Statistica Sinica **33** (2023), 2763-2786
doi:https://doi.org/10.5705/ss.202021.0213

NONPARAMETRIC ESTIMATION AND TESTING FOR PANEL COUNT DATA WITH INFORMATIVE TERMINAL EVENT

Xiangbin Hu[1], Li Liu[2], Ying Zhang[3] and Xingqiu Zhao[1]

[1] *The Hong Kong Polytechnic University,* [2] *Wuhan University and* [3] *University of Nebraska Medical Center*

Abstract: Informative terminal events often occur in long-term recurrent event follow-up studies. To explicitly reflect the effects of such events on recurrent event processes, we propose a reverse nonparametric mean model for panel count data, with a terminal event subject to right censoring. This model enjoys meaningful interpretation for applications and robustness for statistical inference. Treating the distribution of the right-censored terminal event time as a nuisance functional parameter, we develop a two-stage estimation procedure by combining the Kaplan–Meier estimator and nonparametric sieve estimation techniques. We establish the consistency, convergence rate, and asymptotic normality of the proposed nonparametric estimator, and construct a class of new statistics for a two-sample test. We also establish the asymptotic properties of the new tests and evaluate their performance using extensive simulation studies. Lastly, we demonstrate the proposed method by applying it to panel count data from a Chinese longitudinal healthy longevity study.

Key words and phrases: Monotone polynomial spline, nonparametric test, panel count data, terminal event, two-stage estimation.

1. Introduction

In longitudinal follow-up studies, individuals may experience the same recurrent events repeatedly. However, the exact recurrent event times may not be observed in practice. Instead, we can examine individuals at distinct time points, and collect the number of recurrent events between two observation time points. Such data are so-called panel count data, and occur frequently in medical follow-up studies and clinical trials. Many methods have been developed for analyzing panel count data. For example, Sun and Kalbfeisch (1995), Wellner and Zhang (2000), Hu, Lagakos and Lockhart (2009), and Lu, Zhang and Huang (2007) studied the nonparametric estimation of the mean function of an under-

Corresponding author: Xingqiu Zhao, Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China. Email: xingqiu.zhao@polyu.edu.hk.

lying counting process with panel count data, and Thall and Lachin (1988), Sun and Fang (2003), Zhang (2006), Balakrishnan and Zhao (2009), and Zhao and Zhang (2017) developed nonparametric tests for comparing the mean functions of counting processes based on panel count data. In addition, Cheng and Wei (2000), Sun and Wei (2000), Hu, Sun and Wei (2003), Sun, Tong and He (2007), Wellner and Zhang (2007), Lu, Zhang and Huang (2009), He et al. (2017), and Jiang, Su and Zhao (2020) studied semiparametric regression models with panel count data, and Zhao et al. (2019) investigated a nonparametric regression model for analyzing panel count data.

Recurrent events are often truncated by a terminal event, such as death. In such cases, it is common to use a joint modeling approach and a shared frailty variable to deal with the correlation between the terminal event and the recurrent event processes; see, for example, Liu, Wolfe and Huang (2004), Zhao, Zhou and Sun (2011), Sun et al. (2012), Zhao, Li and Sun (2013a,b), and Zhou et al. (2017). However, these approaches do not reflect the explicit effect of the terminal event on recurrent events, which is often of interest in scientific investigations. For example, Chan et al. (1995) found that AIDS-defending events accumulate sharply before the death of an HIV-infected individual, and Lunney et al. (2003) claim that the functional decline changes tremendously in the last year of the lives of HIV-infected individuals. Directly modeling the effect of terminal event on recurrent events is a challenging task. Recently, Chan and Wang (2010) developed a backward model for studying the behavior of recurrent events prior to a terminal event, and Chan and Wang (2017) considered the effects of covariates and extended the backward model to a semiparametric model. Kong et al. (2018) proposed a conditional model for longitudinal data by treating the terminal event time as a covariate. Here, we propose a reverse mean model for panel count data with an informative terminal event, and then develop nonparametric inference procedures for comparing the mean functions of recurrent event processes that are likely truncated by a terminal event.

The main contributions of this work are fourfold. First, we propose a reverse nonparametric mean model for panel count data with a right-censored terminal event, where the nonparametric mean function is increasingly dependent on the terminal event time. Thus, the proposed model provides an intuitive interpretation of the effect of a terminal event on recurrent event processes. Second, we develop a two-stage sieve-based nonparametric estimation procedure by treating the distribution function of the terminal event time as a nuisance functional parameter. Third, we establish the asymptotic properties of the proposed estimator by overcoming the challenges posed by a nuisance functional parameter and a nonparametric estimator with a convergence rate slower than the standard rate $n^{-1/2}$. Fourth, we develop a class of nonparametric tests for comparing the mean functions of recurrent event processes with panel count data in the presence of an informative terminal event.

The remainder of this paper is organized as follows. In Section 2, we present a reverse mean model with a terminal event, and propose a two-stage nonparametric sieve-based estimation procedure. In Section 3, we establish the asymptotic properties of the proposed estimator. In Section 4, we propose a new class of test statistics for two-sample comparisons, and establish their asymptotic normality. In Section 5, we conduct simulation studies to demonstrate the finite-sample performance of the proposed method, and in Section 6, we apply the proposed method to panel count data from a Chinese longitudinal healthy longevity study. We conclude the paper in Section 7. All technical proofs are provided in the Supplementary Material.

2. Model Setting and Estimation Procedure

Suppose that a counting process $\{N(t) : 0 \leq t \leq \tau\}$ denotes the number of recurrent events occurring up to time t, where τ is a fixed time point. Let $\underline{T} = (T_1, T_2, \ldots, T_K)$ be the observation times of $N(t)$, where K represents the total number of observation times. Then, the observed counting process is $\underline{N} = (N_1, N_2, \ldots, N_K) = (N(T_1), N(T_2), \ldots, N(T_K))$. Let U and C be the terminal event time and the censoring time, respectively. The time for the last observation is denoted by $Y = U \wedge C$, and whether the time is a terminal event time is denoted by $\Delta = 1_{\{U \leq C\}}$. The observed data for subject i consist of $X_i = (Y_i, \Delta_i, K_i, \underline{T_i}, \underline{N_i})$, for $i = 1, \ldots, n$, where $\underline{T_i} = (T_{i1}, T_{i2}, \ldots, T_{iK_i})$ and $\underline{N_i} = (N(T_{i1}), N(T_{i2}), \ldots, N(T_{iK_i}))$. Let $\mathcal{X} = \{X_i, i = 1, \ldots, n\}$ denote a sample of such panel count data from n subjects.

To investigate the effect of a terminal event on the recurrent event process, motivated by Chan and Wang (2010, 2017) and Kong et al. (2018), we consider a counting process $\tilde{N}(t; U)$ denoting the event counts from time t to the terminal event U, and propose a reverse nonparametric mean model anchored at the terminal event:

$$E(\tilde{N}(t; U)|U = u) = \Lambda(u - t), \tag{2.1}$$

where $\Lambda(\cdot)$ is an unknown nondecreasing function, with $\Lambda(0) = 0$ to ensure the identifiability of the model. We propose modeling the recurrent events anchored at the terminal event time using (2.1). The advantage of this approach is that we can describe the behavior of the stochastic process from the terminal event in a straightforward manner, and can still use all the event counts to learn the model of the entire process. This model implies that $E(\tilde{N}(t_1;U)|U=u)-E(\tilde{N}(t_2;U)|U=u)=\Lambda(u-t_1)-\Lambda(u-t_2)$, where $0\le t_1\le t_2\le u$. Noting that $N(t_2)-N(t_1)=\tilde{N}(t_1;U)-\tilde{N}(t_2;U)$ and $N(0)=0$, we obtain $E(N(t)|U=u)=\Lambda(u)-\Lambda(u-t)$.

Let $F(u)$ denote the underlying distribution function of U. To make a valid inference for our proposed model, we consider the following data scenarios: (i) U and C are independent; (ii) the censoring event time C is noninformative for Λ; and (iii) given (Y,Δ), the distribution of $(K,\underline{T})$ is noninformative for Λ. Define $\triangle N_j=N(T_j)-N(T_{j-1})$ and $\triangle\Lambda_j(u)=\Lambda(u-T_{j-1})-\Lambda(u-T_j)$, for $j=1,\ldots,K$ with $T_0=0$. A straightforward calculation yields

$$\begin{aligned}&E\left[\sum_{j=1}^{K}\{\triangle N_j-\triangle\Lambda_j(U)\}^2|Y,\Delta,K,\underline{T},\underline{N}\right]\\&=\sum_{j=1}^{K}\Delta\{\triangle N_j-\triangle\Lambda_j(Y)\}^2+\sum_{j=1}^{K}(1-\Delta)\frac{\int_Y^\infty\{\triangle N_j-\triangle\Lambda_j(u)\}^2dF(u)}{1-F(Y)}.\end{aligned}$$

Hence, we propose the following least-squares-based loss function:

$$\begin{aligned}\ell_n(\Lambda,F;\mathcal{X})=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K_i}&\left[\Delta_i\{\triangle N_{i,j}-\triangle\Lambda_{i,j}(Y_i)\}^2\right.\\&\left.+(1-\Delta_i)\frac{\int_{Y_i}^\infty\{\triangle N_{i,j}-\triangle\Lambda_{i,j}(u)\}^2dF(u)}{1-F(Y_i)}\right],\end{aligned}\tag{2.2}$$

where $\triangle N_{i,j}=N_i(T_{ij})-N_i(T_{i(j-1)})$ and $\triangle\Lambda_{i,j}(u)=\Lambda(u-T_{i(j-1)})-\Lambda(u-T_{ij})$. A natural idea is to take the minimizer of $l_n(\Lambda,F;\mathcal{X})$, defined in (2.2) as the estimator of the parameter. However, because the loss function involves an unknown distribution function F, it is difficult to estimate Λ and F simultaneously. To tackle this problem, we propose a two-stage approach. In Stage 1, we estimate F using the Kaplan–Meier (KM) estimator $\hat{F}_n(u)$ (Kaplan and Meier (1958)). In Stage 2, we obtain $\hat{\Lambda}_n$ by minimizing the loss function $\ell_n(\Lambda,\hat{F}_n;\mathcal{X})$ with respect to Λ. Because the observed data of (Y,Δ) are used in both stages, to distinguish them, we use the notation $(\tilde{Y},\tilde{\Delta})$ to represent the data when we obtain the KM estimator in Stage 1, without any ambiguity.

We adopt the spline-based sieve estimation method to estimate the nondecreasing function Λ because of its numerical advantage and good statistical properties (Lu, Zhang and Huang (2007, 2009)). Let $\{t_i: i=1,\ldots,m_n+2d\}$ be a sequence of knots that partition $[0,\tau]$ into m_n+1 subintervals, where

$$0 = t_1 = \cdots = t_d < t_{d+1} < \cdots < t_{m_n+d} < t_{m_n+d+1} = \cdots = t_{m_n+2d} = \tau.$$

Let $q_n = m_n + d$ and $\{I_l(s), l = 1, \ldots, q_n\}$ be I-spline basis functions of order d (Ramsay (1988)). We then define the functional space of the estimator for Λ as

$$\Phi_n = \left\{ \sum_{l=1}^{q_n} \alpha_l I_l(s) : \alpha_l \geq 0, l = 1, \ldots, q_n \right\}.$$

Because $\hat{F}_n$ is a monotone step function, as shown in Section S1 of the online Supplementary Material, minimizing the loss function $l_n(\Lambda, \hat{F}_n; \mathcal{X})$ is a quadratic programming problem with the constraint that $\alpha_l \geq 0$, for $l = 1, \ldots, q_n$. Let $\mathbf{I}(s) = (I_1(s), \ldots, I_{q_n}(s))^T$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{q_n})^T$, and the solution of the quadratic programming problem be $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_{q_n})^T$. Then, the spline estimator of $\Lambda(s)$ is $\hat{\Lambda}_n(s) = \mathbf{I}(s)^T \hat{\boldsymbol{\alpha}}$.

3. Asymptotic Properties

Before presenting the asymptotic results, we introduce some notation. Let $g^{(r)}$ be the rth derivative of the function g. For $r \geq 1$, define

$$\begin{aligned} \mathcal{H}_r &= \{g : |g^{(r-1)}(s) - g^{(r-1)}(t)| \leq c_0|s-t| \ \textit{for all}\ 0 \leq s, t \leq \tau\}, \\ \Phi &= \{\Lambda \in \mathcal{H}_r : \Lambda \text{ is a nondecreasing continuous function} \\ &\qquad \text{on } [0,\tau] \text{ with } \Lambda(0) = 0\}, \\ \mathcal{F} &= \{F : F \text{ is a distribution function on } [0,\infty)\}. \end{aligned}$$

Denote the true value of (Λ, F) as $(\Lambda_0, F_0) \in \Phi \times \mathcal{F}$. For

$$B_1, B_2 \in \mathcal{B}_{[0,\tau]} =: \{B \cap [0,\tau] : B \in \mathcal{B}\},$$

where $\mathcal{B}$ denotes the collection of Borel sets, set

$$\begin{aligned} \mu_1(B_1 \times B_2) &= \int \sum_{k=1}^{\infty} P(K = k | U = u) \\ &\quad \times \sum_{j=1}^{k} P\big((u - T_j) \in B_1, (u - T_{j-1}) \in B_2 | K = k, U = u\big) dF_0(u), \\ \mu_2(B_1 \times B_2) &= \int \sum_{k=1}^{\infty} P(K = k | U = u) \\ &\quad \times P\big((u - T_K) \in B_1, u \in B_2 | K = k, U = u\big) dF_0(u). \end{aligned} \tag{3.1}$$

Then, μ_1 and μ_2 are measures on $([0,\tau]^2, \mathcal{B}^2_{[0,\tau]})$. For any functions $\Lambda_1, \Lambda_2 \in \Phi$, we define the metric as

$$\begin{aligned} d_1(\Lambda_1,\Lambda_2)^2 &= ||\triangle\Lambda_1(s_1,s_2) - \triangle\Lambda_2(s_1,s_2)||^2_{L_2(\mu_1)} \\ &= E\left[\sum_{j=1}^{K}\Big(\triangle\Lambda_{1,j}(U) - \triangle\Lambda_{2,j}(U)\Big)^2\right] \\ &= E\left[\sum_{j=1}^{K}\left\{\Delta\Big(\triangle\Lambda_{1,j}(Y) - \triangle\Lambda_{2,j}(Y)\Big)^2\right.\right. \\ &\qquad \left.\left. + (1-\Delta)\frac{\int_Y^\infty (\triangle\Lambda_{1,j}(u) - \triangle\Lambda_{2,j}(u))^2 dF_0(u)}{1-F_0(Y)}\right\}\right], \end{aligned}$$

where $\triangle\Lambda(s_1,s_2) = \Lambda(s_2) - \Lambda(s_1)$. For any $F_1, F_2 \in \mathcal{F}$, we define the metric as

$$d_2(F_1,F_2) = ||F_1 - F_2||_\infty,$$

where $||\cdot||_\infty$ represents the L_∞ norm. Denote the δ-neighborhood of F_0 by $\mathcal{F}_\delta = \{F \in \mathcal{F} : d_2(F,F_0) \le \delta\}$, for any small $\delta > 0$.

To establish the asymptotic properties of the proposed estimator, we assume the following:

(C1) $0 < \Lambda_0(\tau) < \infty$.

(C2) $0 < F_0(\tau) < 1$. F_0 is absolutely continuous with respect to the Lebesgue measure. Moreover, the density function $f_0(s)$ has a uniform positive lower bound, for all $s \in [M_1, \tau]$, where M_1 is a constant representing the minimum value of the support of F_0.

(C3) $E\big[\sum_{j=1}^K\{\triangle N_j - \triangle\Lambda_{0,j}(U)\}^2\big] < \infty$.

(C4) The probability of censoring $\varrho = P(Y < U)$ satisfies that $0 < \varrho < 1$.

(C5) $m_n = O(n^\nu)$, for $0 < \nu < 1/2$. Moreover, we suppose that

$$\max_{d+1\le i\le m_n+d+1} |t_i - t_{i-1}| = O(n^{-\nu}),$$

and there is a constant $M_2 > 0$ such that

$$\frac{\max_{d+1\le i\le m_n+d+1}|t_i - t_{i-1}|}{\min_{d+1\le i\le m_n+d+1}|t_i - t_{i-1}|} \le M_2$$

uniformly for n.

(C6) There is a constant $M_3 > 0$ such that $P(K \leq M_3) = 1$.

(C7) $P(T_j - T_{j-1} \geq M_4,\ \text{for all } j = 1, \ldots, K) = 1$, with some constant $M_4 > 0$.

Remark 1. Condition (C1) is standard in the literature on nonparametric estimation. Condition (C2) holds, in general, for the cumulative distribution function of a continuous random variable. Condition (C3) requires that $\sum_{j=1}^{K} \triangle N_j$ has finite second-order central moment. Condition (C4) ensures that the censoring rate lies between zero and one, which is commonly assumed in survival data analyses. Condition (C5) is a regularity condition for the monotone spline estimation; see Lu, Zhang and Huang (2007, 2009). Condition (C6) is similar to Condition (C2) in Wellner and Zhang (2007), and indicates that the number of observations is bounded. Condition (C7) requires that the adjacent observation times are separable (Wellner and Zhang (2007)), which is generally the case in practice.

Theorem 1. (Consistency for the two-stage estimator). *Suppose that Conditions* (C1)–(C7) *hold. Then, for every* $0 \leq b_1 \leq b_2 \leq \tau$ *satisfying* $\mu_2([0, b_1] \times [b_2, \tau]) > 0$*, we have*

$$\|\triangle\hat{\Lambda}_n(s_1, s_2)1_{\{(s_1,s_2)\in[b_1,b_2]\times[b_1,b_2]\}} - \triangle\Lambda_0(s_1, s_2)1_{\{(s_1,s_2)\in[b_1,b_2]\times[b_1,b_2]\}}\|^2_{L_2(\mu_1)} = o_p(1).$$

In particular, if $\mu_2(\{0\} \times \{\tau\}) > 0$*, then* $d_1(\hat{\Lambda}_n, \Lambda_0) = o_p(1)$.

To establish the rate of convergence and the asymptotic normality, we further assume the following:

(C8) μ_1 is absolutely continuous with respect to the Lebesgue measure with a derivative $\dot{\mu}_1$, and $\dot{\mu}_1$ has a uniform positive lower bound.

(C9) There is a constant $0 < \lambda_a < \infty$ such that $1/\lambda_a < \Lambda_0'(s) < \lambda_a$, for all $s \in [\tau', \tau]$, with $0 < \tau' \leq \tau$ such that $\Lambda_0(\tau') > 0$.

(C10) $P(U \geq \tau) = \omega_1 > 0$ and $P(C \geq \tau) = \omega_2 > 0$.

(C11) $E(e^{cN(t)})$ is uniformly bounded for $t \in [0, \tau]$ and some constant c.

Remark 2. Condition (C8) implies that the metric μ_1 defined in (3.1) has a strictly positive intensity. Condition (C9) requires that the true conditional mean function be absolutely continuous with a bounded intensity function, which is reasonable, as explained in Wellner and Zhang (2007). Condition (C10) is used as a technical condition in the proof of the uniform weak convergence rate of the KM estimator, following Kong et al. (2018). Condition (C11) holds when $N(t)$ is from a Poisson-type process or is uniformly bounded, conditional on the terminal event time, which is also true in practice, in general.

Theorem 2. (Rate of Convergence). *Suppose that Conditions* (C1)–(C11) *hold and* $\mu_2(\{0\} \times \{\tau\}) > 0$. *Taking* $\nu = 1/(1+2r)$, *we have* $d_1(\hat{\Lambda}_n, \Lambda_0) = O_p(n^{-r/(1+2r)})$.

Before presenting the asymptotic normality, we introduce some additional notation. Let $\mathcal{P}$ and $\mathbb{P}_n$ denote the probability measure and empirical measure, respectively, with $\mathcal{P}f = \int f d\mathcal{P}$ and $\mathbb{P}_n f = n^{-1}\sum_{i=1}^n f(X_i)$. Write $\ell_n(\Lambda, F; \mathcal{X}) = \mathbb{P}_n m(\Lambda, F; X)$. Let $l^\infty(\mathcal{H}_r)$ be the space of bounded functionals on $\mathcal{H}_r$ under the L_∞ norm. For $h \in \mathcal{H}_r$, define a sequence of maps Q_n of a neighborhood of (Λ_0, F_0), denoted by $\mathcal{U}$, in the parameter space for (Λ, F) into $l^\infty(\mathcal{H}_r)$ as the derivative of $\ell_n(\Lambda, F; \mathcal{X})$ with respect to Λ in the direction h:

$$\begin{aligned} Q_n(\Lambda, F)[h] &= \lim_{\varepsilon \to 0} \frac{\ell_n(\Lambda + \varepsilon h, F; \mathcal{X}) - l_n(\Lambda, F; \mathcal{X})}{\varepsilon} \\ &= \mathbb{P}_n \lim_{\varepsilon \to 0} \frac{m(\Lambda + \varepsilon h, F; X) - m(\Lambda, F; X)}{\varepsilon} \\ &= \mathbb{P}_n \psi(\Lambda, F; X)[h], \end{aligned}$$

and define $Q(\Lambda, F)[h] = \mathcal{P}\psi(\Lambda, F; X)[h]$, where

$$\begin{aligned} \psi(\Lambda, F; X)[h] = \sum_{j=1}^K \Bigg[& \Delta\big\{\triangle N_j - \triangle\Lambda_j(Y)\big\}\triangle h_j(Y) \\ & + (1-\Delta)\frac{\int_Y^\infty \{\triangle N_j - \triangle\Lambda_j(u)\}\triangle h_j(u) dF(u)}{1 - F(Y)}\Bigg], \end{aligned}$$

with $\triangle h_j(u) = h(u - T_{j-1}) - h(u - T_j)$, for $j = 1, \ldots, K$. Furthermore, define

$$\begin{aligned} \dot{Q}^{(1)}_{\Lambda_0, \hat{F}_n}(\hat{\Lambda}_n - \Lambda_0)[h] &= \lim_{\varepsilon \to 0} \frac{Q(\Lambda_0 + \varepsilon(\hat{\Lambda}_n - \Lambda_0), \hat{F}_n)[h] - Q(\Lambda_0, \hat{F}_n)[h]}{\varepsilon} \\ &= -\mathcal{P}\varsigma(\hat{\Lambda}_n, \hat{F}_n; X)[h], \end{aligned}$$

with

$$\begin{aligned} \varsigma(\Lambda, F; X)[h] = \sum_{j=1}^K \Bigg[& \Delta\{\triangle\Lambda_j(Y) - \triangle\Lambda_{0,j}(Y)\}\triangle h_j(Y) \\ & + (1-\Delta)\frac{\int_Y^\infty \{\triangle\Lambda_j(u) - \triangle\Lambda_{0,j}(u)\}\triangle h_j(u) dF(u)}{1 - F(Y)}\Bigg], \end{aligned}$$

and

$$\dot{Q}^{(2)}_{\Lambda_0,F_0}(\hat{F}_n - F_0)[h] = \lim_{\varepsilon\to 0} \frac{Q(\Lambda_0, F_0 + \varepsilon(\hat{F}_n - F_0))[h] - Q(\Lambda_0, F_0)[h]}{\varepsilon}$$
$$= \mathcal{P}\left[\int_Y^\infty \bar{\varphi}_{\Lambda_0,F_0}(u;X)[h]d(\hat{F}_n - F_0)(u)\right],$$

with

$$\bar{\varphi}_{\Lambda,F}(u;X)[h] = \frac{1-\Delta}{1-F(Y)}\sum_{j=1}^{K}\Bigg\{\big(\triangle N_j - \triangle\Lambda_j(u)\big)\cdot\triangle h_j(u) - \frac{\int_Y^\infty \big(\triangle N_j - \triangle\Lambda_j(s)\big)\cdot\triangle h_j(s)dF(s)}{1-F(Y)}\Bigg\}.$$

Theorem 3. (Asymptotic Normality). *Suppose that Conditions* (C1)–(C11) *hold and* $\mu_2(\{0\}\times\{\tau\}) > 0$.

(i) Then, for any bounded function $h \in \mathcal{H}_r$, *we have*

$$-\sqrt{n}\dot{Q}^{(1)}_{\Lambda_0,\hat{F}_n}(\hat{\Lambda}_n - \Lambda_0)[h] = \sqrt{n}\dot{Q}^{(2)}_{\Lambda_0,F_0}(\hat{F}_n - F_0)[h] + \sqrt{n}Q_n(\Lambda_0, F_0)[h] + o_p(1).$$

(ii) Moreover, $\sqrt{n}\mathcal{P}\varsigma(\hat{\Lambda}_n, \hat{F}_n; X)[h] \xrightarrow{d} N(0, \sigma_0^2)$, *where*

$$\sigma_0^2 = E\left[\{\mathcal{P}\varphi(\Lambda_0, F_0; X; \tilde{Y}, \tilde{\Delta})[h] + \psi(\Lambda_0, F_0; X)[h]\}^2\right],$$

and $\varphi(\Lambda, F; X; \tilde{Y}, \tilde{\Delta})[h]$ *is defined in the online Supplementary Material.*

4. Two-Sample Test

Suppose that n subjects are drawn from two groups with corresponding sample sizes n_1 and n_2, where $n_1 + n_2 = n$. Denote the observed data of the lth group as $\{X_i^{(l)} : i = 1,\ldots,n_l\} = \{(Y_i^{(l)}, \Delta_i^{(l)}, K_i^{(l)}, \underline{T_i}^{(l)}, \underline{N_i}^{(l)}) : i = 1,\ldots,n_l\}$, for $l = 1, 2$. Given the terminal event time $U^{(l)} = u$, the conditional mean function of $\tilde{N}^{(l)}(t)$ is $\Lambda_l(u - t)$.

4.1. Terminal events with an identical distribution

In this subsection, we assume that the terminal event times share the same distribution function F_0 for all the subjects, so that we can obtain its estimator $\hat{F}_n$ from the pooled data. We investigate a two-sample test with the null hypothesis $H_0 : \Lambda_1 = \Lambda_2 = \Lambda_0$. Denote $\hat{\Lambda}_l$ and $\hat{\Lambda}_n$ as the estimates of Λ_l and Λ_0, respectively, based on the data set of group l and the pooled data, respectively.

Theorem 4. (Two-sample test with an identical distribution of terminal events). *In addition to the conditions in Theorem* 3*, we suppose that* $h_n(\cdot)$ *is a bounded weight process, and there is a bounded function* $h \in \mathcal{H}_r$ *such that*

$$d_1^2(h_n, h) = E\left[\sum_{j=1}^{K}\{\triangle h_{n,j}(U) - \triangle h_j(U)\}^2\right] = o_p(n^{-1/(1+2r)}).$$

Assume that $n_1/n \to p$ *as* $n \to \infty$*, where* $0 < p < 1$*. Then, under* $H_0 : \Lambda_1 = \Lambda_2$*,* $U_n = \sqrt{n}\mathbb{P}_n(\varsigma(\hat{\Lambda}_1, \hat{F}_n; X)[h_n] - \varsigma(\hat{\Lambda}_2, \hat{F}_n; X)[h_n])$ *converges in distribution to* $N(0, (1/p + 1/(1-p))\sigma_0^2)$*, where* $\sigma_0^2 = E[\psi^2(\Lambda_0, F_0; X)[h]]$*. Moreover,* σ_0^2 *can be estimated consistently using* $\hat{\sigma}_n^2 = \mathbb{P}_n[\psi^2(\hat{\Lambda}_n, \hat{F}_n; X)[h_n]]$*.*

Remark 5. Theorem 4 states the asymptotic normality of the new statistic U_n and gives a consistent estimator of its asymptotic variance. Then, the h_n-specific standardized statistic, $T_n(h_n) = U_n\{\hat{\sigma}_n\sqrt{n/n_1 + n/n_2}\}^{-1}$, can be applied to conduct the two-sample hypothesis test. A natural choice is to take $h_n = \hat{\Lambda}_n$. Other possible choices of weight processes are explored and evaluated in the simulation studies. Note that the weight process is only required to be bounded, which is more flexible than the monotone conditions in Zhang (2006) and Balakrishnan and Zhao (2009).

4.2. Terminal events with different distributions

In this subsection, we assume the distribution and density functions of $U^{(l)}$, for $l = 1, 2$, are F_l and f_l, respectively, which may be different for the two groups. Let $\hat{F}_l$ and $\hat{\Lambda}_l$ be estimators of F_l and Λ_l, respectively. Given a partition $0 = t_0^{(l)} < t_1^{(l)} < \cdots < t_{\nu_{n_l}}^{(l)} = \tau$, we define histogram-type estimators of f_l as $\hat{f}_l(u) = (\hat{F}_l(t_{i_l}^{(l)}) - \hat{F}_l(t_{i_l-1}^{(l)}))/(t_i^{(l)} - t_{i-1}^{(l)})$ for $t_{i_l-1}^{(l)} \le u < t_{i_l}^{(l)}$, following Földes, Rejtő and Winter (1981). Set $f_T(u) = p_1 f_1(u) + p_2 f_2(u)$,

$$w_l(u - t_1) - w_l(u - t_2) = (h(u - t_1) - h(u - t_2))\frac{f_T(u)}{f_l(u)},$$
$$w_n^{(l)}(u - t_1) - w_n^{(l)}(u - t_2) = (h_n(u - t_1) - h_n(u - t_2))\left(\frac{n_l}{n} + \frac{n_r}{n}\frac{\hat{f}_r(u)}{\hat{f}_l(u)}\right),$$

where $p_l = \lim_{n\to\infty} n_l/n$, for $l, r = 1, 2$, $l \ne r$.

Theorem 5. (Two-sample test with different distributions of terminal events). *Suppose that the conditions in Theorem* 4 *hold for each group and that* f_l *are Lipschitz continuous. Then, for* $\Lambda_0 \in \mathcal{H}_r, r \ge 2$*, under the null hypothesis* $H_0 : \Lambda_1 = \Lambda_2 = \Lambda_0$*, we have the following:*

(i) $\tilde{U}_n = (1/\sqrt{n})\sum_{l=1}^{2}\sum_{i=1}^{n_l}(\varsigma(\hat{\Lambda}_1, \hat{F}_l; X_i^{(l)})[h_n] - \varsigma(\hat{\Lambda}_2, \hat{F}_l; X_i^{(l)})[h_n])$ *converges in distribution to* $N(0, (\sigma_1^2/p_1 + \sigma_2^2/p_2))$*, where*

$$\sigma_l^2 = E[\{\mathcal{P}\varphi_l(\Lambda_0, F_l; X^{(l)}; \tilde{Y}^{(l)}, \tilde{\Delta}^{(l)})[w_l] + \psi(\Lambda_0, F_l; X^{(l)})[w_l]\}^2],$$

with $\varphi_l(\Lambda, F; X; \tilde{Y}, \tilde{\Delta})[w]$ *defined in the online Supplementary Material.*

(ii) In addition, suppose the knots of the partition satisfy

$$\max_{i=1,\ldots,\nu_{n_l}{}^{(l)}}\{|t_i^{(l)} - t_{i-1}^{(l)}|\} \to 0 \text{ and } \left(\frac{n}{\log n}\right)^{1/4} \min_{i=1,\ldots,\nu_{n_l}{}^{(l)}}\left\{\left|t_i^{(l)} - t_{i-1}^{(l)}\right|\right\} \to \infty$$

as $n \to \infty$*. Then,* σ_l^2 *can be estimated consistently by* $\hat{\sigma}_l^2$*, and the asymptotic variance of* $\tilde{U}_n$ *can be estimated consistently by* $\tilde{\sigma}_n^2 = n(\hat{\sigma}_1^2/n_1 + \hat{\sigma}_2^2/n_2)$*, where*

$$\hat{\sigma}_l^2 = \mathbb{P}_{n_l}\left[\left\{\mathbb{P}_{n_l}\varphi_n^{(l)}(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \tilde{Y}^{(l)}, \tilde{\Delta}^{(l)})[w_n^{(l)}] + \psi(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[w_n^{(l)}]\right\}^2\right],$$

with $\varphi_n^{(l)}(\Lambda, F; X; \tilde{Y}, \tilde{\Delta})[w]$ *defined in the online Supplementary Material.*

According to Theorem 5, we can apply the statistics $\tilde{T}_n(h_n) = \tilde{U}_n/\tilde{\sigma}_n$ to test the equality of the mean function of the counting processes under the scenario of different terminal event time distributions between the two samples.

5. Simulation Studies

5.1. Two-stage estimation for the mean function

We use simulation studies to evaluate the finite-sample performance of the two-stage estimator $\hat{\Lambda}_n$ with sample sizes $n = 100$ and 200. For subject i, the observation $X_i = (Y_i, \Delta_i, K_i, \underline{T_i}, \underline{N_i})$ is generated as follows. The latent terminal event time U_i is from $6 + \exp(1)$, and the censoring time C_i is from $6 + \kappa\exp(3)$. The study ending time is set at $\tau = 10$. The values of κ are chosen as 1.375 and 0.500, such that the terminal events are censored by 20% and 40%, respectively. Let $Y_i = U_i \wedge C_i$ and $\Delta_i = 1_{\{U_i \le C_i\}}$. The number of observations K_i is generated from the discrete uniform distribution in $\{1, \ldots, 6\}$. Given the censored terminal event time Y_i, observation time $\underline{T_i} = (T_{i1}, T_{i2}, \ldots, T_{iK_i})$ is an ordered sample from the uniform distribution $\text{Unif}(0, Y_i)$. Let N_i be a Poisson process with $E(N_i(t)|U = u) = \Lambda_0(u) - \Lambda_0(u-t)$ and $\Lambda_0(s) = s$. Therefore, $N_i(T_{i1})$ is from the Poisson distribution with mean T_{i1}, and $N_i(T_{ij}) - N_i(T_{i(j-1)})$ is from the Poisson distribution with mean $T_{ij} - T_{i(j-1)}$. For the knots of the spline, $d = m_n = 3$,

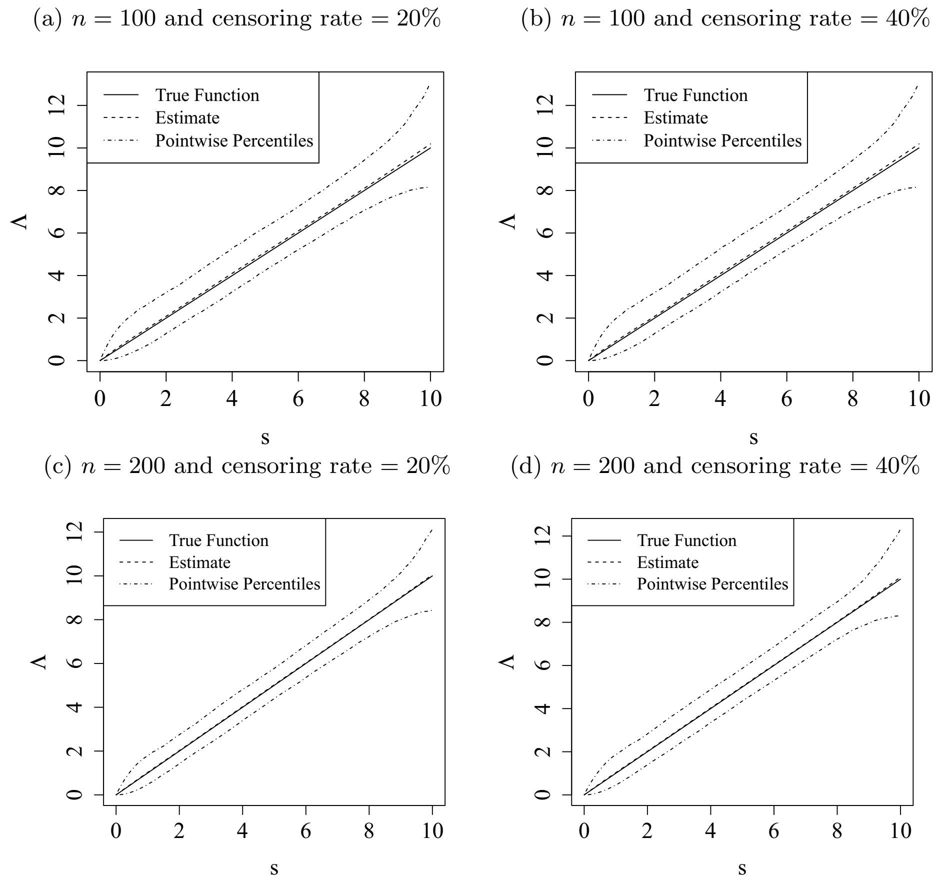

Figure 1. Estimates of the mean function.

and let $t_{d+1}, t_{d+2}, t_{d+3}$ be the quartiles of $\{Y_i - T_{ij} : j = 1, \ldots, K_i, i = 1, \ldots, n\}$. All of our simulation studies are based on 1,000 replications.

Figure 1 shows the results of the two-stage estimators for the conditional mean functions. The solid line represents the true mean function, the dashed line represents the average of the estimated mean function, and the dotted-dash lines represent the 2.5% and 97.5% pointwise percentiles of the functional estimator based on 1,000 replications. From (a)–(d) in Figure 1, we conclude that the sample mean of the estimated mean functions is very close to the true mean, the degree of variability of the functional estimator decreases as the sample size increases, indicating the asymptotic unbiasedness of the proposed method, and the sampling distribution of the estimator does not appear to be affected by the censoring rate.

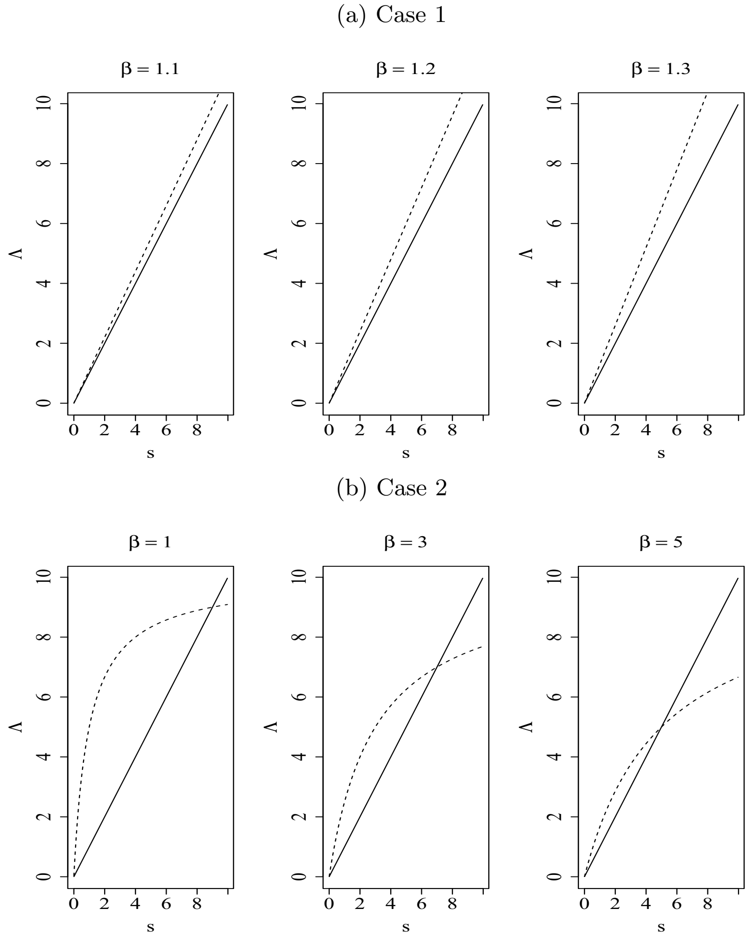

Figure 2. Discrepancy of the mean functions.

5.2. Two-sample test with an identical distribution of terminal event times

Here, we conduct the two-sample test for two groups that have the same distribution of terminal event times. We generate two groups of independent and identically distributed (i.i.d.) samples $\{X_i^{(l)} : i = 1, \ldots, n_l\}$ in the same way as in Subsection 5.1, for $l = 1, 2$, with sample size $n_1 = n_2 = 100$, 150, or 200. Let $N^{(l)}$ be a Poisson process with $E(N^{(l)}(t)|U^{(l)} = u) = \Lambda_l(u) - \Lambda_l(u - t)$ for the lth group. We considered the following two cases:

$$\text{Case 1} : \Lambda_1(s) = s,\ \ \Lambda_2(s) = \beta s;$$
$$\text{Case 2} : \Lambda_1(s) = s,\ \ \Lambda_2(s) = \frac{10s}{s+\beta}.$$

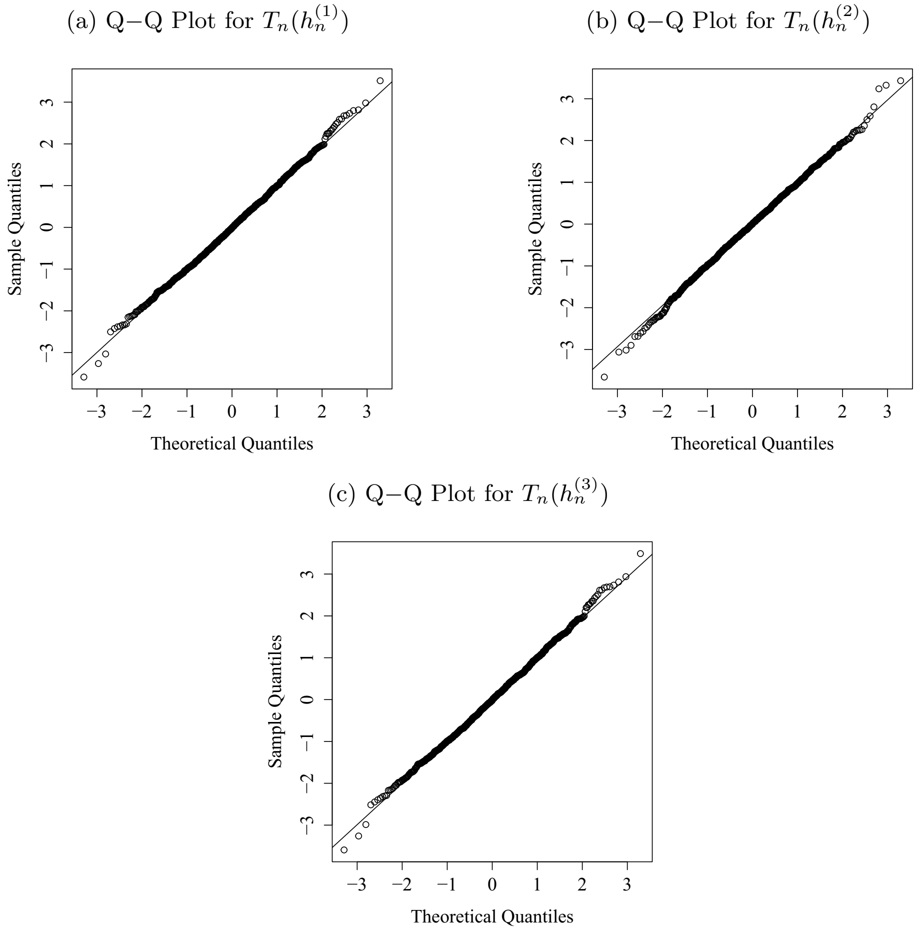

Figure 3. Q–Q plots for n_1=n_2=200, β=1, and censoring rate = 20% when the distribution functions of the terminal event time are identical for the two groups.

We take $\beta = 1, 1.1, 1.2, 1.3$ in Case 1 and $\beta = 1, 3, 5$ in Case 2.

Figure 2 plots the true mean functions for the two cases with different values of β. The conditional mean functions of the two groups do not overlap in Case 1, but do cross over in Case 2. The weight processes $h_n^{(j)}(t)$, for $j = 1, 2, 3$, in Theorem 4 are chosen to be

$$h_n^{(1)}(t) = t,\ h_n^{(2)}(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{(t \geq T_{iK_i})},\ h_n^{(3)}(t) = \hat{\Lambda}_n(t).$$

The results of the simulation study are evaluated based on 1,000 replications.

Figure 3 presents quantile plots of the test statistics with the three weight processes against the standard normal distribution for Case 1 with $n_1 = n_2 = 200$,

Table 1. Simulation results of the two-sample tests with different weights for Case 1 when the distribution functions of the terminal event time are identical for the two groups.

	Censoring rate 20%			Censoring rate 40%		
β	$T_n(h_n^{(1)})$	$T_n(h_n^{(2)})$	$T_n(h_n^{(3)})$	$T_n(h_n^{(1)})$	$T_n(h_n^{(2)})$	$T_n(h_n^{(3)})$
			n_1=n_2=100			
1	0.051	0.047	0.052	0.052	0.048	0.051
1.1	0.242	0.229	0.243	0.258	0.240	0.255
1.2	0.728	0.703	0.728	0.715	0.690	0.717
1.3	0.967	0.956	0.967	0.962	0.956	0.960
			n_1=n_2=150			
1	0.048	0.047	0.048	0.049	0.048	0.048
1.1	0.359	0.344	0.361	0.355	0.338	0.354
1.2	0.873	0.851	0.873	0.877	0.859	0.878
1.3	0.996	0.989	0.996	0.995	0.991	0.995
			n_1=n_2=200			
1	0.042	0.044	0.043	0.043	0.051	0.043
1.1	0.438	0.398	0.443	0.467	0.433	0.467
1.2	0.944	0.924	0.943	0.954	0.939	0.954
1.3	1.000	1.000	1.000	0.999	0.998	0.999

Table 2. Simulation results of the two-sample tests with different weights for Case 2 when the distribution functions of the terminal event time are identical for the two groups.

	Censoring rate 20%			Censoring rate 40%		
β	$T_n(h_n^{(1)})$	$T_n(h_n^{(2)})$	$T_n(h_n^{(3)})$	$T_n(h_n^{(1)})$	$T_n(h_n^{(2)})$	$T_n(h_n^{(3)})$
			n_1=n_2=100			
1	0.999	1.000	0.752	1.000	1.000	0.868
3	0.982	0.971	0.859	0.987	1.000	0.892
5	0.998	1.000	0.991	0.999	1.000	0.991
			n_1=n_2=150			
1	1.000	1.000	0.900	1.000	1.000	0.960
3	0.998	1.000	0.949	0.998	1.000	0.970
5	1.000	1.000	0.998	1.000	1.000	0.999
			n_1=n_2=200			
1	1.000	1.000	0.941	1.000	1.000	0.977
3	1.000	1.000	0.986	1.000	1.000	0.989
5	1.000	1.000	1.000	1.000	1.000	1.000

$\beta = 1$, and censoring rate 20%. The figure shows that the normality of the test statistics is satisfactory, in general, with a sample size of 200. Similar results are obtained for other situations, and hence are not presented here.

Table 3. Simulation results of the two-sample tests with different weights for Case 1 when the distribution functions of the terminal event times are different for the two groups.

	Censoring rate 20%			Censoring rate 40%		
β	$\tilde{T}_n(h_n^{(1)})$	$\tilde{T}_n(h_n^{(2)})$	$\tilde{T}_n(h_n^{(3)})$	$\tilde{T}_n(h_n^{(1)})$	$\tilde{T}_n(h_n^{(2)})$	$\tilde{T}_n(h_n^{(3)})$
			n_1=n_2=100			
1	0.056	0.061	0.061	0.068	0.056	0.072
1.1	0.231	0.221	0.231	0.260	0.241	0.260
1.2	0.688	0.664	0.693	0.721	0.687	0.717
1.3	0.960	0.956	0.960	0.956	0.951	0.955
			n_1=n_2=150			
1	0.043	0.049	0.041	0.056	0.052	0.054
1.1	0.338	0.311	0.341	0.352	0.330	0.348
1.2	0.863	0.835	0.865	0.872	0.851	0.869
1.3	0.993	0.992	0.994	0.994	0.985	0.992
			n_1=n_2=200			
1	0.046	0.048	0.045	0.058	0.057	0.060
1.1	0.447	0.427	0.450	0.469	0.451	0.469
1.2	0.951	0.933	0.953	0.952	0.941	0.954
1.3	1.000	0.998	1.000	0.999	0.999	0.998

Tables 1 and 2 report the size and power of the proposed test statistics $T_n(h_n^{(j)})$ for the three weight processes at a significance level of 0.05 for different values of β for Cases 1 and 2, respectively, where $T_n(h_n^{(j)})$ represents the test statistic with the jth weight process, for $j = 1, 2, 3$. The simulation results show that the proposed test possesses good properties: (i) the estimated sizes are all around the significance level of 0.05; (ii) for any fixed value of β in both cases, the power increases when the sample size increases; (iii) the power increases to one when the discrepancy of the mean functions becomes more evident, implying the consistency of the proposed test; and (iv) the power of the test does not appear to affected by the censoring rate.

5.3. Two-sample test with different distributions of terminal event times

Here, we conduct a parallel simulation study to evaluate the proposed two-sample test for two groups with different distributions of terminal event times. The data sets are generated in a similar way to those in Subsection 5.2, except that the terminal event times $U_i^{(1)}$ and $U_i^{(2)}$ are from $6 + \exp(1)$ and $6 + \exp(2)$, respectively, and the censoring times $C_i^{(1)}$ and $C_i^{(2)}$ are from $6 + \kappa_1 \exp(3)$ and $6 + \kappa_2 \exp(3)$, respectively, where $(\kappa_1, \kappa_2) = (1.375, 5.769)$ and $(0.500, 1.010)$, resulting in censoring rates of 20% and 40%, respectively. The three weight

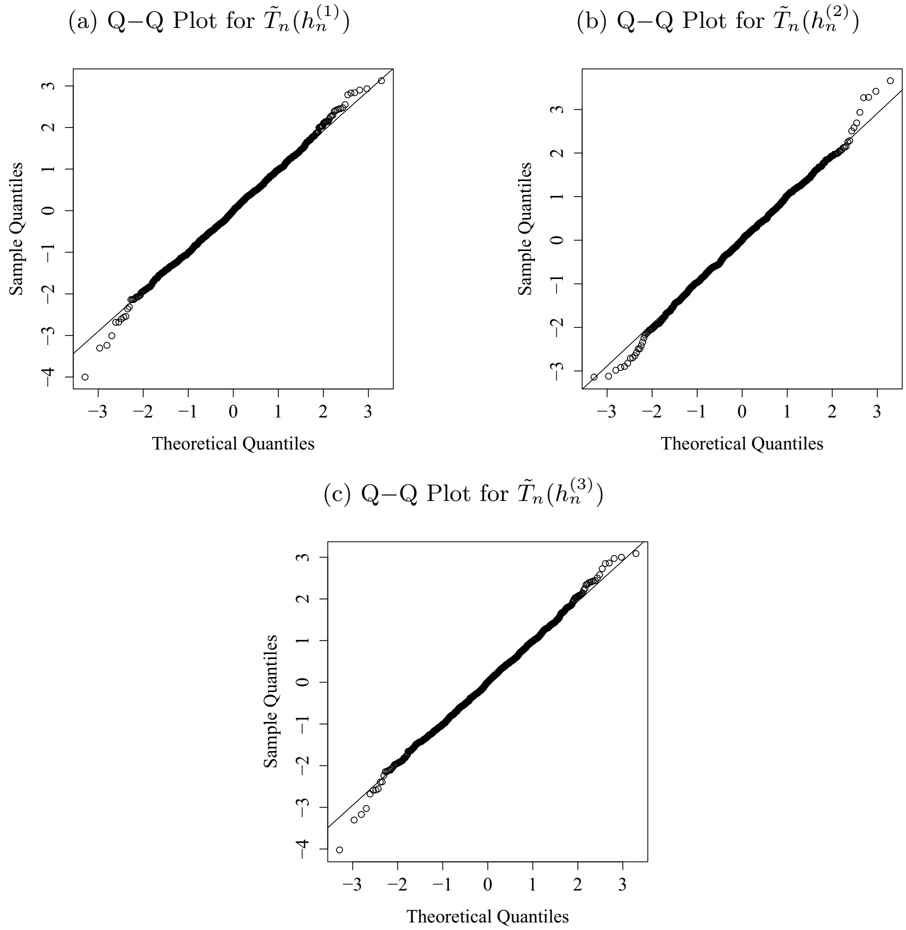

Figure 4. Q–Q plots for n_1=n_2=200, β=1, and censoring rate = 20% when the distribution functions of the terminal event times are different for the two groups.

processes $h_n^{(j)}(t)$, for $j = 1, \ldots, 3$, are chosen as

$$h_n^{(1)}(t) = t,\ h_n^{(2)}(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{(t \geq T_{iK_i})},\ h_n^{(3)}(t) = \frac{\hat{\Lambda}_1(t) + \hat{\Lambda}_2(t)}{2}.$$

To calculate the histogram-type estimators of f_l, we divide $[6, \tau]$ into five equal subintervals. The results of the simulation are again evaluated based on 1,000 replications.

Figure 4 presents quantile plots of the test statistics with the three weight processes against the standard normal distribution with $n_1 = n_2 = 200$, $\beta = 1$, and censoring rate 20% for Case 1. The normality of the test statistics appears to be justified when the sample size is 200. The plots for other situations are

Table 4. Simulation results of the two-sample tests with different weights for Case 2 when the distribution functions of the terminal event times are different for the two groups.

	Censoring rate 20%			Censoring rate 40%		
β	$\tilde{T}_n(h_n^{(1)})$	$\tilde{T}_n(h_n^{(2)})$	$\tilde{T}_n(h_n^{(3)})$	$\tilde{T}_n(h_n^{(1)})$	$\tilde{T}_n(h_n^{(2)})$	$\tilde{T}_n(h_n^{(3)})$
			n_1=n_2=100			
1	1.000	1.000	0.834	1.000	1.000	0.926
3	0.986	1.000	0.884	0.999	1.000	0.947
5	0.997	1.000	0.987	1.000	1.000	0.996
			n_1=n_2=150			
1	1.000	1.000	0.940	1.000	1.000	0.985
3	0.998	1.000	0.976	0.999	1.000	1.000
5	1.000	1.000	0.999	1.000	1.000	0.999
			n_1=n_2=200			
1	1.000	1.000	0.972	1.000	1.000	0.990
3	1.000	1.000	0.994	1.000	1.000	0.995
5	1.000	1.000	1.000	1.000	1.000	1.000

similar, and so are not presented here. Tables 3 and 4 summarize the size and power of the proposed statistics $\tilde{T}_n(h_n^{(j)})$, for $j = 1, 2, 3$, at a significance level of 0.05 for Cases 1 and 2, respectively. The results yield the same conclusions as those in Subsection 5.2.

6. Real-Data Analysis

In this section, we apply the proposed method to a data set from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) for the period 1998 to 2014 (Zeng et al. (2017)). This survey was conducted by the Center for Healthy Aging and Development Studies (CHADS) of the National School of Development at Peking University and the Chinese Center for Disease Control and Prevention (CDC) to study the determinants of healthy human longevity and later-life mortality. The data consist of survey information in seven waves (1998, 2000, 2002, 2005, 2008, 2011, and 2014) from interviews with 9,093 respondents who were 77 years or older in the 1998 baseline survey. In each wave, the respondents were asked to provide information about their health, socioeconomic status, family, lifestyle, and demographic profile on a date randomly chosen in the interview year.

For this data set, we analyze panel count data on serious illness before death, where a serious illness is defined as one that results in a person being bedridden. Then, we compare the incidences of serious illness among older adults living in different areas during the overall study period. The incidences of serious

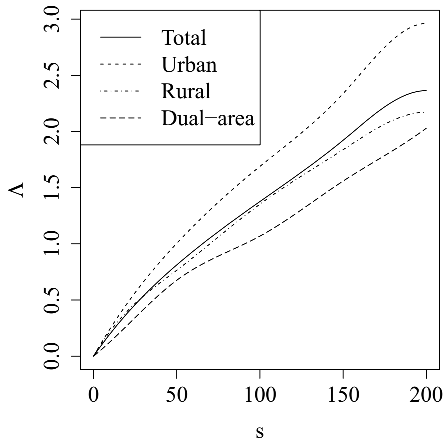

Figure 5. The two-stage estimates of the mean function for the CLHLS data.

illness are counted repeatedly if the respondents were confined to bed again after recovering from the previous serious illness. We analyze data for those subjects who have at least one follow-up observation after 1998 to allow for a meaningful analysis of the panel count data. Therefore, we include data on 4,362 older adults after removing 3,368 individuals, of whom 2,005 died in 2000 or between 1998 and 2000, 894 were lost to follow-up in 2000, and 469 had missing or mistakenly documented records. Of the 4,362 individuals, 1,489 and 1,561 lived only in urban and rural areas, respectively, and 1,312 lived in both areas during this period. For the ith individual, we define T_{ij} as the time (in months) between survey j and the baseline survey, $N_i(t)$ as the accumulated number of serious illnesses suffered by an individual up to time t, and K_i as the total number of follow-up surveys in the study. Death is the terminal event, and loss to follow-up is treated as the censoring event to death. The longest follow-up time is $\tau = 200$ (months), and the censoring rate is 25.72%.

We examine the mean function of the serious illness process using model (2.1). We estimate the mean function using an I-spline with order $d = 3$ and $m_n = 3$ to divide $[0, \tau]$ into four subintervals with internal knots of splines selected at $t_{d+1} = \tau/4$, $t_{d+2} = \tau/2$, $t_{d+3} = 3\tau/4$. Figure 5 plots the estimated mean functions for the combined sample (solid line), urban sample (dash line), rural sample (dotted-dash line), and dual-area sample (long-dash line). The figure shows that urban-area residents tended to experience more serious illnesses than rural-area and dual-area residents did, which may be because the latter two groups had limited access to advanced medical care, compared with their counterparts in urban areas.

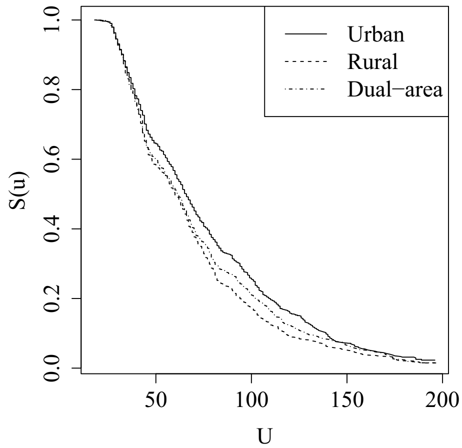

Figure 6. KM estimates of the survival function for the CLHLS data.

We use the same three weight functions $h_n^{(j)}$, for $j = 1, 2, 3$, as those in Subsection 5.2 to conduct the two-sample test for each pair of two mean functions for the urban sample, rural sample, and dual-area sample. That is, the null hypotheses are $H_0^{(1)}$: $\Lambda_U(s) = \Lambda_R(s)$, $H_0^{(2)}$: $\Lambda_U(s) = \Lambda_D(s)$, and $H_0^{(3)}$: $\Lambda_R(s) = \Lambda_D(s)$, where Λ_U, Λ_R, and Λ_D are the mean functions for the urban sample, rural sample, and dual-area sample, respectively. To conduct the test, we first use the log-rank test to conclude that the survival function of death for urban-area residents is significantly different from those of rural-area residents and dual-area residents, with p-values 3×10^{-6} and 0.006, respectively. Furthermore, the log-rank test shows that the survival functions of death for rural-area residents are not significantly different from those of dual-area residents, with a p-value$= 0.09$. Figure 6 shows KM estimates for the survival functions of death for older adults in the three groups. We use the test statistic in Theorem 4 to compare the mean functions of the rural-area and dual-area residents. For the tests $H_0^{(1)}$: $\Lambda_U(s) = \Lambda_R(s)$ and $H_0^{(2)}$: $\Lambda_U(s) = \Lambda_D(s)$, we divide $[0, \tau]$ evenly into eight subintervals to obtain the histogram-type estimators, and then apply the test statistic in Theorem 5. The test results, summarized in Table 5, suggest there is a significant difference in later-life serious illness experience at a level of 0.01 between urban-area and rural-area residents, and between urban-area and dual-area residents. Although the mean functions of the dual-area and rural-area residents are not statistically different (p-value>0.05), the dual-area residents seem to have fewer incidences of severe illness compared with rural-area residents. This may be because many rural-area residents in China have recently moved to urban areas, or because older adults who have lived in both urban and rural areas tend to have

Table 5. Two-sample test results for the three weights for the CLHLS data.

	$h_n^{(1)}$	$h_n^{(2)}$	$h_n^{(3)}$
	$H_0^{(1)}$: $\Lambda_U(s) = \Lambda_R(s)$		
$\tilde{U}_n$	287.904	3.968	4.325
$\tilde{\sigma}_n$	77.849	1.138	1.182
$\tilde{T}_n$	3.698	3.488	3.659
p-value	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
	$H_0^{(2)}$: $\Lambda_U(s) = \Lambda_D(s)$		
$\tilde{U}_n$	368.945	4.132	5.225
$\tilde{\sigma}_n$	79.298	1.193	1.205
$\tilde{T}_n$	4.653	3.465	4.335
p-value	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
	$H_0^{(3)}$: $\Lambda_R(s) = \Lambda_D(s)$		
U_n	84.872	0.384	0.932
σ_n	66.322	0.957	0.833
T_n	1.280	0.401	1.119
p-value	0.201	0.688	0.263

$T_n(h)$ (or $\tilde{T}_n(h)$): the observed value of the test statistic with different weight functions; *** represents a significance level of 0.01.

a better family socioeconomic status and health status compared with rural-area residents. Our conclusions were consistent, regardless of the choice of the weight processes.

7. Discussion

Focusing on the potential associations between recurrent event and terminal event processes that often occur in life sciences, we have developed a conditional nonparametric mean functional model to study the explicit effects of terminal events on a recurrent event process with panel count data. This approach is particularly useful for determining the recurrent event behavior prior to the terminal event, but has challenges in terms of numerical computation and statistical inference because of the nature of panel count data subject to informative termination. We have designed a two-stage spline-based estimation procedure to ease the computational burden, and have examined the asymptotic behavior of the proposed nonparametric estimator using modern empirical process theories. In particular, the asymptotic normality can be readily expanded to address multi-sample hypothesis-testing problems. Our extensive simulation studies demonstrate the good performance of the proposed method in finite-sample settings.

The proposed method can be extended to consider a conditional semiparametric mean function model:

$$E(\tilde{N}(t;U)|U=u,Z)=\exp(\boldsymbol{\beta}^T\mathbf{Z})\Lambda(u-t), 0\leq t\leq u,$$

where $\tilde{N}(t;u)=N(u)-N(u-t)$, $N(t)$ is the number of recurrent events up to time t, U is the terminal event time, $\mathbf{Z}$ is a covariate vector, $\boldsymbol{\beta}$ is an unknown parameter, and $\Lambda(\cdot)$ is an unknown nondecreasing baseline function. This model has a broader application that not only captures the effect of the terminal event, but also allows us to assess external risk factors on a recurrent event process, and is left to future research.

Supplementary Material

The online Supplementary Material provides the calculation of the loss function, some preliminary lemmas, proofs of Theorems 1–5, some additional simulation studies, and a real-data analysis.

Acknowledgments

The authors are grateful to the editor, associate editor, and two referees for their valuable comments and suggestions. This work was supported in part by the Research Grant Council of Hong Kong (15301218) for Xingqiu Zhao, the Natural Science Foundation of China (12171374) for Li Liu, and NIH/NIGMS (2 U54 GM115458-06) for Ying Zhang.

References

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

Chan, K. C. G. and Wang, M.-C. (2010). Backward estimation of stochastic processes with failure events as time origins. *The Annals of Applied Statistics* **4**, 1602–1620.

Chan, K. C. G. and Wang, M.-C. (2017). Semiparametric modeling and estimation of the terminal behavior of recurrent marker processes before failure events. *Journal of the American Statistical Association* **112**, 351–362.

Chan, I. S., Neaton, J. D., Saravolatz, L. D., Crane, L. R. and Osterberger, J. (1995). Frequencies of opportunistic diseases prior to death among HIV-infected persons. Community Programs for Clinical Research on AIDS. *AIDS* **9**, 1145–1151.

Cheng, S. C. and Wei, L. J. (2000). Inferences for a semiparametric model with panel data. *Biometrika* **87**, 89–97.

Földes, A., Rejtő, L. and Winter, B. B. (1981). Strong consistency properties of nonparametric estimators for randomly censored data, II: Estimation of density and failure rate. *Periodica Mathematica Hungarica* **12**, 15–29.

He, X., Feng, X., Tong, X. and Zhao, X. (2017). Semiparametric partially linear varying coefficient models with panel count data. *Lifetime Data Analysis* **23**, 439–466.

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

Hu, X. J., Sun, J. and Wei, L. J. (2003). Regression parameter estimation from panel counts. *Scandinavian Journal of Statistics* **30**, 25–43.

Jiang, H., Su, W. and Zhao, X. (2020). Robust estimation for panel count data with informative observation times and censoring times. *Lifetime Data Analysis* **26**, 65–84.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53**, 457–481.

Kong, S., Nan, B., Kalbfleisch, J. D., Saran, R. and Hirth, R. (2018). Conditional modeling of longitudinal data with terminal event. *Journal of the American Statistical Association* **113**, 357–368.

Liu, L., Wolfe, R. A. and Huang, X. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics* **60**, 747–756.

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. *Journal of the American Statistical Association* **104**, 1060–1070.

Lunney, J. R., Lynn, J., Foley, D. J., Lipson, S. and Guralnik, J. M. (2003). Patterns of functional decline at the end of life. *Journal of Nonparametric Statistics* **289**, 2387–2392.

Ramsay, J. O. (1988). Monotone regression splines in action. *Statistical Science* **3**, 425–441.

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. *Statistica Sinica* **5**, 279–290.

Sun, L., Song, X., Zhou, J. and Liu, L. (2012). Joint analysis of longitudinal data with informative observation times and a dependent terminal event. *Journal of the American Statistical Association* **107**, 688–700.

Sun, J., Tong, X. and He, X. (2007). Regression analysis of panel count data with dependent observation times. *Biometrics* **63**, 1053–1059.

Sun, J. and Wei, L. J. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **62**, 293–302.

Thall, P. F. and Lachin, J. M. (1988). Analysis of recurrent events: Nonparametric methods for random-interval count data. *Journal of the American Statistical Association* **83**, 339–347.

Wellner, J. A. and Zhang, Y. (2000). Two estimator of the mean of a counting process with panel count data. *The Annals of Statistics* **28**, 779–814.

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

Zeng, Y., Vaupel, J., Xiao, Z., Liu, Y. and Zhang, C. (2017). Chinese Longitudinal Healthy Longevity Survey (CLHLS), 1998–2014. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. `https://doi.org/10.3886/ICPSR36692.v1`.

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

Zhao, H., Li, Y. and Sun, J. (2013a). Analyzing panel count data with a dependent observation process and a terminal event. *Canadian Journal of Statistics* **41**, 174–191.

Zhao, H., Li, Y. and Sun, J. (2013b). Semiparametric analysis of multivariate panel count data with dependent observation processes and a terminal event. *Journal of Nonparametric Statistics* **25**, 379–394.

Zhao, H., Zhang, Y., Zhao, X. and Yu, Z. (2019). A nonparametric regression model for panel count data analysis. *Statistica Sinica* **29**, 809–826.

Zhao, X. and Zhang, Y. (2017). Asymptotic normality of nonparametric M-estimators with applications to hypothesis testing for panel count data. *Statistica Sinica* **27**, 931–950.

Zhao, X., Zhou, J. and Sun, L. (2011). Semiparametric transformation models with time-varying coefficients for recurrent and terminal events. *Biometrics* **67**, 404–414.

Zhou, J., Zhang, H., Sun, L. and Sun, J. (2017). Joint analysis of panel count data with an informative observation process and a dependent terminal event. *Lifetime Data Analysis* **23**, 560–584.

Xiangbin Hu

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China.

E-mail: xiang-bin.hu@connect.polyu.hk

Li Liu

School of Mathematics and Statistics, Wuhan University, Wuhan, China.

E-mail: lliu.math@whu.edu.cn

Ying Zhang

Department of Biostatistics, University of Nebraska Medical Center, Omaha NE 68198, USA.

E-mail: ying.zhang@unmc.edu

Xingqiu Zhao

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China.

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

(Received June 2021; accepted April 2022)