 \) and \( G_2 \) are approximately zero.

Since \( H_1 \leq 0 \), we have already known that \( B_1 \rightarrow B_3 \) as \( j \rightarrow \infty \), by \((C_7)\), for large \( j \), we have
\[
\left| \frac{1}{2} u^2(\phi_{A_j}(\theta))^2 H_1 \right| \approx 0
\]
Now, it turns out that to have above (12) approaches to zero uniformly as \( j \rightarrow \infty \), we need the assumptions \( E(N - M) = O \left( \int_{A_j} \lambda(s; \theta) \nu(ds) \right) \) and \( Var(N - M) = O \left( \left( \int_{A_j} \lambda(s; \theta) \nu(ds) \right)^2 \right) \). It is very obvious that when \( M = N \), which is the point model, the above assumptions are satisfied. When \( M \leq N \), the assumptions are also rather unsurprising. Thus, we may conclude that part\( (\Pi') \) converges to zero in probability. Therefore, under two new assumptions, Lemma 6 is proved, that is, the family \( \{P_{\theta}^{A_j} : \theta \in \Theta \} \) is uniformly asymptotically normal (UAN) as \( j \rightarrow \infty \).

Proof of Lemma 7. For \( y \in [0,1] \), let \( u_y = u_0 + (u_1 - u_0) y \), \( \theta(y) = \theta + \phi_{A_j}(\theta) u_y \), and \( R_{\theta(y)}(ds) = N(ds) - \lambda(s; \theta(y)) \). Since \( \lambda(s; \theta(y)) \) is absolutely continuous with respect to \( y \) for almost all \( s \in \mathbb{R}^d \), \( Z_{A_j}(u_y) \) is also absolutely continuous with respect to \( y \). Then
\[
E_{\theta} \left( \frac{1}{2m} \left( (Z_{A_j}(u_1))^{\frac{1}{2m}} - (Z_{A_j}(u_0))^{\frac{1}{2m}} \right) \right)^{2m} = E_{\theta} \left( \int_0^1 \frac{\partial}{\partial y} \left( (Z_{A_j}(u_y))^{\frac{1}{2m}} \right) dy \right)^{2m}.
\] (13)
From the proof of Lemma 6, 

\[
\log (Z_{A_j}(u_y)) = \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u_y)}{\lambda(s; \theta)} N(ds)
\]

\[- \int_{A_j} (\lambda(s; \theta + \phi_{A_j}(\theta)u_y) - \lambda(s; \theta)) \nu(ds)
\]

\[- \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u_y)}{\lambda(s; \theta)} (N - M)(ds)
\]

\[+ (N - M) \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u_y) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)},
\]

so,

\[
\frac{\partial}{\partial y} ((Z_{A_j}(u_y))) = Z_{A_j}(u_y) \left( \int_{A_j} (u_1 - u_0)' \phi_{A_j}(\theta)' \frac{\hat{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} \nu(ds) \right)
\]

\[\times \left( Q_{\theta(y)}(ds) + R_{\theta(y)}^{(1)}(ds) - R_{\theta(y)}^{(2)}(ds) \right)
\]

\[= Z_{A_j}(u_y) \left( \int_{A_j} (u_1 - u_0)' \phi_{A_j}(\theta)' \frac{\hat{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} \left( Q_{\theta(y)}(ds) + R_{\theta(y)}(ds) \right) \right)
\]

\[= Z_{A_j}(u_y) \left( \int_{A_j} (u_1 - u_0)' \phi_{A_j}(\theta)' \frac{\hat{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} S_{\theta(y)}(ds) \right),
\]

where \(R_{\theta(y)}^{(1)}(ds) = (N - M) \int_{A_j} \frac{\lambda(s; \theta(y))}{\lambda(s; \theta(y))} \nu(ds)\), \(R_{\theta(y)}^{(2)}(ds) = (N - M)(ds)\), \(R_{\theta(y)}(ds) = R_{\theta(y)}^{(1)}(ds) - R_{\theta(y)}^{(2)}(ds)\), \(Q_{\theta(y)}(ds) = N(ds) - \lambda(s; \theta(y)) \nu(ds)\), and \(S_{\theta(y)}(ds) = Q_{\theta(y)}(ds) + R_{\theta(y)}(ds)\). Thus, by Fubini’s theorem and Jensen’s inequality, (13) on page 176 can be processed as

\[
E_\theta \left( \left( (Z_{A_j}(u_1)) \right)^{2m} - \left( (Z_{A_j}(u_0)) \right)^{2m} \right)^{2m} = E_\theta \left( \int_0^1 \frac{\partial}{\partial y} \left( (Z_{A_j}(u_y)) \right)^{2m} dy \right)^{2m}
\]

\[
\leq (2m)^{-2m} \int_0^1 E_\theta(y) \left( \int_{A_j} (u_1 - u_0)' \phi_{A_j}(\theta)' \frac{\hat{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} S_{\theta(y)}(ds) \right)^{2m} dy
\]

since \(Z_{A_j}(u_y) = dP_{\theta(y)}/dP_\theta\).

Now, assume that for \(a > 0, b > 0\), and \(p > 1\), let’s verify an inequality as follows first,

\[(a + b)^p \leq 2^{p-1}(a^p + b^p).
\]
Actually, let \( f(x) = x^p \), and \( X = a \) or \( b \) with the same probability. Since \( f(x) \)

is a convex function here, by Jensen’s Inequality, we may have above (15). For the current case \( p = 2m \), since \( f''(x) = 2m(2m - 1)(x^{m-1})^2 \), we only need to require \( m > 1/2 \) to have inequality (15).

By inequality (15), the above (14) now is,

\[
(2m)^{-2m} \int_0^1 E_{\theta(y)} \left( \int_{A_j} (u_1 - u_0)' \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} S_{\theta(y)}(ds) \right)^{2m} dy \\
+ R_{\theta(y)}(ds) \right)^{2m} dy \\
\leq \frac{(2m)^{-2m+1}}{2m} \int_0^1 E_{\theta(y)} \left( \int_{A_j} (u_1 - u_0)' \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} Q_{\theta(y)}(ds) \right)^{2m} dy \\
+ (2m)^{-2m+1} \int_0^1 E_{\theta(y)} \left( \int_{A_j} (u_1 - u_0)' \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} R_{\theta(y)}(ds) \right)^{2m} dy.
\]

\( T_1 \) has been taken care of by Rathbun and Cressie [20] as follows. By Lemma 4, \( T_1 \) is less than or equal to the maximum of the following two quantities

\[
c_{2m}(2m)^{-2m+1} \int_0^1 \int_{A_j} \left( u_1 - u_0 \right)' \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} \right)^{2m} \lambda(s; \theta(y)) \nu(ds) dy,
\]

and

\[
c_{2m}(2m)^{-2m+1} \int_0^1 \left( \int_{A_j} \left( u_1 - u_0 \right)' \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} \right)^{2m} \lambda(s; \theta(y)) \nu(ds) \right)^{m} dy,
\]

for some constant \( c_{2m} \). The first term above is less than or equal to

\[
c_{2m}(2m)^{-2m+1} |u_1 - u_0|^{2m} \sup_{\theta_0 \in \Theta} \int_{A_j} \phi_{A_j}(\theta) \frac{\dot{\lambda}(s; \theta_0)}{\lambda(s; \theta_0)} |^{2m} \lambda(s; \theta_0) \nu(ds).
\]

By condition (\( C_0 \)), we know there exists a constant \( c_1 \), such that above expression is less than or equal to \( c_1 |u_1 - u_0|^{2m} \). By assumption (\( C_2 \)), we know there exists a constant \( c_2 \), such that the second term above is less than or equal to \( c_2 |u_1 - u_0|^{2m} \).
Now, we start working on $T_2$. Without losing generality, starting from now on, we assume our case is scalar parameter (i.e., $\theta$ is one dimensional parameter, thus, $k = 1$). We may have the same results analogously for multi-parameter case ($k \geq 1$). Choose $m = 1$ (obviously, $m = 1 > \frac{k}{2} = \frac{1}{2}$), it is obvious that

$$T_2 = \frac{1}{2} \int_0^1 E_{\theta(y)} \left( \int_{A_j} (u_1 - u_0) \phi_{A_j}(\theta) \frac{\hat{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} R_{\theta(y)}(ds) \right)^2 dy$$

$$= \frac{1}{2} (u_1 - u_0)^2 \left( \phi_{A_j}(\theta) \right)^2 \int_0^1 T'_2 dy,$$

where

$$T'_2 = E_{\theta(y)} \left( \int_{A_j} \frac{\hat{\lambda}(s; \theta(y))}{\lambda(s; \theta(y))} \left( R^{(1)}_{\theta(y)}(ds) + R^{(2)}_{\theta(y)}(ds) \right) \right)^2$$

$$= E_{\theta(y)} \left( (N - M) \int_{A_j} \frac{\hat{\lambda}(s; \theta(y)) \nu(ds)}{\lambda(s; \theta(y)) \nu(ds)} - \int_{A_j} \frac{\hat{\lambda}(s; \theta(y)) \nu(ds)}{\lambda(s; \theta(y)) \nu(ds)} (N - M)(ds) \right)^2.$$

For simplicity, $\hat{\lambda}(s; \theta(y))$ and $\lambda(s; \theta(y))$ are denoted as $\hat{\lambda}$ and $\lambda$ respectively in the following, it can be easily verified that

$$T_2 \leq (2)^{-1} (u_1 - u_0)^2 \left( \phi_{A_j}(\theta) \right)^2$$

$$\times \sup_{0 \leq y \leq 1} \left( \frac{E_{\theta(y)} (N - M)}{\int_{A_j} \lambda \nu(ds)} \left( \int_{A_j} \frac{\hat{\lambda}^2}{\lambda} \nu(ds) - \left( \int_{A_j} \frac{\hat{\lambda} \nu(ds)}{\lambda \nu(ds)} \right)^2 \right) \right).$$

By Casella and Berger’s [2] lemma and condition $(C_2)$, we may know that there exists a constant $c_3$, the right hand side of the above expression is less than or equal to $c_3|u_1 - u_0|^2$. Analogously for general case, we may also have $c_3|u_1 - u_0|^{2m}$. Therefore, if we choose $c = \max\{c_1, c_2\} + c_3$, Lemma 7 is proved.

**Proof Lemma 8.**

$$((Z_{A_j}(u))^{\frac{1}{2}}$$

$$= \exp \left( \frac{1}{2} \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} N(ds) - \frac{1}{2} \int_{A_j} (\lambda(s; \theta + \phi_{A_j}(\theta)u) - \lambda(s; \theta)) \nu(ds) \right.$$

$$- \frac{1}{2} \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} (N - M)(ds)$$

$$\left. - \frac{1}{2} \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} (N - M)(ds) \right).$$
\[ + \frac{1}{2}(N - M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta)u) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \]

\[ = \exp \left( \frac{1}{2} \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} N(ds) - \frac{1}{2} \int_{A_j} (\lambda(s; \theta + \phi_{A_j}(\theta)u) - \lambda(s; \theta)) \nu(ds) \right) \]

\[ \times \exp \left( \frac{1}{2}(N - M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta)u) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right) \]

\[ \times \exp \left( -\frac{1}{2} \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} (N - M)(ds) \right) \]

\[ = W_1 \times W_2, \text{ where } W_2 = W_{2.1} \times W_{2.2}. \]

By Chebychev’s inequality, Cauchy-Schwarz inequality and Jensen’s inequality, we know that

\[ P_\theta^{A_j} \{ Z_{A_j}(u) > \exp (-c_1 g(u)) \} \]

\[ \leq \exp \left( \frac{c_1}{4} g(u) \right) (E_\theta W_1)^{\frac{1}{2}} (E_\theta W_2)^{\frac{1}{2}}. \]

By Lemma 2, we know that

\[ E_\theta W_1 = E_\theta \left( \exp \left( \frac{1}{2} \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} N(ds) \right) \right. \]

\[ - \frac{1}{2} \int_{A_j} \left( (\lambda(s; \theta + \phi_{A_j}(\theta)u) - \lambda(s; \theta)) \nu(ds) \right) \]

\[ = \exp \left( -\frac{1}{2} \int_{A_j} \left( (\lambda(s; \theta + \phi_{A_j}(\theta)u))^\frac{1}{2} - (\lambda(s; \theta))^\frac{1}{2} \right)^2 \nu(ds) \right). \] (16)

Now, let’s calculate \( E_\theta W_2^2 \) as follows,

\[ E_\theta W_2^2 \]

\[ = E_\theta \left( \exp \left( (N - M) \log \frac{\int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta)u) \nu(ds)}{\int_{A_j} \lambda(s; \theta) \nu(ds)} \right) \right. \]

\[ - \int_{A_j} \log \frac{\lambda(s; \theta + \phi_{A_j}(\theta)u)}{\lambda(s; \theta)} (N - M)(ds) \right) \]
\[ E_{\theta} \left( \exp \left( (N - M) \log \int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta) u) \nu(ds) \right. \right. \]
\[ \left. \left. - \int_{A_j} \log \left( \frac{\lambda(s; \theta + \phi_{A_j}(\theta) u)}{\lambda(s; \theta)} \right) (N - M)(ds) \right| N, M \right) \right) \]
\[ = E_{\theta} \left( \left( \int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta) u) \nu(ds) \right)^{N-M} \right. \]
\[ \left. \left( \int_{A_j} \frac{\lambda(s; \theta)}{\lambda(s; \theta + \phi_{A_j}(\theta) u)} \int_{A_j} \frac{\lambda(s; \theta)}{\lambda(s; \theta + \phi_{A_j}(\theta) u)} \nu(ds) \right)^{N-M} \right) \]
\[ = E_{\theta} \left( \left( W_2' \right)^{N-M} \right) \leq E_{\theta} \left( \left( W_2' \right)^N \right), \]

since \( N - M \geq 0 \) and \( W_2' \geq 1 \) by Casella and Berger’s [2] lemma. Here,
\[ W_2' = \int_{A_j} \lambda(s; \theta + \phi_{A_j}(\theta) u) \nu(ds) \int_{A_j} \frac{(\lambda(s; \theta))^2}{\lambda(s; \theta + \phi_{A_j}(\theta) u)} \nu(ds) \]
\[ = \int_{A_j} \lambda(s; \theta) \nu(ds) \int_{A_j} \frac{(\lambda(s; \theta))^2}{\lambda(s; \theta + \phi_{A_j}(\theta) u)} \nu(ds) \]

Since we know that \( N \sim \text{Poisson} \left( \int_{A_j} \lambda(s; \theta) \nu(ds) \right) \), we have
\[ E_{\theta} \left( \left( W_2' \right)^N \right) = E_{\theta} \left( \exp \left( N \log \left( W_2' \right) \right) \right) = \text{MGF}_{\text{Poisson}} \log \left( W_2' \right) \]
\[ = \exp \left( \int_{A_j} \lambda(s; \theta) \nu(ds) (W_2' - 1) \right) \tag{17} \]

Define \( U_1 = \{ u : |u| < \mu(A_j) \} \) and \( U_2 = \{ u : |u| > \mu(A_j) \} \cap \{ u : \theta + \phi_{A_j}(\theta) u \in \Theta \} \). By the similar strategy in Rathbun and Cressie’s [20] Lemma 10, for \( u \in U_1 \), we know that for large \( j \), i.e., there exists a \( j_1 \), when \( j \geq j_1 \), (16) is as follows,
\[ \exp \left( -\frac{1}{2} \int_{A_j} \left( \frac{(\lambda(s; \theta + \phi_{A_j}(\theta) u))^2}{\lambda(s; \theta)} - (\lambda(s; \theta))^{1/2} \right)^2 \nu(ds) \right) \]
\[ \leq \exp \left( -\frac{1}{8} \int_{A_j} \left( u' \phi_{A_j}(\theta)' \lambda(s; \theta) \right) (\lambda(s; \theta))^{1/2} \nu(ds) \right), \]
thus, we have

\[
(E_\theta W_1)^{1/2} = \exp \left(-\frac{1}{4} \int_{A_j} \left( \left( \lambda(s; \theta + \phi_{A_j}(\theta) u) \right)^{1/2} - \left( \lambda(s; \theta) \right)^{1/2} \right)^2 \nu(ds) \right)
\]

\[
\leq \exp \left(-\frac{1}{16} \int_{A_j} \left( u' \phi_{A_j}(\theta)' \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)^{1/2}} \right)^2 \nu(ds) \right). \tag{18}
\]

Now, for large \( j \), i.e., there exists a \( j_2 \), when \( j \geq j_2 \), by Taylor expansion and \((C_7)\), let’s look at \( \int_{A_j} \lambda(s; \theta) \nu(ds) (W'_2 - 1) \) as follows,

\[
\int_{A_j} \lambda(s; \theta) \nu(ds) (W'_2 - 1) 
\approx \int_{A_j} \lambda(s; \theta) \nu(ds) - \int_{A_j} \lambda(s; \theta) (\log \frac{\lambda(s; \theta)}{\lambda(s; \theta + \phi_{A_j}(\theta) u)} + 1) \nu(ds) 
\leq \frac{1}{2} \int_{A_j} \left( u' \phi_{A_j}(\theta)' \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)^{1/2}} \right)^2 \nu(ds), \text{ by } B_3 \rightarrow B_1,
\]

thus, we have

\[
(E_\theta (W^2_2))^{1/4} \leq \left( E_\theta \left( (W'_2)^N \right) \right)^{1/4} = \left( \exp \left( \int_{A_j} \lambda(s; \theta) \nu(ds) (W'_2 - 1) \right) \right)^{1/4}
\]

\[
\approx \exp \left( \frac{1}{8} \int_{A_j} \left( u' \phi_{A_j}(\theta)' \frac{\dot{\lambda}(s; \theta)}{\lambda(s; \theta)^{1/2}} \right)^2 \nu(ds) \right). \tag{19}
\]

By (18) and (19) and assumption \((C_2)\), there exists a constant \( c_2 \), such that
\[(E_\theta W_1)^{\frac{1}{2}} (E_\theta W_2^{2})^{\frac{1}{4}} \leq \exp \left( \frac{1}{16} \int_{A_j} \left( u' \phi_{A_j}(\theta) \frac{\lambda(s; \theta)}{(\lambda(s; \theta))^{1/2}} \right)^2 \nu(ds) \right) \leq \exp \left( \frac{c_2}{16} |u|^2 \right). \]

Hence, for all \( j \geq \max (j_1, j_2), \)
\[P_{\theta}^{A_j} \{ Z_{A_j}(u) > \exp (-c_1 g(u)) \} \leq \exp \left( \frac{c_1}{4} g(u) \right) (E_\theta W_1)^{\frac{1}{2}} (E_\theta W_2^{2})^{\frac{1}{4}} \]
\[\leq \exp \left( \left( \frac{c_1}{2} - \frac{c_2}{16} \right) g(u) \right).\]

This means that the lemma holds for \( c_1 = \frac{c_2}{24}, \) where \( u \in U_1. \)

Now, for \( u \in U_2, \) by the boundness of \( \Theta, \) there exists a constant \( c_3 \) such that \( |u| < c_3 |\phi_{A_j}(\theta)|^{-1} \), for all \( \theta \in K. \) Let’s denote the lower bound in assumption \((C_5)\) by \( \chi. \) Then, by \((C_5),\) there exists \( \alpha > 0 \) and \( j_3 \leq 1, \) such that for all \( j \geq j_3, \)
\[\int_{A_j} \left( (\lambda(s; \theta + \phi_{A_j}(\theta)u))^{\frac{1}{2}} - (\lambda(s; \theta))^{\frac{1}{2}} \right)^2 \nu(ds) \]
\[= \frac{1}{4} \int_{A_j} (\psi(s; \theta + \phi_{A_j}(\theta)u) - \psi(s; \theta))^2 \nu(ds) \]
\[= \frac{1}{4} \left| |\psi(s; \theta + \phi_{A_j}(\theta)u) - \psi(s; \theta)|^2 \right| \]
\[\text{assumption(C}_5) \]
\[\geq \frac{1}{4} |\phi_{A_j}(\theta)|^{-2\alpha} \chi^2 \]
\[\geq \frac{1}{4} |u|^{-2\alpha} |c_3|^{-2\alpha} \chi^2 = \frac{1}{4} |c_3|^{-2\alpha} \chi^2 |u|^{2\alpha} = c_k |u|^{2\alpha}, \]

where \( c_k = \frac{1}{4} |c_3|^{-2\alpha} \chi^2. \) Then, for all \( j > j_3, \)
\[P_{\theta}^{A_j} \{ Z_{A_j}(u) > \exp (-c_1 g(u)) \} \leq \exp \left( \frac{c_1}{2} g(u) \right) E_\theta \left( (Z_{A_j}(u))^{\frac{1}{2}} \right) \]
\[= \exp \left( \frac{c_1}{2} g(u) \right) \exp \left( -\frac{1}{2} \int_{A_j} \left( (\lambda(s; \theta + \phi_{A_j}(\theta)u))^{\frac{1}{2}} - (\lambda(s; \theta))^{\frac{1}{2}} \right)^2 \nu(ds) \right) \]
\[\leq \exp \left( \frac{c_1}{2} g(u) \right) \exp \left( -\frac{c_k}{2} |u|^{2\alpha} \right) = \exp \left( \left( \frac{c_1}{2} - \frac{c_k}{2} \right) g(u) \right). \]

This means that the lemma holds for \( c_1 = \frac{c_k}{2}, \) where \( u \in U_2. \) Finally, the lemma follows by choosing \( j_0 = \max (j_1, j_2, j_3) \) and \( c_1 = \min \left( \frac{c_k}{24}, \frac{c_k}{2} \right). \)