doi:10.5705/ss.202014.0021

ASYMPTOTIC NORMALITY OF NONPARAMETRIC M-ESTIMATORS WITH APPLICATIONS TO HYPOTHESIS TESTING FOR PANEL COUNT DATA

Xingqiu Zhao and Ying Zhang

The Hong Kong Polytechnic University and Indiana University

Abstract: In semiparametric and nonparametric statistical inference, the asymptotic normality of estimators has been widely established when they are $\sqrt{n}$-consistent. In many applications, nonparametric estimators are not able to achieve this rate. We have a result on the asymptotic normality of nonparametric M-estimators that can be used if the rate of convergence of an estimator is $n^{-1/2}$ or slower. We apply this to study the asymptotic distribution of sieve estimators of functionals of a mean function from a counting process, and develop nonparametric tests for the problem of treatment comparison with panel count data. The test statistics are constructed with spline likelihood estimators instead of nonparametric likelihood estimators. The new tests have a more general and simpler structure and are easy to implement. Simulation studies show that the proposed tests perform well even for small sample sizes. We find that a new test is always powerful for all the situations considered and is thus robust. For illustration, a data analysis example is provided.

Key words and phrases: Asymptotic normality; M-estimators; Nonparametric maximum likelihood; Nonparametric maximum pseudo-likelihood; Nonparametric tests; Spline.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

1 Introduction

Weak convergence theory and empirical theory (van der Vaart and Wellner (1996)) have been widely used to study the asymptotic properties of estimators in semiparametric and nonparametric models. When the convergence rate of estimators is $n^{-1/2}$, the asymptotic distribution of the estimators can be derived by using the weak convergence theorem on Z-estimators (van der Vaart and Wellner (1996, page 310)). For example, Zeng, Lin, and Yin (2005) and Zeng and Lin (2006, 2007) obtained the desired asymptotic normality of the estimators for the proportional odds model and the semiparametric transformation models by verifying the conditions of Theorem 3.3.1 of van der Vaart and Wellner (1996). When the convergence rate is slower than $n^{-1/2}$, for example, the convergence rates of nonparametric maximum likelihood estimators of cumulative distribution functions based on interval-censored data are $n^{-1/3}$ (Groeneboom and Wellner (1992)), this theorem is no longer applicable. For such situations, it is difficult to derive the asymptotic distribution of nonparametric estimators. Zhang (2006) and Balakrishnan and Zhao (2009) investigated the asymptotic normality of functionals of the nonparametric maximum pseudo-likelihood and likelihood estimators for panel count data. We have a general theorem dealing with the asymptotic normality of nonparametric M-estimators and we apply this to panel count models as illustrative examples.

For the nonparametric inference of panel count data, several estimation and testing methods have been developed. Sun and Kalbfeisch (1995), Wellner and Zhang (2000), Lu, Zhang, and Huang (2007), and Hu, Lagakos, and Lockhart (2009a, b) studied the nonparametric estimation of the mean function of the underlying counting process with panel count data by using isotonic regression techniques, the likelihood approach, the spline likelihood approach,

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

the estimating equation approach, and the generalized least squares method, respectively; Thall and Lachin (1988), Sun and Fang (2003), Zhang (2006), and Balakrishnan and Zhao (2009) presented some nonparametric tests for nonparametric comparison of mean function of counting process with panel count data. For a comprehensive review about the analysis of panel count data, see Sun and Zhao (2013).

Lu, Zhang, and Huang (2007, 2009) showed that the spline likelihood estimators have a convergence rate slower than $n^{-1/2}$ but faster than $n^{-1/3}$, and are more efficient both statistically and computationally than the nonparametric maximum likelihood estimators in simulations. In this paper, we explore asymptotic normality of functionals of spline likelihood estimators of mean functions, and propose some new nonparametric tests based on them to compare with existing tests for the nonparametric comparison of counting processes with panel count data.

The remainder of the article is organized as follows. In Section 2, we present a general theorem regarding the asymptotic normality of nonparametric M-estimators. In Section 3, we briefly review the nonparametric spline-based likelihood estimators for panel count data and establish the asymptotic normality of their functionals. Section 4 presents two classes of nonparametric test statistics for comparing two treatment groups with respect to their mean functions. The asymptotic normality of the proposed test statistics are established. Section 5 reports some simulation results to assess the finite-sample properties of the proposed test procedure and to compare them with the tests based on the nonparametric likelihood estimators. A data analysis example is provided in Section 6. Section 7 contains some concluding remarks. The proofs of theorems are given in the Supplementary Materials available at the *Statistica Sinica* journal website.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

2 Asymptotic Distributional Theory of Nonparametric M-Estimators

Suppose $\mathbf{X} = (X_1, \ldots, X_n)$ is a random sample taken from the distribution of X, and $l_n(\Lambda; \mathbf{X}) = \sum_{i=1}^n m(\Lambda; X_i)$ is an objective function based on $\mathbf{X}$, where Λ is an unknown function in the class $\mathcal{F}$. Let $\mathcal{F}_n$ be the sieve parameter space satisfying

$$\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \cdots \subseteq \mathcal{F}, \quad \text{for } n \geq 1.$$

Assume that $\hat{\Lambda}_n$ is the estimator of Λ_0 that maximizes $l_n(\Lambda; \mathbf{X})$ with respect to $\mathcal{F}_n$.

Suppose Λ_η is a parametric path in $\mathcal{F}$ through Λ, $\Lambda_\eta \in \mathcal{F}$ and $\Lambda_\eta|_{\eta=0} = \Lambda$. Let $\mathcal{H} = \{h : h = \frac{\partial \Lambda_\eta}{\partial \eta}|_{\eta=0}\}$ and $l^\infty(\mathcal{H})$ be the space of bounded functionals on $\mathcal{H}$ under the supermum norm $||f||_\infty = \sup_{h \in \mathcal{H}} |f(h)|$. For $h \in \mathcal{H}$, we define a sequence of maps G_n of a neighborhood of Λ_0, denoted by $\mathcal{U}$, in the parameter space for Λ into $l^\infty(\mathcal{H})$ by

$$\begin{aligned} G_n(\Lambda)[h] &= n^{-1} \frac{\partial}{\partial \eta} l_n(\Lambda_\eta; \mathbf{X})|_{\eta=0} \\ &= n^{-1} \sum_{i=1}^n \frac{\partial}{\partial \eta} m(\Lambda_\eta; X_i)|_{\eta=0} \\ &\equiv \mathbb{P}_n \psi(\Lambda; X)[h], \end{aligned}$$

and take $G(\Lambda)[h] = P\psi(\Lambda; X)[h]$, where P and $\mathbb{P}_n$ denote the probability measure and empirical measure with $Pf = \int f dP$ and $\mathbb{P}_n f = n^{-1} \sum_{i=1}^n f(X_i)$, respectively.

To establish the asymptotic normality, we need the following conditions.

A1. $\sqrt{n}(G_n - G)(\hat{\Lambda}_n)[h] - \sqrt{n}(G_n - G)(\Lambda_0)[h] = o_p(1)$.

A2. $G(\Lambda_0)[h] = 0$ and $G_n(\hat{\Lambda}_n)[h] = o_p(n^{-1/2})$.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

A3. $\sqrt{n}(G_n - G)(\Lambda_0)[h]$ converges in distribution to a tight Gaussian process on $l^\infty(\mathcal{H}_r)$.

A4. $G(\Lambda)[h]$ is Fréchet-differentiable at Λ_0 with a continuous derivative $\dot{G}_{\Lambda_0}[h]$.

A5. $G(\hat{\Lambda}_n)[h] - G(\Lambda_0)[h] - \dot{G}_{\Lambda_0}(\hat{\Lambda}_n - \Lambda_0)[h] = o_p(n^{-1/2})$.

Theorem 2.1 *If A1-A5 hold, then for any $h \in \mathcal{H}$,*

$$-\sqrt{n}\dot{G}_{\Lambda_0}(\hat{\Lambda}_n - \Lambda_0)[h] = \sqrt{n}(G_n - G)(\Lambda_0)[h] + o_p(1).$$

REMARK 1. The above theorem does not require the $\hat{\Lambda}_n$ be $\sqrt{n}$-consistent, while the conditions stated in Theorem 3.3.1 of van der Vaart and Wellner (1996) imply that the estimator has the usual convergence rate $n^{-1/2}$.

REMARK 2. Assumptions A2-A4 are the analytical conditions given in Theorem 3.3.1 of van der Vaart and Wellner (1996). Assumptions A1 and A5 require the remainder in a Taylor expansion be negligible; they are weaker than those required by van der Vaart and Wellner (1996).

The theorem can be widely used to establish the asymptotic normality of nonparametric estimators no matter whether the rate of convergence is $n^{-1/2}$, or is slower. We focus on counting process models with panel count data to illustrate applications of the theorem.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

3 Asymptotic Normality of Functionals of Nonparametric Spline-based Likelihood Estimators for Panel Count Data

3.1 Nonparametric Spline-based Likelihood Estimation

Consider a recurrent event study that consists of n independent subjects and let $N_i(t)$ denote the number of occurrences of the recurrent event of interest before or at time t for subject i. For subject i, suppose that $N_i(\cdot)$ is observed only at finite time points $T_{K_i,1} < \cdots < T_{K_i,K_i} \leq \tau$, where K_i denotes the potential number of observation times, $i = 1, \ldots, n$, and τ is the length of the study.

In the following, we assume that $(K_i; T_{K_i,1}, ..., T_{K_i,K_i})$ are independent of the counting processes N_i's. Let $\mathbf{X} = (K, \mathbf{T}, \mathbf{N})$, where $\mathbf{T} = (T_{K,1}, ..., T_{K,K})$ and $\mathbf{N} = (N(T_{K,1}), \ldots, N(T_{K,K}))$. Then $\{\, \mathbf{X}_i = (K_i, \mathbf{T}_i, \mathbf{N}_i),\ i = 1, ..., n\}$ is a random sample of size n from the distribution of $\mathbf{X}$, where $\mathbf{T}_i = (T_{K_i,1}, ..., T_{K_i,K_i})$ and $\mathbf{N}_i = (N_i(T_{K_i,1}), \ldots, N_i(T_{K_i,K_i}))$.

Suppose that for each subject, $N_i(t)$ is a non-homogeneous Poisson process with the mean function $\Lambda(t)$. The log pseudo-likelihood and the log-likelihood functions for Λ are

$$l_n^{ps}(\Lambda) = \sum_{i=1}^{n} \sum_{j=1}^{K_i} \left[N_i(T_{K_i,j}) \log \left\{ \Lambda(T_{K_i,j}) \right\} - \Lambda(T_{K_i,j}) \right],$$

$$l_n(\Lambda) = \sum_{i=1}^{n} \sum_{j=1}^{K_i} \left[\Delta N_i(T_{K_i,j}) \log \left\{ \Delta\Lambda(T_{K_i,j}) \right\} - \Delta\Lambda(T_{K_i,j}) \right],$$

after omitting the parts independent of Λ, where $T_{K_i,0} = 0$, $\Delta\Lambda(T_{K_i,j}) = \Lambda(T_{K_i,j}) -$

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

$\Lambda(T_{K_i,j-1})$, and $\Delta N_i(T_{K_i,j}) = N_i(T_{K_i,j}) - N_i(T_{K_i,j-1})$.

For estimation of the smooth function $\Lambda_0(t)$, we use B-spline function approximation (Lu, Zhang, and Huang (2007)). Let $\mathcal{T} = \{s_i, i = 1, \ldots, m_n + 2l\}$, with

$$\tau_0 = s_1 = \cdots = s_l < s_{l+1} < \cdots < s_{m_n+l} < s_{m_n+l+1} = \cdots = s_{m_n+2l} = \tau,$$

be a sequence of knots that partition $[\tau_0, \tau]$ into $m_n + 1$ subintervals $I_i = [s_{l+i}, s_{l+i+1}]$, for $i = 0, 1, \ldots, m_n$. Let Φ_n be the class of polynomial splines of order $l \geq 1$ with the knot sequence $\mathcal{T}$. Then Φ_n can be linearly spanned by the normalized B-spline basis functions $\{B_i, i = 1, \ldots, \alpha_{q_n}\}$ with $q_n = m_n + l$ (Schumaker (1981)). Define a subclass of Φ_n,

$$\Psi_n = \left\{ \sum_{i=1}^{q_n} \alpha_i B_i : 0 \leq \alpha_1 \leq \cdots \leq \alpha_{q_n} \right\}.$$

Following Lu, Zhang and Huang (2007), the estimators $\hat{\Lambda}_n^{ps}$ and $\hat{\Lambda}_n$ are the values that maximize $l_n^{ps}(\Lambda)$ and $l_n(\Lambda)$ with respect to $\Lambda \in \Psi_n$, respectively.

We denote the spline pseudo-likelihood and spline likelihood estimators of Λ by $\hat{\Lambda}_n^{ps} = \sum_{i=1}^{q_n} \hat{\alpha}_{in}^{ps} B_i$ and $\hat{\Lambda}_n = \sum_{i=1}^{q_n} \hat{\alpha}_{in} B_i$.

3.2 Asymptotic Normality

Let $\mathcal{B}$ denote the collection of Borel sets in $\mathbb{R}$, and let $\mathcal{B}_{[0,\tau]} = \{B \cap [0, \tau] : B \in \mathcal{B}\}$. Following Wellner and Zhang (2000), define measures μ_1 and μ_2 as follows: for $B, B_1, B_2 \in \mathcal{B}_{[0,\tau]}$,

$$\begin{aligned} \mu_1(B) &= \sum_{k=1}^{\infty} P(K = k) \sum_{j=1}^{k} P(T_{k,j} \in B | K = k) \\ &= E\left\{ \sum_{j=1}^{K} I(T_{K,j} \in B) \right\}, \end{aligned}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

$$\begin{aligned}\mu_2(B_1 \times B_2) &= \sum_{k=1}^{\infty}\{P(K=k)\sum_{j=1}^{k} P(T_{k,j-1} \in B_1, T_{k,j} \in B_2 | K=k)\} \\ &= E\left\{\sum_{j=1}^{K} I(T_{K,j-1} \in B_1, T_{K,j} \in B_2)\right\}.\end{aligned}$$

Define the L_2-metrics d_1 and d_2 as

$$d_1(\Lambda_1, \Lambda_2) = \left\{\int |\Lambda_1(t) - \Lambda_2(t)|^2 d\mu_1(t)\right\}^{1/2},$$

$$d_2(\Lambda_1, \Lambda_2) = \left\{\int\int |(\Lambda_1(u) - \Lambda_1(v)) - (\Lambda_2(u) - \Lambda_2(v))|^2 d\mu_2(u,v)\right\}^{1/2}.$$

To establish the asymptotic properties of the estimators, we need the following regularity conditions.

C1. The maximum spacing of the knots, $\Delta \equiv \max_{l+1 \le i \le m_n+l+1} |s_i - s_{i-1}| = O(n^{-v})$ with $m_n = O(n^v)$ for $0 < v < 0.5$. There exists a constant $M > 0$ such that $\Delta/\delta \le M$ uniformly in n, where $\delta \equiv \min_{l+1 \le i \le m_n+l+1} |s_i - s_{i-1}|$.

C2. The true mean function Λ_0 is a nondecreasing function over $[0, \tau]$ with $\Lambda(0) = 0$, with a bounded rth derivative in $[0, \tau]$ for $r \ge 1$, and $\Lambda_0'(t) \ge a_0$ for some $a_0 \in (0, \infty)$.

C3. There exists a positive integer K_0 such that $P(K \le K_0)$=1.

C4. For some positive constant k_0, $E[\exp\{k_0 N(\tau)\}] < \infty$.

C5. $P(\cap_{j=1}^{K}\{T_{K,j} \in [\tau_0, \tau]\}) = 1$ with $\tau_0 > 0$, $\Lambda_0(\tau_0) > 0$, and $\Lambda_0(\tau) \le M_0$ for some constant $M_0 > 0$.

C6. $\mu_1(\tau_0) > 0$, and for all $\tau_0 < \tau_1 < \tau_2 < \tau$, $\mu_1((\tau_1, \tau_2)) > 0$.

C7. There exists a positive constant s_0 such that $P(\min_{1 \le j \le K}\{T_{K,j} - T_{K,j-1}\} \ge s_0) = 1$.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

C8. μ_1 is absolutely continuous with respect to Lebesgue measure, with derivative $\dot{\mu}_1$.

C9. μ_2 is absolutely continuous with respect to Lebesgue measure, with derivative $\dot{\mu}_2$.

C10. If with probability 1, $h(T_{K,j}) = 0, j = 1, \ldots, K$ for some h, then $h = 0$.

Conditions C1-C5 and C7 are required by Lu, Zhang and Huang (2007); condition C6 is required by Balakrishnan and Zhao (2009). Conditions C8 and C9 are similar to C11 in Wellner and Zhang (2007). Condition C10 is needed for identifiability of the model.

Theorem 3.1 *Suppose C1-C6 and C10 hold, and let*

$$\mathcal{H}_r = \left\{ g(\cdot) : |g^{(r-1)}(s) - g^{(r-1)}(t)| \leq c_0 |s - t| \text{ for all } \tau_0 \leq s, t \leq \tau \right\}.$$

where $g^{(r-1)}$ is the $(r-1)$th derivative function of g, and c_0 is a constant.

(i) If C8 holds, then for $h \in \mathcal{H}_r$,

$$\sqrt{n} \int \{\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)\} dh(t) \to_d N(0, \sigma_{ps}^2), \tag{3.1}$$

where σ_{ps}^2 is given at (S2.4) of Supplementary Materials.

(ii) If C7 and C9 hold, then for $h \in \mathcal{H}_r$,

$$\sqrt{n} \int \{\hat{\Lambda}_n(t) - \Lambda_0(t)\} dh(t) \to_d N(0, \sigma^2), \tag{3.2}$$

where σ^2 is given at (S2.7) of Supplementary Materials.

Corollary 3.1 *Suppose the conditions in Theorem 3.1 hold.*

(i)

$$\sqrt{n} \int h(t) \frac{\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)}{\Lambda_0(t)} d\mu_1(t) \to_d N(0, \sigma_1^2), \tag{3.3}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

where $h \in \mathcal{H}_r$, *and*

$$\sigma_1^2 = E\left[\sum_{j=1}^{K} h(T_{K,j}) \frac{N(T_{K,j}) - \Lambda_0(T_{K,j})}{\Lambda_0(T_{K,j})}\right]^2. \tag{3.4}$$

(ii)

$$\sqrt{n}\int \{h(u) - h(v)\} \frac{\{\hat{\Lambda}_n(u) - \hat{\Lambda}_n(v)\} - \{\Lambda_0(u) - \Lambda_0(v)\}}{\{\Lambda_0(u) - \Lambda_0(v)\}} d\mu_2(u, v)$$

$$\rightarrow_d \ N(0, \sigma_2^2), \tag{3.5}$$

where $h \in \mathcal{H}_r$, *and*

$$\sigma_2^2 = E\left[\sum_{j=1}^{K} \Delta h(T_{K,j}) \frac{\Delta N(T_{K,j}) - \Delta\Lambda_0(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right]^2. \tag{3.6}$$

REMARK 3. These results can be used to construct new tests for the problem of multi-sample nonparametric comparison of counting processes with panel count data.

REMARK 4. We can show that under some regularity conditions, (3.1)-(3.6) hold for the two nonparametric likelihood-based estimators proposed by Wellner and Zhang (2000).

4 Nonparametric Two-sample Tests

Consider a longitudinal study with some recurrent event and n independent subjects from two groups, n_l in the lth group with $n_1+n_2 = n$. Let $N_i(t)$ denote the counting process arising from subject i and $\Lambda_l(t)$ denote the mean function of the counting process corresponding to group l, $l = 1, 2$. Here, the problem of interest is to test the null hypothesis $H_0 : \Lambda_1(t) = \Lambda_2(t)$. Suppose subject i is observed only at distinct time points $0 < T_{K_i,1} < \cdots < T_{K_i,K_i}$ and that no information is available about $N_i(t)$ between them.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

Let $\hat{\Lambda}_l^{ps}$ and $\hat{\Lambda}_l$ denote the spline pseudo-likelihood and spline likelihood estimators of Λ_l based on samples from all the subjects in the lth group. Let $\Lambda_0(t)$ denote the common mean function of the $N_i(t)$'s under H_0, and let $\hat{\Lambda}_0^{ps}$ and $\hat{\Lambda}_0$ be the spline pseudo-likelihood and spline likelihood estimators of Λ_0 based on the pooled data. Clearly, $\mu_1(t)$ and $\mu_2(u,v)$ can be consistently estimated by

$$\hat{\mu}_1(t) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K_i} I(T_{K_i,j} \le t),$$

$$\hat{\mu}_2(u,v) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K_i} I(T_{K_i,j-1} \le v, T_{K_i,j} \le u),$$

respectively.

To test the hypothesis H_0, Zhang (2006) and Balakrishnan and Zhao (2009) proposed to use the two statistics

$$U_Z = \sqrt{n}\int_0^{\tau} W_n(t)\{\hat{\Lambda}_{1,mple}(t) - \hat{\Lambda}_{2,mple}(t)\}d\hat{\mu}_1(t),$$

$$\begin{aligned} U_{BZ} \;=\; & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Bigg[\sum_{j=1}^{K_i-1} W_n(T_{K_i,j})\hat{\Lambda}_{0,mle}(T_{K_i,j}) \\ & \times\Bigg\{\left(\frac{\Delta\hat{\Lambda}_{1,mle}(T_{K_i,j+1})}{\Delta\hat{\Lambda}_{0,mle}(T_{K_i,j+1})} - \frac{\Delta\hat{\Lambda}_{1,mle}(T_{K_i,j})}{\Delta\hat{\Lambda}_{0,mle}(T_{K_i,j})}\right) \\ & \quad - \left(\frac{\Delta\hat{\Lambda}_{2,mle}(T_{K_i,j+1})}{\Delta\hat{\Lambda}_{0,mle}(T_{K_i,j+1})} - \frac{\Delta\hat{\Lambda}_{2,mle}(T_{K_i,j})}{\Delta\hat{\Lambda}_{0,mle}(T_{K_i,j})}\right)\Bigg\} \\ & + W_n(T_{K_i,K_i})\hat{\Lambda}_{0,mle}(T_{K_i,K_i}) \\ & \times\Bigg\{\left(1 - \frac{\Delta\hat{\Lambda}_{1,mle}(T_{K_i,K_i})}{\Delta\hat{\Lambda}_{0,mle}(T_{K_i,K_i})}\right) - \left(1 - \frac{\Delta\hat{\Lambda}_{2,mle}(T_{K_i,K_i})}{\Delta\hat{\Lambda}_{0,mle}(T_{K_i,K_i})}\right)\Bigg\}\Bigg], \end{aligned}$$

where $W_n(t)$ is a bounded weight process, $\hat{\Lambda}_{l,mple}$ and $\hat{\Lambda}_{l,mle}$ denote the maximum pseudo-likelihood and maximum likelihood estimators of Λ_l based on samples from all the subjects in the lth group, and $\hat{\Lambda}_{0,mle}$ denotes the maximum likelihood estimator of Λ_0 based on the

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

pooled data. We propose the two test statistics

$$\begin{aligned} U_n^{ps} &= \sqrt{n}\int h_n^{ps}(t)\frac{\hat{\Lambda}_1^{ps}(t)-\hat{\Lambda}_2^{ps}(t)}{\hat{\Lambda}_0^{ps}(t)}d\hat{\mu}_1(t) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{K_i} h_n^{ps}(T_{K_i,j})\frac{\hat{\Lambda}_1^{ps}(T_{K_i,j})-\hat{\Lambda}_2^{ps}(T_{K_i,j})}{\hat{\Lambda}_0^{ps}(T_{K_i,j})}, \end{aligned}$$

$$\begin{aligned} U_n &= \sqrt{n}\int \{h_n(u)-h_n(v)\}\frac{\{\hat{\Lambda}_1(u)-\hat{\Lambda}_1(v)\}-\{\hat{\Lambda}_2(u)-\hat{\Lambda}_2(v)\}}{\{\hat{\Lambda}_0(u)-\hat{\Lambda}_0(v)\}}d\hat{\mu}_2(u,v) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{K_i}\{\Delta h_n(T_{K_i,j})\}\frac{\Delta\hat{\Lambda}_1(T_{K_i,j})-\Delta\hat{\Lambda}_2(T_{K_i,j})}{\Delta\hat{\Lambda}_0(T_{K_i,j})}, \end{aligned}$$

where $h_n^{ps}(t)$ and $h_n(t)$ are bounded weight processes. For the propose of comparison, we consider three choices of the weight processes $h_n^{ps}(t)$ and $h_n(t)$: $h_n^{ps}(t) = \hat{\Lambda}_0^{ps}(t)W_n^{(k)}(t)$ and $h_n(t) = \hat{\Lambda}_0(t)W_n^{(k)}(t)$, where $W^{(1)}(t) = 1$, $W_n^{(2)}(t) = \sum_{i=1}^n I(t \le T_{K_i,K_i})$, and $W_n^{(3)}(t) = 1 - W_n^{(2)}(t) = \sum_{i=1}^n I(t > T_{K_i,K_i})$. Other choices of weight processes can be made. For example, if we take $h_n(t) = \{\hat{\Lambda}_0(t)\}^2$, the structure of U_n^{ps} is similar to U_Z, while the structure of U_n is much simpler than that of U_{BZ}.

Theorem 4.1 *Suppose the conditions of Theorem 3.1 hold. Suppose that $h_n(t)$'s are bounded weight processes and that there exists a bounded function $h(t)$ such that $h \in \mathcal{H}_r$, and*

$$\left[\int_0^\tau \{h_n(t) - h(t)\}^2 d\mu_1(t)\right]^{1/2} = o_p(n^{-\frac{1}{2(1+2r)}}) .$$

If $n_1/n \to p$ as $n \to \infty$, where $0 < p < 1$, then, under $H_0 : \Lambda_1 = \Lambda_2 = \Lambda_0$,

(i) U_n^{ps} has an asymptotic normal distribution $N(0, \sigma_{ps}^2)$, where

$$\sigma_{ps}^2 = \left(\frac{1}{p} + \frac{1}{1-p}\right)\sigma_1^2$$

with σ_1^2 as given as (3.4);

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

(ii) U_n *has an asymptotic normal distribution* $N(0, \sigma^2)$*, where*

$$\sigma^2 = \left(\frac{1}{p} + \frac{1}{1-p}\right)\sigma_2^2$$

with σ_2^2 *as given as (3.6);*

(iii) If

$$\max_{1\le i\le n} E\left[\sum_{j=1}^{K_i}\{h_n(T_{K_i,j}) - h(T_{K_i,j})\}^2\right] \longrightarrow 0,$$

then σ_1^2*,* σ_2^2 *can be consistently estimated by*

$$\hat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\sum_{j=1}^{K_i} h_n(T_{K_i,j})\frac{N_i(T_{K_i,j}) - \hat{\Lambda}_0^{ps}(T_{K_i,j})}{\hat{\Lambda}_0^{ps}(T_{K_i,j})}\right]^2,$$

$$\hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\sum_{j=1}^{K_i} \Delta h_n(T_{K_i,j})\frac{\Delta N_i(T_{K_i,j}) - \Delta\hat{\Lambda}_0(T_{K_i,j})}{\Delta\hat{\Lambda}_0(T_{K_i,j})}\right]^2, \quad \textit{respectively.}$$

Remark 5. For the asymptotic normality of the proposed test statistics, we do not need the condition that $h \circ \Lambda_0^{-1}$ is a bounded Lipschitz function as required by Zhang (2006) and Balakrishnan and Zhao (2009).

Remark 6. We can show that, under some regularity conditions, (i)-(iii) hold if the spline pseudo-likelihood and spline likelihood estimators in the expression of U_n^{ps} and U_n are replaced with the nonparametric maximum pseudo-likelihood and nonparametric maximum likelihood estimators proposed by Wellner and Zhang (2000), respectively.

5 Simulation Study

We conducted a simulation study to investigate the finite-sample properties of the proposed test statistics and to make comparisons with those of the tests presented by Zhang (2006)

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

and Balakrishnan and Zhao (2009). We let $T^{ps} = U_n^{ps}/\hat{\sigma}_{ps}$ and $T = U_n/\hat{\sigma}$, where

$$\hat{\sigma}_{ps} = \left\{ \left(\frac{n}{n_1} + \frac{n}{n_2} \right) \hat{\sigma}_1^2 \right\}^{1/2},$$

$$\hat{\sigma} = \left\{ \left(\frac{n}{n_1} + \frac{n}{n_2} \right) \hat{\sigma}_2^2 \right\}^{1/2},$$

and U_n^{ps}, U_n, and $\hat{\sigma}_l^2$ be as given in Section 4. By Theorem 4.1, the null hypothesis can be tested by T^{ps} and T, which have asymptotic standard normal distributions. For the generation of panel count data, denoted by $\{k_i, t_{ij}, n_{ij}, j = 1, \ldots, k_i, i = 1, \ldots, n\}$, we first generated the number of observation times k_i from the uniform distribution $U\{1, \ldots, 10\}$, and then, given k_i, we took the observation times t_{ij}'s to be the order statistics of a random sample of size k_i drawn from $U\{1, \ldots, 12\}$. To generate the n_{ij}'s, we assumed that, given a nonnegative random variable γ_i, $N_i(t)$ is a Poisson process with mean function $\Lambda_i(t|\gamma_i) = E(N_i(t)|\gamma_i)$. Let S_l denote the set of indices for subjects in group l, $l = 1, 2$. For comparison, we considered cases representing two patterns of the mean functions:

Case 1. $\Lambda_i(t|\gamma_i) = \gamma_i t$ for $i \in S_1$, $\Lambda_i(t) = \gamma_i t \exp(\beta)$ for $i \in S_2$.

Case 2. $\Lambda_i(t|\gamma_i) = \gamma_i t$ for $i \in S_1$, $\Lambda_i(t) = \gamma_i \sqrt{\beta t}$ for $i \in S_2$.

As shown in Figures 1-2 of Balakrishnan and Zhao (2009), the two mean functions do not overlap in Case 1 and they cross over in Case 2.

For each case, we took $\gamma_i = 1$ and $\gamma_i \sim$ Gamma$(2, 1/2)$ corresponding to Poisson and mixed Poisson processes, respectively. For each setting, we took $n_1 = 30, n_2 = 50$ and $n_1 = 50, n_2 = 70$. We considered the weight processes $h_n^{(j)}(t) = \hat{\Lambda}_0(t) W_n^{(j)}(t)$, $j = 1, 2, 3$, with

$$W_n^{(1)}(t) = 1, \qquad W_n^{(2)}(t) = \frac{1}{n} \sum_{i=1}^{n} I(t \le t_{k_i, k_i}), \quad \text{and}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

$$W_n^{(3)}(t) = \frac{1}{n}\sum_{i=1}^{n} I(t > t_{k_i,k_i}),$$

and $h_n^{(4)}(t) = \{\hat{\Lambda}_0(t)\}^2$, denoting the corresponding tests by T_j with $h_n^{(j)}$ $(j = 1, 2, 3, 4)$ and T_{BZj} with $W_n^{(j)}$ $(j = 1, 2, 3)$. Here, the nonparametric maximum likelihood estimators $\hat{\Lambda}_{l,mle}$, $\hat{\Lambda}_{0,mle}$ were computed by using the modified iterative convex minorant algorithm in Wellner and Zhang (2000); the spline likelihood estimators $\hat{\Lambda}_l$ and $\hat{\Lambda}_0$ were computed by using the algorithm in Lu, Zhang and Huang (2007). The results reported here are based on 1000 Monte Carlo replications using R software.

Tables 1 and 2 present the estimated sizes and powers of the proposed test statistics T_j's and T_{BZj}'s (Balakrishnan and Zhao (2009)) at significance level $\alpha = 0.05$ for different values of β and the different weight processes based on the simulated data for the two cases with $\gamma_i = 1$ and $\gamma_i \sim$ Gamma$(2, 1/2)$, respectively. The two parts of each table include the comparison of T_j and T_{BZj} with the sample sizes $n_1 = 30, n_2 = 50$ and $n_1 = 50, n_2 = 70$ in Cases 1 and 2, respectively. To see what happens when the difference between n_1 and n_2 becomes large, we also considered the sample sizes $n_1 = 50, n_2 = 100$. The simulation results shown in the tables suggest that the tests based on the spline likelihood estimators have similar sizes and powers to those of the tests based on the nonparametric maximum likelihood estimators.

The new test procedure is easy to implement and performs well for all the situations considered. However, for Case 2 with $n_1 = 30$ and $n_2 = 50$, we note that the estimated powers of test T_{BZj}'s display "NA" often when running the simulation program. In this case, we chose to report the simulation results when the estimated powers of test T_{BZj}'s were available. It is surprising that the proposed test T_1 with the simplest structure has

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

Table 1: Estimated size and power of the tests for Poisson processes

β	T_1	T_2	T_3	T_4	T_{BZ1}	T_{BZ2}	T_{BZ3}
			$n_1 = 30, n_2 = 50$				
0.0	0.053	0.049	0.054	0.054	0.053	0.051	0.062
0.1	0.321	0.189	0.147	0.255	0.324	0.184	0.159
0.2	0.859	0.578	0.396	0.750	0.857	0.568	0.408
0.3	0.990	0.903	0.713	0.966	0.989	0.901	0.712
			$n_1 = 50, n_2 = 70$				
0.0	0.055	0.051	0.057	0.047	0.059	0.054	0.057
0.1	0.445	0.278	0.184	0.360	0.447	0.268	0.187
0.2	0.948	0.736	0.555	0.898	0.948	0.730	0.555
0.3	1.000	0.979	0.858	1.000	1.000	0.976	0.853
			$n_1 = 50, n_2 = 100$				
0.0	0.041	0.044	0.053	0.043	0.043	0.045	0.055
0.1	0.448	0.277	0.164	0.351	0.443	0.279	0.174
0.2	0.961	0.776	0.579	0.923	0.961	0.771	0.583
0.3	1.000	0.991	0.926	1.000	1.000	0.990	0.926
			$n_1 = 30, n_2 = 50$				
3	1.000	0.580	1.000	1.000	1.000	0.593	1.000
5	0.997	0.081	1.000	1.000	0.997	0.082	1.000
8	0.483	0.543	0.995	0.998	0.483	0.544	0.991
			$n_1 = 50, n_2 = 70$				
3	1.000	0.823	1.000	1.000	1.000	0.932	1.000
5	1.000	0.073	1.000	1.000	1.000	0.076	1.000
8	0.619	0.714	1.000	1.000	0.619	0.713	1.000
			$n_1 = 50, n_2 = 100$				
3	1.000	0.760	1.000	1.000	1.000	0.761	1.000
5	1.000	0.083	1.000	1.000	1.000	0.085	1.000
8	0.699	0.822	1.000	1.000	0.701	0.820	1.000

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

Table 2: Estimated size and power of the tests for mixed Poisson processes

β	T_1	T_2	T_3	T_4	T_{BZ1}	T_{BZ2}	T_{BZ3}
			$n_1 = 30, n_2 = 50$				
0.0	0.059	0.051	0.061	0.058	0.061	0.048	0.062
0.1	0.187	0.126	0.091	0.157	0.187	0.123	0.098
0.2	0.568	0.378	0.286	0.492	0.564	0.381	0.286
0.3	0.869	0.637	0.560	0.785	0.871	0.638	0.549
			$n_1 = 50, n_2 = 70$				
0.0	0.050	0.049	0.054	0.061	0.049	0.053	0.054
0.1	0.267	0.185	0.134	0.218	0.270	0.189	0.133
0.2	0.751	0.498	0.387	0.659	0.750	0.500	0.391
0.3	0.968	0.827	0.737	0.942	0.968	0.828	0.738
			$n_1 = 50, n_2 = 100$				
0.0	0.043	0.041	0.053	0.055	0.043	0.041	0.059
0.1	0.264	0.158	0.150	0.226	0.263	0.157	0.152
0.2	0.747	0.501	0.425	0.675	0.746	0.508	0.433
0.3	0.989	0.868	0.756	0.963	0.989	0.865	0.762
			$n_1 = 30, n_2 = 50$				
3	1.000	0.509	1.000	1.000	1.000	0.527	1.000
5	0.951	0.076	1.000	1.000	0.951	0.077	1.000
8	0.317	0.384	0.998	0.973	0.315	0.375	0.997
			$n_1 = 50, n_2 = 70$				
3	1.000	0.747	1.000	1.000	1.000	0.750	1.000
5	0.994	0.065	1.000	1.000	0.995	0.065	1.000
8	0.429	0.485	1.000	0.996	0.428	0.479	1.000
			$n_1 = 50, n_2 = 100$				
3	1.000	0.677	1.000	1.000	1.000	0.681	1.000
5	0.996	0.078	1.000	1.000	0.996	0.078	1.000
8	0.456	0.631	1.000	1.000	0.455	0.624	1.000

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

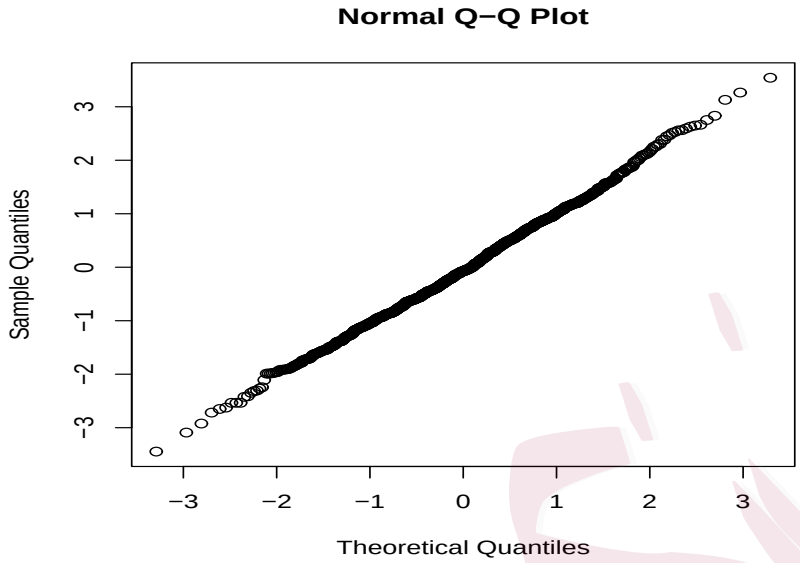

Figure 1: Simulation study. Normal quantile plot for T_4 ($n = 30$ and $n_2 = 50$).

similar size and power to T_{BZ1} with a complicated structure when the simulated values of T_{BZ1} are available. The new test T_4 has similar power to T_1 and T_{BZ1} in Case 1 and to T_3 and T_{BZ3} in Case 2. We conclude that T_4, with a simple structure, is always powerful for the two cases considered, and thus robust.

To evaluate the asymptotic normality of Theorem 4.1, we constructed the quantile plots of the test statistics against the standard normal. Figures 1 and 2 present such plots for T_4 and they reveal that the asymptotic normality is justified for the given finite sample sizes. Similar plots were obtained for test statistics T_1, T_2, and T_3 and other situations as well.

6 An Application

We applied the proposed tests to a set of panel count data arising from a skin cancer chemoprevention trial conducted by the University of Wisconsin Comprehensive Cancer Center

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

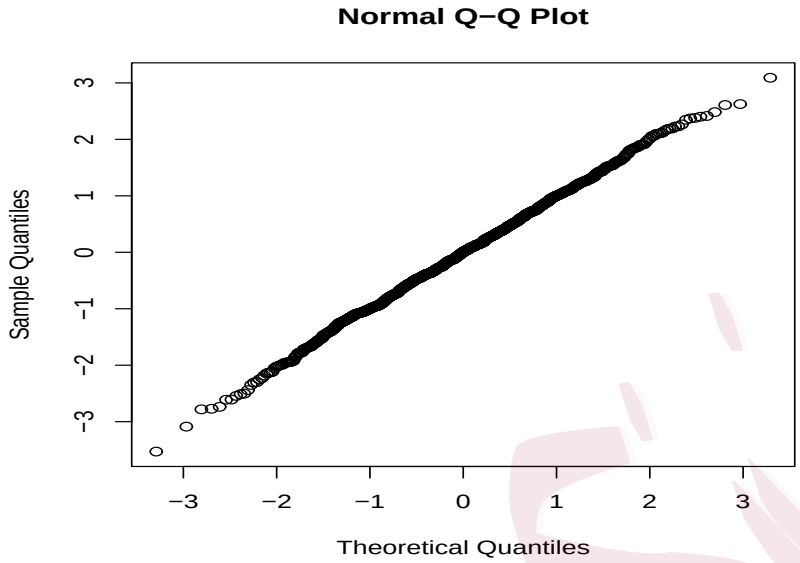

Figure 2: Simulation study. Normal quantile plot for T_4 ($n_1 = 50$ and $n_2 = 70$).

in Madison, Wisconsin. It was a double-blinded and placebo-controlled randomized phase III clinical trial. The primary objective of this trial was to evaluate the effectiveness of 0.5g/m2/day PO difluoromethylornithine (DFMO) in reducing new skin cancers in a population of the patients with a history of non-melanoma skin cancers: basal cell carcinoma and squamous cell carcinoma. The study consisted of 290 patients who were randomized to two groups: DFMO group (143) or the placebo group (147). The observed data included a sequence of observation times in days and the numbers of occurrences of both basal cell carcinoma and squamous cell carcinoma between the observation times for patients in different treatment groups (see Table 9.3 of Sun and Zhao (2013)). Sun and Zhao (2013) analyzed these data and found that the overall DFMO treatment seemed to have some mild effects in reducing the recurrence rates of basal cell carcinoma and quamous cell carcinoma. In addition, they presented a graphical comparison of the two groups and concluded that the

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

DFMO treatment seemed to have some effects in reducing the recurrence rate of basal cell carcinoma but not to have any effect on the recurrence rate of squamous cell carcinoma. For this reason, we focused on comparing two treatment groups in terms of the recurrence rates of basal cell carcinoma.

To test the difference between the two groups, we treated the DFMO group as group 1 and the placebo group as group 2. Let $N_i(t)$ represent the number of the occurrences of basal cell carcinoma up to time t for patient i, $i = 1, \ldots, 290$. Let $\Lambda_l(t)$ denote the expected occurrences of basel cell carcinoma up to time t for group l. The goal is to test $H_0 : \Lambda_1(t) = \Lambda_2(t) = \Lambda_0(t)$. The spline likelihood estimates $\hat{\Lambda}_l$ and $\hat{\Lambda}_0$ of Λ_l and Λ_0 based on samples from all the patients in the l-th group and the pooled data are shown in Figure 3. We applied the test procedure of Section 4 to this problem and obtained $T_1 = -2.2285$, $T_2 = -0.9245$, $T_3 = -2.0245$, and $T_4 = -2.1940$ where T_j's are as defined in Section 5, giving p-values of 0.0258, 0.3552, 0.0429, and 0.0282 based on the standard normal. The test results from T_1, T_3 and T_4 suggest that the incidence rates of basal cell carcinoma were significantly reduced by the DFMO treatment, while test T_2 fails to reject H_0. This can be easily understood by looking at the behavior of the estimates. From Figure 3, the difference of mean functions at later times dominate the difference at earlier times so that the test with $W_n^{(2)}$ could not detect the difference between two groups.

7 Concluding Remarks

For semiparametric models, Wellner and Zhang (2007) developed a general theorem for deriving the asymptotic normality of semiparametric M-estimators of regression parameters.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

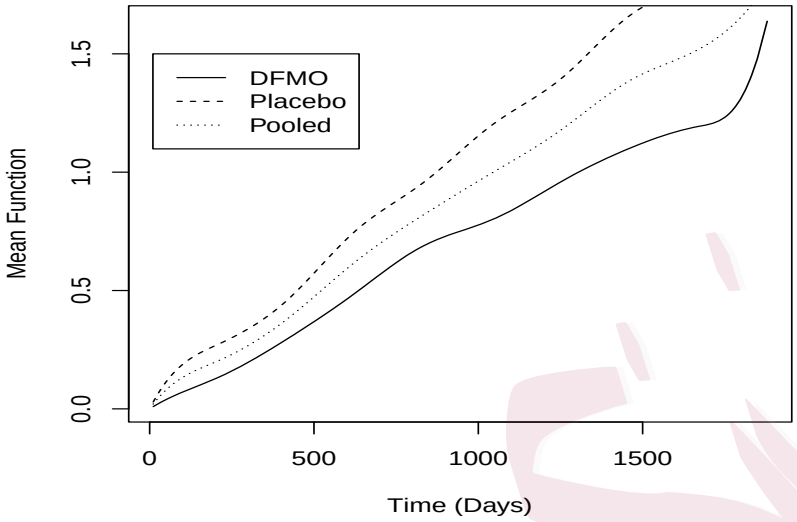

Figure 3: The estimated mean functions for the skin cancer study.

We can establish similar theory. For example, we have the following results about the asymptotic normality of estimators in the semiparametric model considered by Wellner and Zhang (2007) and Lu, Zhang and Huang (2009). Suppose that for each subject, given a d-dimensional vector of covariates Z_i, $N_i(t)$ is a non-homogeneous Poisson process with the mean function $\Lambda_i(t|Z_i) = \Lambda_0(t)\exp\{Z_i'\beta\}$, where Λ_0 is an unknown baseline mean function and β is a d-dimensional vector of unknown regression parameters.

Let $\hat{\theta}_n^{ps} = (\hat{\beta}_n^{ps}, \hat{\Lambda}_n^{ps})$ and $\hat{\theta}_n = (\hat{\beta}_n, \hat{\Lambda}_n)$ be the semiparametric pseudo-likelihood and likelihood estimators of Lu, Zhang, and Huang (2009). Let $\mathcal{B}_d$ denote the collection of Borel sets in $\mathbb{R}^d$, and $\mathcal{B}$ and $\mathcal{B}_{[0,\tau]}$ as defined in Section 3. Let F be the cumulative distribution function of Z. We considered measures ν_1 and ν_2 as follows: for $B, B_1, B_2 \in \mathcal{B}_{[0,\tau]}$, and

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

$B_3 \in \mathcal{B}_d$,

$$\nu_1(B \times B_3) = \int_{B_3} \sum_{k=1}^{\infty} P(K=k|Z=z) \sum_{j=1}^{k} P(T_{k,j} \in B|K=k, Z=z) dF(z),$$

$$\begin{aligned}
&\nu_2(B_1 \times B_2 \times B_3) \\
&= \int_{B_3} \sum_{k=1}^{\infty} \{P(K=k|Z=z) \\
&\qquad \times \sum_{j=1}^{k} P(T_{k,j-1} \in B_1, T_{k,j} \in B_2|K=k, Z=z) \Bigg\} dF(z).
\end{aligned}$$

Take $\tilde{\mathcal{H}}_r = \left\{(\mathbf{h}_1, h_2) : \mathbf{h}_1 \in \mathbb{R}^d, ||\mathbf{h}_1|| \leq 1,\ h_2 \in \mathcal{H}_r, h_2(0) = 0\right\}$. Under some regularity conditions,

(i) For $(\mathbf{h}_1, h_2) \in \tilde{\mathcal{H}}_r$, $\mathbf{h}_1' \sqrt{n}(\hat{\beta}_n^{ps} - \beta_0) + \sqrt{n} \int \{\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)\} dh_2(t)$ is asymptotically normal.

(ii) For $(\mathbf{h}_1, h_2) \in \tilde{\mathcal{H}}_r$, $\mathbf{h}_1' \sqrt{n}(\hat{\beta}_n - \beta_0) + \sqrt{n} \int \{\hat{\Lambda}_n(t) - \Lambda_0(t)\} dh_2(t)$ is asymptotically normal.

(iii) (*Asymptotic Normality of* $\hat{\beta}_n^{ps}$) $\sqrt{n}(\hat{\beta}_n^{ps} - \beta_0) \to_d N(0, \Sigma_{ps})$, where $\Sigma_{ps} = (A^{ps})^{-1} B^{ps} ((A^{ps})^{-1})'$ with

$$A^{ps} = E\left[\sum_{j=1}^{K} \Lambda_0(T_{K,j}) e^{\beta_0' Z} \left\{ Z - \frac{E\left(Z e^{\beta_0' Z} | K, T_{K,j}\right)}{E\left(e^{\beta_0' Z} | K, T_{K,j}\right)} \right\}^{\otimes 2}\right],$$

$$\begin{aligned}
B^{ps} \quad = \quad & E\left[\sum_{j=1}^{K} \sum_{j'=1}^{K} \left\{ N(T_{K,j}) - \Lambda_0(T_{K,j}) e^{\beta_0' Z} \right\}\right. \\
& \times \left\{ N(T_{K,j'}) - \Lambda_0(T_{K,j'}) e^{\beta_0' Z} \right\} \\
& \times \left\{ Z - \frac{E\left(Z e^{\beta_0' Z} | K, T_{K,j}\right)}{E\left(e^{\beta_0' Z} | K, T_{K,j}\right)} \right\} \\
& \left. \times \left\{ Z - \frac{E\left(Z e^{\beta_0' Z} | K, T_{K,j'}\right)}{E\left(e^{\beta_0' Z} | K, T_{K,j'}\right)} \right\}' \right],
\end{aligned}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

and (*Asymptotic Normality of Functional of* $\hat{\Lambda}_n^{ps}$) for $h \in \mathcal{H}_r$,

$$\begin{aligned}
&\sqrt{n}\int \{\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)\} e^{\beta_0' z} \\
&\quad \times \left\{ \frac{h(t)}{\Lambda_0(t)} - z' \left(\Gamma^{ps}\right)^{-1} E\left(Z e^{\beta_0' Z} \sum_{j=1}^{K} h(T_{K,j}) \right) \right\} d\nu_1(t,z) \\
&\to_d \; N(0, \tilde{\sigma}_1^2),
\end{aligned}$$

where

$$\Gamma^{ps} = E\left[Z Z' \sum_{j=1}^{K} \Lambda_0(T_{K,j}) e^{\beta_0' Z} \right],$$

$$\begin{aligned}
\tilde{\sigma}_1^2 \;=\; & E\left[\sum_{j=1}^{K} \left\{ N(T_{K,j}) - \Lambda_0(T_{K,j}) e^{\beta_0' Z} \right\} \right. \\
& \left. \times \left\{ \frac{h(T_{K,j})}{\Lambda_0(T_{K,j})} - Z' \left(\Gamma^{ps}\right)^{-1} E\left(Z e^{\beta_0' Z} \sum_{j'=1}^{K} h(T_{K,j'}) \right) \right\} \right]^2.
\end{aligned}$$

(iv) (*Asymptotic Normality of* $\hat{\beta}_n$) $\sqrt{n}(\hat{\beta}_n - \beta_0) \to_d N(0, \Sigma)$, where $\Sigma = (A)^{-1} B ((A)^{-1})'$

with

$$A = E\left[\sum_{j=1}^{K} \Delta\Lambda_0(T_{K,j}) e^{\beta_0' Z} \left\{ Z - \frac{E\left(Z e^{\beta_0' Z} | K, T_{K,j-1}, T_{K,j} \right)}{E\left(e^{\beta_0' Z} | K, T_{K,j-1}, T_{K,j} \right)} \right\}^{\otimes 2} \right],$$

$$\begin{aligned}
B \;=\; & E\left[\sum_{j=1}^{K} \sum_{j'=1}^{K} \left\{ \Delta N(T_{K,j}) - \Delta\Lambda_0(T_{K,j}) e^{\beta_0' Z} \right\} \right. \\
& \times \left\{ \Delta N(T_{K,j'}) - \Delta\Lambda_0(T_{K,j'}) e^{\beta_0' Z} \right\} \\
& \times \left\{ Z - \frac{E\left(Z e^{\beta_0' Z} | K, T_{K,j-1}, T_{K,j} \right)}{E\left(e^{\beta_0' Z} | K, T_{K,j-1}, T_{K,j} \right)} \right\} \\
& \left. \times \left\{ Z - \frac{E\left(Z e^{\beta_0' Z} | K, T_{K,j'-1}, T_{K,j'} \right)}{E\left(e^{\beta_0' Z} | K, T_{K,j'-1}, T_{K,j'} \right)} \right\}' \right]
\end{aligned}$$

and (*Asymptotic Normality of Functional of* $\hat{\Lambda}_n$) for $h \in \mathcal{H}_r$,

$$\sqrt{n}\int \left\{ \left(\hat{\Lambda}_n(t) - \hat{\Lambda}_n(s) \right) - \left(\Lambda_0(t) - \Lambda_0(s) \right) \right\} e^{\beta_0' z}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

$$\times \left\{ \frac{h(t)-h(s)}{\Lambda_0(t)-\Lambda_0(s)} - z'\Gamma^{-1}E\left(Ze^{\beta_0'Z}\sum_{j=1}^{K}\Delta h(T_{K,j})\right)\right\} d\nu_2(s,t,z)$$
$$\to_d \ N(0,\tilde{\sigma}_2^2),$$

where

$$\Gamma = E\left[ZZ'\sum_{j=1}^{K}\Delta\Lambda_0(T_{K,j})e^{\beta_0'Z}\right],$$

$$\begin{aligned}\tilde{\sigma}_2^2 \ &= \ E\left[\sum_{j=1}^{K}\left\{\Delta N(T_{K,j})-\Delta\Lambda_0(T_{K,j})e^{\beta_0'Z}\right\}\right. \\ &\quad \left.\times\left\{\frac{\Delta h(T_{K,j})}{\Delta\Lambda_0(T_{K,j})} - Z'\Gamma^{-1}E\left(Ze^{\beta_0'Z}\sum_{j'=1}^{K}\Delta h(T_{K,j'})\right)\right\}\right]^2.\end{aligned}$$

Here, the obtained asymptotic distributions for $\hat{\beta}_n^{ps}$ and $\hat{\beta}_n$ are the same as those in Theorem 3.3 of Wellner and Zhang (2007) and Theorem 3 of Lu, Zhang, and Huang (2009). The new results about the baseline mean function can be used to conduct statistical hypothesis tests. The proofs of the above conclusions are available from the authors.

Supplementary Materials

The supplementary materials include proofs of theorems.

Acknowledgements

The authors would like to thank the Editor, an associate editor and the two referees for helpful comments and suggestions that greatly improved the paper. Zhao's research was supported in part by the Research Grant Council of Hong Kong (503513), the Natural Science Foundation of China (11371299), and The Hong Kong Polytechnic University. Zhang's research was partially supported by the NSFC grant (471420107023).

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

References

Balakrishnan, N. and Zhao, X. (2009). New multi-sample nonparametric tests for panel count data. *Ann. Statist.* **37**, 1112–1149.

Groeneboom, P. (1996). Lectures on Inverse Problems. In *Lecture Notes in Mathematics* **1648**, 67–164. Springer-Verlag, Berlin.

Groeneboom, P. and Wellner, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Birkhäuser, New York.

Hu X. J., Lagakos, S. W., and Lockhart, R. A. (2009a). Marginal analysis of panel counts through estimating functions. *Biometrika* **96**, 445–456.

Hu, X. J., Lagakos, S. W., and Lockhart, R. A. (2009b). Generalized least squares estimation of the mean function of a counting process based on panel counts. *Statist. Sinica* **19**, 561–580.

Lu, M., Zhang, Y., and Huang, J. (2007). Estimation of the mean function with panel count data using monotone polynomial splines. *Biometrika* **94**, 705–718.

Lu, M., Zhang, Y., and Huang, J. (2009). Semiparametric estimation methods for panel count data using monotone B-splines. *J. Amer. Statist. Assoc.* **104**, 1060–1070.

Schick, A. and Yu, Q. (2000). Consistency of the GMLE with mixed case interval-censored data. *Scand. J. Statist.* **27**, 45–55.

Schumaker, L. (1981). *Spline Functions: Basic Theory.* Wiley: New York. .

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

Shen, X. and Wong, W. H. (1994). Convergence rate of sieve estimates. *Ann. Statist.* **18**, 580–615.

Sun, J. and Fang, H. B. (2003). A nonparametric test for panel count data. *Biometrika* **90**, 199–208.

Sun, J. and Kalbfleisch, J. D. (1995). Estimation of the mean function of point processes based on panel count data. *Statist. Sinica* **5**, 279–290.

Sun, J. and Zhao, X. (2013). *Statistical Analysis of Panel Count Data.* Springer, New York.

Thall, P. F. and Lachin, J. M. (1988). Analysis of recurrent events: nonparametric methods for random-interval count data. *J. Amer. Statist. Assoc.* **83**, 339–347.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag: New York.

Wellner, J. A. and Zhang, Y. (2000). Two estimators of the mean of a counting process with panel count data. *Ann. Statist.* **28**, 779–814.

Wellner, J. A. and Zhang, Y. (2007). Two likelihood-based semiparametric estimation methods for panel count data with covariates. *Ann. Statist.* **35**, 2106–2142.

Zeng, D. and Lin, D. Y. (2006). Efficient estimation of semiparametric transformation models for counting processes. *Biometrika* **93**, 627–640.

Zeng, D. and Lin, D. Y. (2007). Semiparametric transformation models with random effects for recurrent events. *J. Amer. Statist. Assoc.* **102**, 1387–1396.

Statistica Sinica: Preprint
doi:10.5705/ss.202014.0021

Zeng, D., Lin, D. Y., and Yin, G. S. (2005). Maximum likelihood estimation for the proportional odds model with random effects. *J. Amer. Statist. Assoc.* **100**, 470–482.

Zhang, Y. (2006). Nonparametric k-sample tests for panel count data. *Biometrika* **93**, 777–790.

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China and The Hong Kong Polytechnic University Shenzhen Research Institute, China.

E-mail: xingqiu.zhao@polyu.edu.hk

Department of Biostatistics, Indiana University, Indianapolis, IN 46202, USA.

E-mail: yz73@iu.edu