NEW NONPARAMETRIC TESTS FOR PANEL COUNT DATA

NEW NONPARAMETRIC TESTS FOR PANEL COUNT DATA

By

XINGQIU ZHAO, Ph.D.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2008) McMaster University

(Mathematics) Hamilton, Ontario

TITLE: New Nonparametric Tests for Panel Count Data

AUTHOR: Xingqiu Zhao, Ph.D.

(Wuhan University, P. R. China)

SUPERVISOR: Professor N. Balakrishnan

NUMBER OF PAGES: xv, 146

Abstract

Statistical analysis of panel count data is an important topic to a number of applied fields including biology, engineering, econometrics, medicine, and public health. Panel count data include observations on subjects over multiple time points where the response variable is a count or recurrent event process when only the numbers of events occurring between observation time points are available. The choice of method for analyzing panel count data usually depends on the relationship between the observation times and the response variable and questions of interest. Most of the previous research was done when the observation times are fixed. If the observation times are random, the data structure becomes more challenging since the observation times for individual subjects vary in addition to the incompleteness of observations. The model-based approach was used to deal with such data. However, this method relies on extra assumptions on the observation scheme and thus is restrictive in practice. In this dissertation, we discuss the problem of multi-sample nonparametric comparison of counting processes with panel count data, which arise naturally when recurrent events are considered. For the problem considered, we develop some new nonparametric tests.

First, we construct a class of nonparametric test statistics based on the integrated weighted differences between the estimated mean functions of the count processes,

where the isotonic regression estimate is used for the mean functions. The asymptotic distributions of the proposed statistics are derived and their finite-sample properties are examined through Monte Carlo simulations. A panel count data from a cancer study is analyzed and presented as an illustrative example.

As shown through Monte Carlo simulations, the nonparametric maximum likelihood estimator (NPMLE) of the mean function is more efficient than the nonparametric maximum pseudo-likelihood estimator (NPMPLE). However, no nonparametric tests have been discussed in the literature for panel count data based on the NPMLE since the NPMLE is more complicated both theoretically and computationally. It is, therefore, particularly important to develop nonparametric tests based on the NPMLE for panel count data.

In the second part of the dissertation, we focus on the situation when treatment indicators can be regarded as independent and identically distributed random variables and propose a nonparametric test in this case using the maximum likelihood estimator. The asymptotic property of the test statistic is derived. Simulation studies are carried out which suggest that the proposed method works well for practical situations, and is more powerful than the existing tests based on the NPMPLEs of the mean functions.

In the third part of the dissertation, we consider more general situations. We construct a class of nonparametric tests based on the accumulated weighted differences between the rates of increase of the estimated mean functions of the counting processes over observation times, where the nonparametric maximum likelihood approach is used to estimate the mean functions instead of the nonparametric maximum pseudo-likelihood. The asymptotic distributions of the proposed statistics are derived and their finite-sample properties are evaluated by means of Monte Carlo simulations. The

simulation results show that the proposed methods work quite well and the tests based on NPMLE are more powerful than those based on NPMPLE. Two real data sets are analyzed and presented as illustrative examples.

The last part of the dissertation discusses a special type of panel count data, namely, current status or case 1 interval-censored data. Such data often occur in tumorigenicity experiments. For nonparametric two-sample comparison based on censored or interval-censored data, most of the existing methods have focused on testing the hypothesis that specifies the two population distributions to be identical under the assumption that observation or censoring times have the same distribution. We consider the nonparametric Behrens-Fisher hypothesis (NBFH) under this settings. For this purpose, we study the asymptotic property of the nonparametric maximum likelihood estimator of the probability that an observation from the first distribution exceeds an observation from the second distribution. A nonparametric test for the NBFH is proposed and the asymptotic normality of the proposed test is established. The method is evaluated using simulation studies and illustrated by a set of real data from a tumorigenicity experiment.

In this thesis, Chapters 2-5 are based on the following papers:

1. Balakrishnan, N. and Zhao, X. (2008). A class of multi-sample nonparametric tests for panel count data. *Annals of the Institute of Statistical Mathematics* (revised).

2. Balakrishnan, N. and Zhao, X. (2008). A new nonparametric test for the equality of counting processes with panel count data. *Canadian Journal of Statistics* (submitted).

3. Balakrishnan, N. and Zhao, X. (2008). New multi-sample nonparametric tests for panel count data. *The Annals of Statistics* (to appear).

4. Balakrishnan, N. and Zhao, X. (2008). Nonparametric Behrens-Fisher Hypothesis Testing for Case 1 Interval-Censored Data. *Under preparation.*

Acknowledgements

I would like to express my sincere appreciation to my supervisor, Professor N. Balakrishnan, for his guidance, support, encouragement, great patience, and careful reading of the manuscript. My appreciation is beyond all words and speech to him.

I am very grateful to members of my supervisory committee: Dr. Roman Viveros-Aguilera and Dr. Rong Zhu for their helpful discussions and comments on my work.

I would also like to thank all the professors in McMaster University who have ever taught me, for their guidance, patience and understanding.

I thank Dr. Ying Zhang and Dr. Minggen Lu for supplying the R code for the modified iterative convex minorant algorithm of Wellner and Zhang (2000).

Finally, I would like to acknowledge the financial support for my PhD studies provided by NSERC (CGS D3).

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Panel Count Data

Statistical analysis of panel count data is an important topic in applied fields including biology, engineering, econometrics, medicine, and public health. By panel count data, we mean that, for each subject at study, observations are taken at several distinct time points and only the numbers of recurrent events that have occurred before observation times are known or only the numbers of events occurring between observation time points are available. No information is available on the exact times of the recurrent event during the study. In addition, the number of observation times and observation times may vary from subject to subject. Such data frequently occur in medical follow-up studies, reliability experiments, AIDS clinical trials, animal tumorigenicity experiments, and sociological studies, for example.

When each subject is observed at a single time point, panel count data are referred to as current status or case 1 interval-censored data; when each subject is observed

exactly twice, panel count data are referred to as case 2 interval-censored data.

1.1.1 Bladder Tumor Study

The bladder tumor study presented here was conducted by the Veterans Administration Co-operative Urological Research Group (VACURG). It is a follow-up study on patients with bladder tumors. The data are presented in Andrews and Herzberg (1985). The data were obtained from a randomized clinical trial. All patients had superficial bladder tumors when they entered the trial, and they were assigned randomly to one of three treatments: placebo, thiotepa and pyridoxine. At subsequent follow-up visits, any tumors noticed were removed and treatment was continued. The data consist of the number of recurrences experienced for each of 116 patients, the number of tumors present initially at the time of randomization in the trial and the diameter of the largest of these, the months from the beginning of the study until each recurrence, the number of tumors present at each recurrence, and the diameter of the largest of these. The main objective of this study is to determine the effect of treatment on the frequency of tumor recurrence. For this purpose, many authors analyzed these data using different methods of inference; see Byar et al. (1977), Byar (1980), Wellner and Zhang (2000), Sun and Wei (2000), Zhang (2002), Sun and Fang (2003), Zhang (2006), and Park et al. (2007), among others. Sun and Wei (2000) and Zhang (2002) concluded that thiotepa effectively reduces the recurrence of tumors. However, the results from the analysis in Sun and Fang (2003), Zhang (2006), and Park et al. (2007) suggested that the effect of treatment is not significant on the rate of tumor recurrence.

1.1.2 Floating Gallstone Study

Schoenfield et al. (1981) and Thall and Lachin (1988) described the recurrent biliary-symptom data from the National Cooperative Gallstone Study (NCGS). It is a follow-up study on patients with floating gallstones. The NCGS was a ten year multicenter, double blind, and placebo-controlled clinical trial on the usage of chenodiol for the dissolution of cholesterol gallstones (Schonefield et al., 1981). In this study, there were 916 patients who were randomized to placebo, low dose, or high dose group. The problem of interest here is to determine difference between the high-dose chenodiol and placebo groups in terms of the incidence rates of nausea since the low dose proved to be ineffective. The patients were scheduled to return for clinical visits at prespecified times. However, actual visit times differ from patient to patient and for each patient, observations include the numbers of nausea, a symptom relating to the disease, between clinical visits. So the data consist of the successive visit times and the associated counts of episodes of nausea for 113 patients with floating gallstones in the high dose and placebo groups. Thall and Lachin (1988) analyzed the study using some grouping techniques and suggested that the incidence of nausea in the high-dose chenodiol group differs significantly from that of the placebo group over the first year. Schoenfield et al. (1981) concluded that chenodiol was not more effective than placebo in reducing the number of episodes of nausea based on a simpler analysis. Sun and Fang (2003) and Park et al. (2007) suggested that the incidence rates of nausea do not differ significantly between the high-dose chenodiol and placebo groups by applying the nonparametric tests based on isotonic regression and the nonparametric maximum pseudo-likelihood, respectively.

1.1.3 Lung Tumors in Mice

Hoel and Walburg (1972) described an example from a tumorigenicity study; see also Dinse and Lagakos (1983), Finkelstein and Wolfe (1985), Finkelstein (1986). 144 RFM mice were assigned randomly to one of two groups: the germ-free and conventional environment, and they ware examined at sacrifice or death for evidence of malignancy. The purpose of this study is to compare the time from beginning of the study until the time to observe a tumor to determine whether a suspected agent accelerates the time to lung tumor onset. Since lung tumors are nonlethal and cannot be observed before death in RFM mice, it is appropriate to treat these data as current status or case 1 interval censored data.

In this study, data on only the age at death and whether the tumor is present at that time are available. The dataset includes 96 conventional mice and 48 germ- free mice, of which 27 conventional mice and 35 germ-free mice developed lung rumors. These data have been examined by Hoel and Walburg (1972) for illustrating the importance of identifying left-censored data, and was also analyzed by Dinse and Lagakos (1983), Finkelstein and Wolfe (1985), Huang (1996), Sun (1999), and Shen (2000).

1.2 Analysis of Panel Count Data

The analysis of Panel Count Data has recently attracted considerable attention (e.g. Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000; Sun and Wei, 2000; Cheng and Wei, 2000; Zhang, 2002; Sun and Fang, 2003; Hu et al.; 2003; Zhang and Jamshidian, 2003; Wellner et al. 2004; Sinha and Maiti, 2004; Zhang, 2006; Huang et al., 2006; Park

et al., 2007; Sun et al., 2007). Panel count data include the number of observation times (K), discrete observation times ($T = (T_{K,1}, \ldots, T_{K,K})$) and the counts of recurrent event ($N = (N(T_{K,1}), \ldots, N(T_{K,K}))$ for each study subject i, where $N(t)$ represents the total number of recurrent events occurring up to time t. The main goals are to estimate the mean function of $N(t)$ and compare the effects of treatment.

When observation times are fixed and the same for each subject, a number of methods have been developed. Kalbfleisch and Lawless (1985) discussed the fitting of a finite state Markov model to panel count data. Hinde (1982) and Breslow (1984) considered regression analysis of Poisson count data. Thall (1988) investigated some regression models for mixed Poisson processes. Liang and Zeger (1986) and Thall and Vail (1990) presented quasi-likelihood regression models with a generalized estimating equation (GEE) approach by treating panel count data as longitudinal count data. Also, Cameron and Trivedi (1998) provided a review of parametric and semiparametric methods for the regression analysis of panel count data.

When observation times are random, there exists limited research. Sun and Kalbfleisch (1995) applied isotonic regression techniques to estimate the mean function. Wellner and Zhang (2000) studied two nonparametric estimators based on the maximum pseudo-likelihood and the maximum likelihood approaches assuming that the counting process $N(t)$ is a non-homogeneous Poisson process and is independent of K and T. Wellner and Zhang (2000) also investigated the asymptotic properties of both estimators and showed that the isotonic regression estimator is equivalent to a pseudo-maximum likelihood estimator. However, the pseudo-maximum likelihood estimator is established by ignoring the dependence among the counts within a subject and is less efficient than the maximum likelihood estimator, a fact seen through simulation stud-

ies. Zhang and Jamshidian (2003) introduced the gamma frailty variable to account for correlation among the panel counts and still used the maximum pseudo-likelihood approach to estimate the mean function.

For treatment comparison based on panel count data, most of the existing methods focus on semiparametric regression analysis based on the models of rate and mean functions. Sun and Wei (2000) and Zhang (2002) discussed regression analysis of panel count data by using estimating equation-based methods and the semiparametric pseudo-likelihood approach, respectively, under the assumption that the observation times are independent of occurrences of the recurrent event under study given covariates. Wellner et al. (2004) also investigated maximum likelihood estimation for regression analysis under this assumption. Cheng and Wei (2000) and Hu et al. (2003) also investigated estimating equation approaches for the case when observation and censoring times are independent of the event process and the case when censoring time is independent of observation times, event process and covariates, but observation times may depend on the event process through covariates, respectively. Sinha and Maiti (2004) developed a Bayesian analysis of panel count data when the censoring time may be correlated with the underlying counting process of interest by assuming that all the subjects have fixed observation times. Huang et al. (2006) studied nonparametric and semiparametric models that allow observation times to be correlated with the event process through a frailty variable, and used the conditional likelihood approach to estimate the baseline function and the regression parameters. Sun et al. (2007) also investigated semiparametric models for rate functions of the observation process and the event process, where both processes may be correlated through a subject-specific latent variable or frailty, and used the same approach as that in Sun and Wei (2000)

for estimation of regression parameters. However, the estimating equation approach still ignored correlations among event counts within a subject.

For the analysis of panel count data, three model-free approaches are available in the literature. One is given by Thall and Lachin (1988) who suggested to use specially defined intervals. The test result may depend on the specific way of grouping the data. Sun and Fang (2003) presented a model-free approach based on the assumption that the treatment indicators can be regarded as independent and identically distributed random variables, where the isotonic regression estimator is used for the mean function. However, this assumption is quite strong and may not hold in practice. Park et al. (2007) presented a class of nonparametric two-sample tests based on the isotonic regression estimator. Zhang (2006) also presented nonparametric k-sample tests by using nonparametric maximum pseudo-likelihood approach. To relax the assumptions about weight processes required by Zhang (2006), one needs to modify the proofs for asymptotic properties of the test statistic given by Zhang (2006). This issue will be discussed in Chapter 2. No methods have been developed based on the nonparametric maximum likelihood approach. This motivates our research in Chapters 3 and 4.

1.3 Analysis of Interval-Censored Data

Interval-censored data are a special type of panel count data. When each subject experiencing recurrent events over time in an experiment is observed at only one time point and no information is available on subjects between their entry time and observation time point, such data are referred to as current status or case 1 interval-censored data (Groeneboom and Wellner, 1992). When each subject during the study is observed

at only two time points and no information is available on subjects among their entry time and two observation time points, such data are referred to as case 2 interval-censored (Groeneboom and Wellner, 1992). One field in which interval-censored data often occur is periodic follow-up studies where patients are supposed to be inspected at prespecified observation times. In this case, it is common that patients miss some prespecified observations and/or are observed at different times rather than the prespecified times, thus resulting in interval-censored failure time. One such example from a cancer study is provided in Finkelstein (1986). Another field which commonly produces interval-censored failure time data is tumorigenicity experiments. An example from a tumorigenicity experiment is provided in Dinse and Lagakos (1983) and will be discussed in more detail in Chapter 5.

Survival comparison is usually one of the main goals in survival studies. For the problem, when right-censored failure time data are available, a number of well-established methods have been developed (Fleming and Harrington, 1991; Kalbfleisch and Prentice, 2002). For the case of interval-censored failure time data, many authors have discussed the problem. For example, Peto and Peto (1972) considered the two-sample comparison problem under the Lehmann-type alternatives $G_2(t) = G_1^{\theta}(t)$, where G_1 and G_2 are survival functions corresponding to the two different samples and θ is a parameter. In this case, the comparison problem reduces to testing $\theta = 1$ and they suggested to use the score test, which they referred to as the log-rank test. Assuming the proportional hazards model, a special case of Lehmann-type alternatives, Finkelstein (1986) investigated the general k-sample comparison problem. For the problem, she also suggested to apply the score test for testing regression parameters equal to zero. Following Finkelstein (1986), Sun (1996) studied the same problem

without assuming the proportional hazards model and developed a nonparametric test using the idea of the log-rank test for right-censored data (Kalbfleisch and Prentice, 2002). Groeneboom and Wellner (1992) studied nonparametric maximum likelihood estimation of the distribution function with interval-censored data. The nonparametric maximum likelihood estimator converges to the true distribution at rate $n^{1/3}$, unlike the empirical distribution function, or the Kaplan-Meier estimator, both of which converge at the more familiar $n^{1/2}$ rate. Its limiting distribution is not Gaussian, but a more complex distribution associated with two-sided Brownian motion (Groenboom and Wellner, 1992). Furthermore, Huang and Wellner (1995) show that nonparametric maximum likelihood estimates of smooth functionals for case 1 interval-censored data converge at rate $n^{1/2}$. Andersen and Ronn (1995) constructed a nonparametric two-sample test based on the asymptotic results for case 1 interval-censored data. Zhang and Liu (2001) developed a nonparametric two-sample test based on a smooth functional of the nonparametric maximum pseudo-likelihood estimator for case 2 interval-censored data. Sun et al. (2005) presented a class of nonparametric tests to the case of interval-censored data, which are generalizations of the log-rank test statistic given in Peto and Peto (1972). Zhao et al. (2008) also developed nonparametric test procedures for partly interval-censored data which are generalizations of the log-rank test statistic discussed in Peto and Peto (1972) and those in Sun et al. (2005). Some other existing test procedures for interval-censored data can be found in Sun (1998, 2006) and Zhu et al. (2008).

For the problem of two-sample comparison with interval-censored data, the existing test procedures usually focus on the identity hypothesis. No methods have been developed for testing other null hypothesis based on interval-censored data. In particular,

Troendle and Yu (2006) applied nonparametric likelihood techniques to obtain tests for either the identity hypothesis or the nonparametric Behrens-Fisher hypothesis for right-censored data, but the asymptotic distributions of the test statistics under the null hypothesis are not established. This motivates our research in Chapter 5.

1.4 Counting Processes

In this section, we give a brief introduction of some basic concepts about counting processes, which play an essential role in the development of statistical models for event history analysis. Aalen (1975, 1978) made a significant contribution for event history analysis based on counting processes. He showed how the theory of multivariate counting processes provides a general framework for event history analysis (Andersen and Borgan, 1985). Andersen et al. (1993) provided a detailed description for use of counting processes in statistical analysis.

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and $\mathcal{T} = [0, \tau)$ be a continuous time interval, where τ is a given terminal time, $0 < \tau \leq \infty$. A stochastic process X is a family of random variables $\{X(t) : t \in \mathcal{T}\}$. A filtration or history, $(\mathcal{F}_t : t \in \mathcal{T})$, is as an increasing right-continuous family of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_t$ contains all the information generated by the stochastic process X on $[0, t]$. The process X is said to be adapted to the filtration if $X(t)$ is $\mathcal{F}_t$-measurable for every $t \in \mathcal{T}$. A process X is predictable with respect to $\mathcal{F}_t$ if $X(t)$ is known given the history $\mathcal{F}_{t-}$ where $\mathcal{F}_{t-}$ is generated by $(X(s), 0 \leq s < t)$.

A counting process is a stochastic process $\{N(t); t \geq 0\}$ such that $N(0) = 0$ and $N(t) < \infty$ a.s., and the paths are right-continuous with probability one, piecewise

constant, and have only jump discontinuities with jumps of size +1.

A Poisson process is a counting process $\{N(t); t \geq 0\}$ such that

$$P\{N(t+dt) - N(t) = 1 | \mathcal{F}_{t-}\} = \lambda(t)dt + o(dt)$$

and

$$P\{N(t+dt) - N(t) \geq 2 | \mathcal{F}_{t-}\} = o(dt)$$

where $\lambda(t) \geq 0$ is a left continuous function satisfying $\int_0^t \lambda(s)ds = \Lambda(t) < \infty$. Here, $\lambda(t)$ and $\Lambda(t)$ are called the intensity and cumulative intensity functions of the Poisson process. The above conditions can also be equivalently written as

$$P\{N(t+dt) - N(t) = 1 | \mathcal{F}_{t-}\} = \lambda(t)dt + o(dt)$$

and

$$P\{N(t+dt) - N(t) = 0 | \mathcal{F}_{t-}\} = 1 - \lambda(t)dt + o(dt).$$

The Poisson process defined above is also known as a nonhomogeneous Poisson process. If $\lambda(t)$ is time invariant, it is called a homogeneous Poisson process. For a Poisson process $\{N(t); t \geq 0\}$, we have $N(t) \sim \text{Poisson}(\Lambda(t))$. Note that $E(N(t)) = \Lambda(t)$. Therefore, $\Lambda(t)$ is also called the mean function of the Poisson process.

1.5 Empirical Processes

Empirical process theory has provided a collection of extremely powerful tools for establishing asymptotic properties of test statistics. In this dissertation, large sample properties of all the proposed test statistics will be studied by using empirical process theory. Here, we briefly introduce some notation, definitions and useful results on

empirical processes. The detailed theory can be found in Var der Vaart and Wellner (1996).

Let $(\Omega, \mathcal{A}, P)$ be an arbitrary probability space and $T : \Omega \mapsto \bar{R}$ an arbitrary map. The *outer integral* of T with respect to P is defined as

$$E^*T = \inf\left\{EU : U \geq T, U : \Omega \mapsto \bar{R} \text{ measurable and } EU \text{ exists}\right\}.$$

For any $B \in \Omega$, the *outer probability* of B is

$$P^*(B) = \inf\left\{P(A) : A \supset B, A \in \mathcal{A}\right\}.$$

Let $X_1, \ldots, X_n$ be a random sample on a measurable space $(\mathcal{X}, \mathcal{A})$. The empirical measure P_n is defined by

$$P_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$

where δ_x is the dirac measure. Let $\mathcal{F}$ be a collection of measurable functions $f : \mathcal{X} \to R$. Then the empirical measure induces a map from $\mathcal{F}$ to R given by

$$f \mapsto P_n f.$$

Here, we use the abbreviation $Qf = \int f dQ$ for $f \in \mathcal{F}$ and measure Q. Let P be the common distribution of the X_i. The $\mathcal{F}$-indexed empirical process $\mathbb{G}_n$ is defined by

$$\mathbb{G}_n f = \sqrt{n}(P_n - P)f = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{f(X_i) - Pf\}.$$

Define

$$||Q||_{\mathcal{F}} = \sup\{|Qf| : f \in \mathcal{F}\}.$$

A class $\mathcal{F}$ is called a *P-Glivenko-Cantelli* class if

$$||P_n - P||_{\mathcal{F}} \longrightarrow 0.$$

A class $\mathcal{F}$ is called a *P-Donsker class* if $\mathbb{G}_n = \sqrt{n}(P_n - P)$ converges weakly to $\mathbb{G}$, where the limit $\mathbb{G}$ is a tight Borel measurable and uniformly bounded function on $\mathcal{F}$.

The *covering number* $N(\varepsilon, \mathcal{F}, ||\cdot||)$ is defined as the minimal number of balls $\{g : ||g - f|| \leq \varepsilon\}$ of radius ε needed to cover the set $\mathcal{F}$. The centers of the balls need not belong to $\mathcal{F}$, but they should have finite norms. The *entropy* is the logarithm of the covering number.

Given two functions l and u, the bracket $[l, u]$ is the set of all functions f with $l \leq f \leq u$. An ε-bracket is a bracket $[l, u]$ with $||u - l|| < \varepsilon$. The *bracketing number* $N_{[]}(\varepsilon, \mathcal{F}, ||\cdot||)$ is the minimum number of ε-brackets needed to cover $\mathcal{F}$. The *entropy with bracketing* is the logarithm of the bracketing number. Here the upper and lower bounds u and l of brackets need not belong to $\mathcal{F}$ themselves but they are assumed to have finite norms. Obviously,

$$N(\varepsilon, \mathcal{F}, ||\cdot||) \leq N_{[]}(2\varepsilon, \mathcal{F}, ||\cdot||).$$

We write $N(\varepsilon, \mathcal{F}, L_r(Q))$ and $N_{[]}(\varepsilon, \mathcal{F}, L_r(Q))$ for covering and bracketing numbers with respect to $L_r(Q)$ norm

$$||f||_{Q,r} = \left(\int |f|^r\right)^{1/r},$$

respectively.

An *envelope function* of a class $\mathcal{F}$ is any function $x \mapsto F(x)$ such that $|f(x)| \leq F(x)$, for every x and f. The minimal envelope function is $x \mapsto \sup_f |f(x)|$.

The *uniform entropy numbers* with L_r are defined as

$$\sup_Q \log N(\varepsilon ||F||_{Q,r}, \mathcal{F}, L_r(Q)),$$

where the supremum is over all probability measures Q on $(\mathcal{X}, \mathcal{A})$, with $0 < QF^r < \infty$.

Define a seminorm ρ_P by

$$\rho_P(f) = \left(P(f - Pf)^2\right)^{1/2}.$$

The empirical process is *asympotically continuous*, if for every $\varepsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P^* \left\{ \sup_{\rho_P(f-g) < \delta} |\mathbb{G}_n(f - g)| \right\} = 0.$$

It is equivalent to the following statement: for every decreasing sequence $\delta_n \downarrow 0$,

$$||\mathbb{G}_n||_{\mathcal{F}_{\delta_n}} \xrightarrow{P^*} 0,$$

where

$$\mathcal{F}_\delta = \{f - g : f, g \in \mathcal{F}, \rho_P(f - g) < \delta\}.$$

Var der Vaart and Wellner (1996) showed that

(1) if

$$\int_0^\infty \sqrt{N_{[]}(\varepsilon, \mathcal{F}, L_2(P))}\, d\varepsilon < \infty,$$

then $\mathcal{F}$ is P-Donsker.

(2) If

$$\int_0^\infty \sup_Q \sqrt{\log N(\varepsilon ||F||_{Q,2}, \mathcal{F}, L_2(Q))}\, d\varepsilon < \infty,$$

then $\mathcal{F}$ is P-Donsker for P such that $P^*F^2 < \infty$ for some envelope function F and the classes $\mathcal{F}_\delta$ and $\mathcal{F}_\infty^2$ are measurable for every $\delta > 0$.

(3) If $\mathcal{F}$ is P-Donsker, then $\mathbb{G}_n$ is asymptotically continuous.

1.6 Outline of the Dissertation

This dissertation discusses the problem of multi-sample nonparametric comparison of counting processes with panel count data, which arise naturally when recurrent events are considered. For the problem considered, we develop some new nonparametric tests. The rest of the dissertation is organized as follows.

In Chapter 2, we construct a class of nonparametric test statistics based on the integrated weighted differences between the estimated mean functions of the count processes, where the isotonic regression estimate is used for the mean functions. The asymptotic distributions of the proposed statistics are derived and their finite-sample properties are examined through Monte Carlo simulations. A panel count data from a cancer study is analyzed and presented as an illustrative example.

In Chapter 3, we focus on the situation when treatment indicators can be regarded as independent and identically distributed random variables and propose a nonparametric test using the maximum likelihood estimator. The asymptotic distribution of the test statistic is derived. Simulation studies are conducted to evaluate the performance of the proposed test. The proposed test procedure is then applied to the analysis of a floating gallstones study to present as an illustrative example.

In Chapter 4, we consider more general situations. We construct a class of nonparametric tests based on the accumulated weighted differences between the rates of increase of the estimated mean functions of the counting processes over observation times, where the nonparametric maximum likelihood approach is used to estimate the mean functions instead of the nonparametric maximum pseudo-likelihood. The asymptotic distributions of the proposed statistics are derived and their finite-sample

properties are evaluated through Monte Carlo simulations. Two real data sets are analyzed and presented as illustrative examples.

Chapter 5 discusses a special type of panel count data, namely, current status or case 1 interval-censored data. We consider the nonparametric Behrens-Fisher hypothesis (NBFH) in this situation. For this purpose, we study the asymptotic property of the nonparametric maximum likelihood estimator of the probability that an observation from the first distribution exceeds an observation from the second distribution. A nonparametric test for the NBFH is proposed and the asymptotic normality of the proposed test is established. The method is evaluated by means of simulations and illustrated by a set of real data from a tumorigenicity experiment.

Chapter 6 presents brief conclusions and some directions for possible future research.

Chapter 2

Multi-sample Nonparametric Comparison of Counting Processes with Panel Count Data

2.1 Introduction

Consider a study that concerns some recurrent event and suppose that each subject in the study gives rise to a point process $N(t)$, denoting the total number of occurrences of the event of interest up to time t. Also suppose that for each subject, observations include only the values of $N(t)$ at discrete observation times or the numbers of occurrences of the event between the observation times. Such data are usually referred to as *panel count data* (Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000). Our focus here will be on the situation when such a study involves k groups. Let $\Lambda_l(t)$ denote the mean function of $N(t)$ corresponding to the lth group for $l = 1, \ldots, k$. The problem

of interest is then to test the hypothesis $H_0 : \Lambda_1(t) = \cdots = \Lambda_k(t)$.

Several authors have discussed the analysis of recurrent event data when each subject in the study is observed continuously over an interval or when the exact times of occurrences of the recurrent event are known. For example, the book by Andersen et al. (1993) presents many of the commonly used statistical methods for the analysis of recurrent event data. In contrast, there exists limited research on the analysis of panel count data. Sun and Kalbfleisch (1995) and Wellner and Zhang (2000) studied estimation of the mean function of $N(t)$. Sun and Wei (2000) and Zhang (2002) discussed regression analysis for such data. To test the hypothesis H_0, Thall and Lachin (1988) suggested to transform the problem to a multivariate comparison problem and then apply a multivariate Wilcoxon-type rank test. Sun and Fang (2003) proposed a nonparametric procedure for this problem, but their procedure depends on the assumption that treatment indicators can be regarded as independent and identically distributed random variables, which may not be the case in practice. In addition to follow-up studies and reliability experiments, panel count data are also encountered in AIDS clinical trials, animal tumorgenicity experiments, and sociological studies.

The remainder of this chapter is organized as follows. Section 2.2 discusses a nonparametric test for the hypothesis H_0 when only panel count data are available and then presents a class of nonparametric test statistics. The statistics, motivated by similar statistics in survival analysis, are formulated as the integrated weighted difference between the estimated mean functions corresponding to the pooled data and each group. To estimate the mean function, the isotonic regression estimate is used (Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000). In Section 2.3, the asymptotic normality of these test statistics is established. In Section 2.4, finite-sample properties

of the proposed test statistics are examined through Monte Carlo simulations. In Section 2.5, we apply the proposed methods to a data from a bladder tumor study. Finally, in Section 2.6, some concluding remarks are made.

2.2 Statistical Methods

Consider a longitudinal study that is concerned with some recurrent event and involves n independent subjects, n_l in the lth group with $n_1 + \cdots + n_k = n$. Let $N_i(t)$ denote the point process arising from subject i and $\Lambda_l(t)$ $(l = 1, \ldots, k)$ be defined as before, for $i = 1, \ldots, n$. Suppose that each subject is observed only at discrete time points $0 < t_{i,1} < \cdots < t_{i,k_i}$ and that no information is available about $N_i(t)$ between observation times; that is, only panel count data are available. Let $n_{i,j} = N_i(t_{i,j})$ be the observed value of N_i at $t_{i,j}$, $j = 1, \ldots, k_i$, $i = 1, \ldots, n$.

To propose the test statistics, we first introduce the isotonic regression estimator of the mean functions (Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000). For simplicity, assume that H_0 is true, and let $\Lambda_0(t)$ denote the common mean function of the $N_i(t)$'s. Further, let $s_1, \ldots, s_m$ denote the ordered distinct observation times in the set $\{ t_{i,j} ;\ j = 1, \ldots, k_i,\ i = 1, \ldots, n \}$ and w_ℓ and $\bar{n}_\ell$ be the number and mean value, respectively, of observations made at time s_ℓ, $\ell = 1, \ldots, m$. Then, the isotonic regression estimator $\hat{\Lambda}_n(t)$ is defined as a nondecreasing step function with possible jumps at the s_ℓ's, and is given by

$$\hat{\Lambda}_n(s_\ell) = \max_{r \le \ell} \min_{s \ge \ell} \frac{\sum_{v=r}^{s} w_v \bar{n}_v}{\sum_{v=r}^{s} w_v} = \min_{s \ge \ell} \max_{r \le \ell} \frac{\sum_{v=r}^{s} w_v \bar{n}_v}{\sum_{v=r}^{s} w_v}, \ \ell = 1, \ldots, m,$$

the isotonic regression of the $\bar{n}_\ell$'s with weights w_ℓ's (Robertson et al., 1988). Wellner and Zhang (2000) established its consistency and also derived its asymptotic distribu-

tion at a fixed time point. Note that the well-known Nelson-Aalen estimator is not available here, since it is applicable only for recurrent event data (Andersen et al., 1993).

Let $\hat{\Lambda}_{n_l}$ denote the isotonic regression estimate of Λ_l based on samples from all the subjects in the lth group. To test the hypothesis H_0, motivated by an idea commonly used in survival analysis (Pepe and Fleming, 1989; Cook et al., 1996; Zhang et al., 2001), we propose the statistic

$$U_n^{(l)} = \sqrt{n} \int_0^{\tau} W_n^{(l)}(t) \{ \hat{\Lambda}_n(t) - \hat{\Lambda}_{n_l}(t) \} \, d\, G_n(t) \ , \ l = 1, \ldots, k,$$

where τ is the largest observation time, $W_n^{(l)}(t)$'s are bounded weight processes, and $G_n(t)$ is defined by

$$G_n(t) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{k_i} I(t_{i,j} \leq t),$$

where $I(s \leq t) = 1$, if $t \geq s$; otherwise, $I(s \leq t) = 0$. The statistic $U_n^{(l)}$ is the integrated weighted difference between $\hat{\Lambda}_n$ and $\hat{\Lambda}_{n_l}$. It is important to mention that some statistics similar to $U_n^{(l)}$ are commonly used in survival analysis. For the two-sample survival comparison with right-censored data, for example, Pepe and Fleming (1989) proposed some test statistics that have a form similar to $U_n^{(l)}$ with $\hat{\Lambda}_n$ and $\hat{\Lambda}_{n_l}$ replaced by the corresponding estimated survival functions. Petroni and Wolfe (1994) and Zhang et al. (2001) used similar methods for the comparison of treatments based on interval-censored data. Cook et al. (1996) presented similar tests for treatment comparisons based on recurrent event data.

When we rewrite the test statistic $U_n^{(l)}$ as

$$U_n^{(l)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{k_i} W_n^{(l)}(t_{i,j}) \{\hat{\Lambda}_n(t_{i,j}) - \hat{\Lambda}_{n_l}(t_{i,j})\},$$

we observe that $U_n^{(l)}$ is also a Wilcoxon-type statistic. Similar statistics are often used in the analysis of repeated measurement data; see, for example, Davis and Wei (1988).

Based on Zhang (2006), we can consider a class of test statistics as follows

$$V_n^{(l)} = \sqrt{n} \int_0^{\tau} W_n^{(l)}(t) \{ \hat{\Lambda}_{n_1}(t) - \hat{\Lambda}_{n_l}(t) \} \, d\, G_n(t) \ , \ l = 2, \ldots, k.$$

Zhang (2006) obtained the asymptotic distribution of $V_n^{(l)}$ when $W_n^{(l)} = W_n$ and W_n satisfies a stronger condition (the composition of the inverse function of Λ_0 and the limit process $W(t)$ of $W_n(t)$ and $W_n(t)$ themselves are monotone). In addition, the proof of Theorem 1 of Zhang (2006) may not be rigorous. These will be discussed with more details in next section.

For the selection of the weight process $W_n^{(l)}(t)$, a simple and natural choice is $W_n^{(1,l)}(t) = 1$, $l = 1, \ldots, k$. Another natural choice is $W_n^{(2,l)}(t) = Y_n(t) = \sum_{i=1}^n I(t \leq t_{i,k_i}) / n$, $l = 1, \ldots, k$, in which case weights are proportional to the number of subjects under observation. Yet another choice for the weight process $W_n^{(l)}(t)$ is

$$W_n^{(3,l)}(t) = g(Y_{n_l}(t), Y_n(t)),$$

where g is a fixed function, and $Y_{n_l}(t)$ $(l = 1, \ldots, k)$ are defined as $Y_n(t)$ with the summation being only over subjects in the lth group. Some weight processes similar to $W_n^{(3)}$ have been used when recurrent event data are observed; see Andersen et al. (1993).

In the next section, we will present the asymptotic distributions of

$$\mathbf{U}_n = (U_n^{(1)}, \ldots, U_n^{(k)})^T$$

and

$$\mathbf{V}_n = (V_n^{(2)}, \ldots, V_n^{(k)})^T$$

in order to construct the tests for the null hypothesis.

2.3 Asymptotic Results

Let $\Lambda_0(t)$ denote the true mean function of the $N_i(t)$'s under H_0. Suppose that K is an integer-value random variable and $T = \{T_{k,j},\ j = 1, \ldots, k,\ k = 1, 2, \ldots\}$ is a random triangular array, and that k_i and $t_{i,j} = t_{k_i,j}$'s are realizations of them. We assume that $\{(K_i; T_{K_i,1}, \ldots, T_{K_i,K_i})\,;\ i = 1, \ldots, n\}$ are independent and identically distributed, and are independent of the N_i's. Let $\mathbf{X} = (K, T_K, N_K)$, where T_k is the kth row of the triangular array T and $N_k = (N(T_{k,1}), \ldots, N(T_{k,k}))$. Then, $\mathbf{X}_i = (K_i, T_{K_i}, N_{i,K_i})$, $i = 1, \ldots, n$, is a random sample of size n from the distribution of $\mathbf{X}$. For establishing asymptotic results on $\hat{\Lambda}_n(t)$ and $\mathbf{U}_n$, we need the following regularity conditions:

A. The mean function Λ_0 is strictly increasing such that $\Lambda_0(\tau) \leq M$ for some constant $M \in (0, \infty)$;

B. There exists a constant K_0 such that $\Pr\{K \leq K_0\} = 1$ and that the random variables $T_{k,j}$'s take values in a bounded set $[\tau_0, \tau]$ and $\Pr\{T_{K,1} = \tau_0\} > 0$, where $0 < \tau_0 < \tau < \infty$;

C. $\Pr\{\limsup_{n\to\infty} \max_i N_i(\tau) < \infty\} = 1$ and $E\left((N_i(\tau))^4\right) \leq M_1$, where M_1 is a constant.

Now, let $\hat{\Lambda}_n(t)$ be the isotonic regression estimate of $\Lambda_0(t)$ under H_0 given in Section 1. Also let Λ_0^{-1} denote the inverse function of Λ_0, and let $W \circ \Lambda_0^{-1}$ denote composition of two functions W and Λ_0^{-1}. First, we present the asymptotic normality of functional of $\hat{\Lambda}_n$.

Theorem 2.3.1 *Suppose that Conditions A, B and C hold. Further, suppose that $W(t)$ is a bounded weight process such that $W \circ \Lambda_0^{-1}$ is a bounded Lipschitz function. Let $G(t) = E\left[\sum_{j=1}^{K} I(T_{K,j} \leq t)\right]$. Then as $n \to \infty$,*

$$\sqrt{n}\int_0^{\tau} W(t)\left\{\hat{\Lambda}_n(t) - \Lambda_0(t)\right\} dG(t) \longrightarrow U_w$$

in distribution, where U_w has a normal distribution with mean zero and variance

$$\sigma_w^2 = E\left[\sum_{j=1}^{K} W(T_{K,j})\left\{N(T_{K,j}) - \Lambda_0(T_{K,j})\right\}\right]^2 \tag{2.1}$$

that can be consistently estimated by

$$\hat{\sigma}_w^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\sum_{j=1}^{K_i} W(T_{K_i,j})\left\{N_i(T_{K_i,j}) - \hat{\Lambda}_n(T_{K_i,j})\right\}\right]^2.$$

Proof. First, note that

$$\sqrt{n}\int_0^{\tau} W(t)\{\hat{\Lambda}_n(t) - \Lambda_0(t)\}\, dG(t) = I_{1n} + I_{2n} + I_{3n},$$

where

$$I_{1n} = \sqrt{n}(P_n - P)\left[\sum_{j=1}^{K} W(T_{K,j})\{\Lambda_0(T_{K,j}) - \hat{\Lambda}_n(T_{K,j})\}\right],$$

$$I_{2n} = \sqrt{n}P_n\left[\sum_{j=1}^{K} W(T_{K,j})\{\hat{\Lambda}_n(T_{K,j}) - N(T_{K,j})\}\right],$$

and

$$I_{3n} = \sqrt{n}P_n\left[\sum_{j=1}^{K} W(T_{K,j})\{N(T_{K,j}) - \Lambda_0(T_{K,j})\}\right],$$

where P_n is the empirical measure corresponding to (N, T, K), P is the corresponding underlying true measure, $P_n f = \frac{1}{n}\sum_{i=1}^{n} f_i$ and $Pf = \int f dP$. It is easy to see that I_{3n} is a U-statistic and has an asymptotic normal distribution with mean zero and variance

that can be consistently estimated by $\hat{\sigma}_w^2$ in the theorem. Hence, it is sufficient to show that both I_{1n} and I_{2n} converge in probability to zero.

We will show the convergence of I_{1n} first. Note that Condition C implies

$$\limsup_{n\to\infty} \hat{\Lambda}_n(\tau) < \infty$$

almost surely. Note that $\hat{\Lambda}_n(\tau_0) \to \Lambda_0(\tau_0)$ a.s. So, for every $0 < \varepsilon < \Lambda_0(\tau_0)$, there exist two positive constants $M_\varepsilon > \Lambda_0(\tau)$ and $L_\varepsilon < \Lambda_0(\tau_0)$ such that

$$\sup_n \Pr\{\hat{\Lambda}_n(\tau_0) < L_\varepsilon\} + \sup_n \Pr\{\hat{\Lambda}_n(\tau) > M_\varepsilon\} < \varepsilon.$$

Let

$$\mathcal{F} = \{\Lambda : [0,\tau] \longrightarrow [0,\infty) \,|\Lambda \text{ is nondecreasing, } \Lambda(0) = 0\}$$

and

$$\mathcal{F}_\varepsilon = \{\Lambda : \Lambda \in \mathcal{F},\ \Lambda(\tau_0) \geq L_\varepsilon,\ \Lambda(\tau) \leq M_\varepsilon\}.$$

Define $\hat{\Lambda}_{n,\varepsilon}$ as

$$\hat{\Lambda}_{n,\varepsilon} = \text{argmax}_{\Lambda\in\Omega\cap\mathcal{F}_\varepsilon} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{K_i} \{N_i(T_{K_i,j}) \log \Lambda(T_{K_i,j}) - \Lambda(T_{K_i,j})\} \right\},$$

where Ω is the class of nondecreasing step functions with possible jumps only at the observation time points $\{T_{K_i,j},\ j = 1,\ldots,K_i,\ i = 1,\ldots,n\}$. Let $I_{1n,\varepsilon}$ denote the version of I_{1n} obtained by replacing $\hat{\Lambda}_n$ with $\hat{\Lambda}_{n,\varepsilon}$. Then, to prove that I_{1n} converges to zero in probability, it is sufficient to show that $I_{1n,\varepsilon} = o_p(1)$ since

$$\Pr\{I_{1n,\varepsilon} \neq I_{1n}\} \leq \Pr\{\hat{\Lambda}_n(\tau_0) < L_\varepsilon\} + \Pr\{\hat{\Lambda}_n(\tau) > M_\varepsilon\} < \varepsilon.$$

By using arguments similar to those in Sun and Fang (2003), it can be shown that $I_{1n,\varepsilon} = o_p(1)$.

Next, we show the convergence of I_{2n}. By using the same block argument as in Proposition 1.2 in Part II of Groeneboom and Wellner (1993), we have for any real function h,

$$\sum_{\ell=1}^{m} h(\hat{\Lambda}_n(s_\ell))w_\ell\{\bar{N}_\ell - \hat{\Lambda}_n(s_\ell)\} = 0,$$

where the s_ℓ's, w_ℓ's and $\bar{N}_\ell = \bar{n}_\ell$ are as defined in Section 2.2. Hence, we can rewrite I_{2n} as

$$I_{2n} = \sqrt{n}P_n\left[\sum_{j=1}^{K}\{W_0(\Lambda_0(T_{K,j})) - W_0(\hat{\Lambda}_n(T_{K,j}))\}\{\hat{\Lambda}_n(T_{K,j}) - N(T_{K,j})\}\right],$$

where $W_0 = W \circ \Lambda_0^{-1}$. By Condition C, there exists a constant N_ε such that

$$\sup_n \Pr\{\max_{1\le i\le n} N_i(\tau) > N_\varepsilon\} < \varepsilon.$$

Let

$$A_n = \{\max_{1\le i\le n} N_i(\tau) \le N_\varepsilon\},$$

and for $\Lambda \in \mathcal{F}_\varepsilon$, let

$$f_\Lambda(\mathbf{X}) = \sum_{j=1}^{K}\{W_0(\Lambda_0(T_{K,j})) - W_0(\Lambda(T_{K,j}))\}\{\Lambda(T_{K,j}) - N(T_{K,j})\},$$

$$g_\Lambda(\mathbf{X}) = \sum_{j=1}^{K}\{W_0(\Lambda_0(T_{K,j})) - W_0(\Lambda(T_{K,j}))\}\{\Lambda(T_{K,j}) - \Lambda_0(T_{K,j})\},$$

and

$$h_\Lambda(\mathbf{X}) = \sum_{j=1}^{K}\{W_0(\Lambda_0(T_{K,j})) - W_0(\Lambda(T_{K,j}))\}\{\Lambda_0(T_{K,j}) - N(T_{K,j})\}.$$

Then, we have

$$I_{2n} = (\Delta_{1n} + \Delta_{2n} + \Delta_{3n})1_{A_n} + \Delta_{4n},$$

where

$$\Delta_{1n} = \sqrt{n}(P_n - P)\left\{f_{\hat{\Lambda}_n}(\mathbf{X})I(N(\tau) \le N_\varepsilon)\right\},$$

$$\Delta_{2n} = \sqrt{n}P\left\{g_{\hat{\Lambda}_n}(\mathbf{X})I(N(\tau) \leq N_\varepsilon)\right\},$$

$$\begin{aligned}
\Delta_{3n} &= \sqrt{n}P\left\{h_{\hat{\Lambda}_n}(\mathbf{X})I(N(\tau) \leq N_\varepsilon)\right\} \\
&= \sqrt{n}P\left\{h_{\hat{\Lambda}_n}(\mathbf{X})\right\} - \sqrt{n}P\left\{h_{\hat{\Lambda}_n}(\mathbf{X})I(N(\tau) > N_\varepsilon)\right\} \\
&= -\sqrt{n}P\left\{h_{\hat{\Lambda}_n}(\mathbf{X})I(N(\tau) > N_\varepsilon)\right\},
\end{aligned}$$

and

$$\Delta_{4n} = \sqrt{n}P_n\left\{f_{\hat{\Lambda}_n}(\mathbf{X})\right\}1_{A_n^c}.$$

For Δ_{3n} and Δ_{4n}, we have $\forall\delta > 0$,

$$P\left\{|\Delta_{3n}| > \delta\right\} \leq P\left\{N(\tau) > N_\varepsilon\right\} < \varepsilon$$

and

$$P\left\{|\Delta_{4n}| > \delta\right\} \leq P\left(A_n^c\right) < \varepsilon\,.$$

Let $\Delta_{1n,\varepsilon}$ denote the version of Δ_{1n} obtained by replacing $\hat{\Lambda}_n$ by $\hat{\Lambda}_{n,\varepsilon}$. Since W_0 is a bounded Lipschitz function, it can be shown that

$$\mathcal{H}_\varepsilon = \left\{f_\Lambda(\mathbf{X})1_{\{N(\tau)\leq N_\varepsilon\}} : \Lambda \in \mathcal{F}_\varepsilon\right\}$$

is P-Donsker using the bracket entropy theorem of Van der Vaart and Wellner (1996, pp. 127-159) and arguments similar to those in Huang and Wellner (1995). Moreover, Theorem 4.1 of Wellner and Zhang (2000) yields

$$d(\hat{\Lambda}_{n,\varepsilon}, \Lambda_0) \leq d(\hat{\Lambda}_n, \Lambda_0) \longrightarrow 0,$$

where

$$d(\Lambda_1, \Lambda_2) = \left\{\int_0^\tau |\Lambda_1(t) - \Lambda_2(t)|^2\, dG(t)\right\}^{1/2}.$$

Hence, it follows from the uniform asymptotic equicontinuity of the empirical process (Van der Vaart and Wellner, 1996, pp. 168-171) that $\Delta_{1n,\varepsilon} = o_p(1)$. Then, we have $\Delta_{1n} = o_p(1)$ since

$$P\{\Delta_{1n} \neq \Delta_{1n,\varepsilon}\} \leq P\{\hat{\Lambda}_n(\tau) > M_\varepsilon\} < \varepsilon.$$

For Δ_{2n}, since W_0 is a bounded Lipschitz function, it follows that

$$\begin{aligned} |\Delta_{2n}| &= \left|\sqrt{n}\int_0^\tau \{W_0(\Lambda_0(t) - W_0(\hat{\Lambda}_n(t))\}\{\hat{\Lambda}_n(t) - \Lambda_0(t)\}\, dG(t)\right| \\ &\leq c_1\sqrt{n}\, d^2(\hat{\Lambda}_n, \Lambda_0), \end{aligned}$$

where c_1 is a constant. To prove that $\sqrt{n}\, d^2(\hat{\Lambda}_n, \Lambda_0) = o_p(1)$, we only need to show that $\sqrt{n}\, d^2(\hat{\Lambda}_{n,\varepsilon}, \Lambda_0) = o_p(1)$. We shall now show that $d(\hat{\Lambda}_{n,\varepsilon}, \Lambda_0) = O_p(n^{-\frac{1}{3}})$.

To establish the rate of convergence for $\hat{\Lambda}_{n,\varepsilon}$, we shall apply Theorem 3.2.5 of Van der Vaart and Wellner (1996). Define

$$m_\Lambda(X) = \sum_{j=1}^{K}\{N(T_{K,j})\log\Lambda(T_{K,j}) - \Lambda(T_{K,j})\}$$

and $\mathbb{M}(\Lambda) = Pm_\Lambda(X)$. Let $\varphi(x) = x(\log x - 1) + 1$. Then $\varphi(x) \geq \frac{1}{5}(x-1)^2$ for x in a neighbourhood of $x = 1$. Thus, in a neighbourhood of Λ_0,

$$\begin{aligned} &\mathbb{M}(\Lambda_0) - \mathbb{M}(\Lambda) \\ &= P\left[\sum_{j=1}^{K}\{\Lambda_0(T_{K,j})\log\Lambda_0(T_{K,j}) - \Lambda_0(T_{K,j})\}\right] \\ &\quad -P\left[\sum_{j=1}^{K}\{\Lambda_0(T_{K,j})\log\Lambda(T_{K,j}) - \Lambda(T_{K,j})\}\right] \\ &= P\left[\sum_{j=1}^{K}\Lambda(T_{K,j})\varphi\left(\frac{\Lambda_0(T_{K,j})}{\Lambda(T_{K,j})}\right)\right] \\ &= \int \Lambda(t)\varphi\left(\frac{\Lambda_0(t)}{\Lambda(t)}\right) dG(t) \end{aligned}$$

$$
\begin{aligned}
&\geq \frac{1}{5}\int \frac{(\Lambda_0(t)-\Lambda(t))^2}{\Lambda(t)}\,dG(t) \\
&\geq \frac{1}{5M_\varepsilon}\,d^2(\Lambda,\Lambda_0),
\end{aligned}
$$

and hence the separation condition of the theorem is satisfied. Also, let

$$\mathcal{F}_{\delta,\varepsilon} = \{\Lambda : d(\Lambda,\Lambda_0) \leq \delta, \Lambda \in \mathcal{F}_\varepsilon\}\ (\delta > 0)$$

and

$$\mathcal{M}_{\delta,\varepsilon} = \{m_\Lambda(X) - m_{\Lambda_0}(X) : \Lambda \in \mathcal{F}_{\delta,\varepsilon}\}.$$

For $\Lambda \in \mathcal{F}_{\delta,\varepsilon}$, it is easily shown that $P|m_\Lambda(X) - m_{\Lambda_0}(X)|^2 \leq c_2\delta^2$ and $||m_\Lambda(X) - m_{\Lambda_0}(X)||_\infty \leq c_3$ for some constatnts c_2 and c_3. Since we have

$$\log N_{[]}\left(\eta, \mathcal{M}_{\delta,\varepsilon}, L_2(P)\right) \leq c_4\eta^{-1},$$

where c_4 is a constant which depends only on M_ε, then

$$
\begin{aligned}
&\int_0^\delta \sqrt{1+\log N_{[]}\left(\eta, \mathcal{M}_{\delta,\varepsilon}, L_2(P)\right)}\,d\eta \\
&\leq \int_0^\delta \sqrt{1+c_4\eta^{-1}}\,d\eta \\
&\leq \sqrt{\delta+c_4}\int_0^\delta \eta^{-\frac{1}{2}}\,d\eta \\
&\leq c_5\delta^{\frac{1}{2}}.
\end{aligned}
$$

for some constant c_5. Hence, by applying Lemma 3.4.2 of Van der Vaart and Wellner (1996), we have

$$E^*||\sqrt{n}(P_n - P)||_{\mathcal{M}_{\delta,\varepsilon}} \leq c_6\phi_n(\delta)$$

for some constant c_6, where E^* denotes the outer expectation, and $\phi_n(\delta) = \delta^{\frac{1}{2}} + \delta^{-1}n^{-\frac{1}{2}}$. Now, upon using Theorem 3.2.5 of Van der Vaart and Wellner (1996), $d(\hat{\Lambda}_{n,\varepsilon}, \Lambda_0)$

converges in probability to zero of order at least $n^{-\frac{1}{3}}$. This shows that $\Delta_{2n} = o_p(1)$ which completes the proof of the theorem. $\square$

Remark. For the proof of Theorem 1 of Zhang (2006), the author claimed that $||m_i^r(X) - m_i^l(X)||_{P,B}^2 \leq C\varepsilon^2$ where $m_i^r(X)$, and $m_i^l(X)$ are as defined in Zhang (2006, pp. 786). From

$$m_i^r(X) - m_i^l(X) = \sum_{j=1}^{K} \left\{ N(T_{K,j}) \log \frac{\Lambda_i^r(T_{K,j})}{\Lambda_i^l(T_{K,j})} - \Lambda_i^r(T_{K,j}) + \Lambda_i^l(T_{K,j}) \right\},$$

one can see that a positive lower bound for $\Lambda_i^l(t)$ is required to derive $||m_i^r(X) - m_i^l(X)||_{P,B}^2 \leq C\varepsilon^2$. However, this property may not be obtained by the setting of $\Lambda_i^l(X)$ given in Zhang (2006, pp. 786). It is for this reason we have considered a smaller class $\mathcal{F}_\varepsilon$ and $\hat{\Lambda}_{n,\varepsilon}$ instead of $\hat{\Lambda}_n$.

Now, we derive the asymptotic distributions of $\mathbf{U}_n$ and $\mathbf{V}_n$. Let S_l denote the set of indices for subjects in group l, $l = 1, \ldots, k$.

Theorem 2.3.2 *Suppose that Conditions A, B and C hold. Further, suppose that $W_n^{(l)}(t)$'s are bounded weight processes and that there exists a bounded function $W(t)$ such that $W \circ \Lambda_0^{-1}$ is a bounded Lipschitz function, and*

$$\left[\int_0^\tau \{W_n^{(l)}(t) - W(t)\}^2 \, dG(t) \right]^{1/2} = o_p(n^{-1/6}) \ , \ l = 1, \ldots, k. \tag{2.2}$$

Also suppose that $n_l/n \to p_l$ as $n \to \infty$, where $0 < p_l < 1$, $l = 1, \ldots, k$, and $p_1 + \cdots + p_k = 1$. Then under $H_0 : \Lambda_1 = \cdots = \Lambda_k = \Lambda_0$,

(i) $\mathbf{U}_n$ *has an asymptotic normal distribution with mean vector* $\mathbf{0}$ *and covariance matrix*

$$\mathbf{\Sigma}_{\mathbf{U}_w} = \mathbf{\Gamma}\, \mathbf{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2)\, \mathbf{\Gamma}', \tag{2.3}$$

where

$$\mathbf{\Gamma} = \begin{pmatrix} \sqrt{p_1} - \sqrt{\frac{1}{p_1}} & \sqrt{p_2} & \cdots & \sqrt{p_k} \\ \sqrt{p_1} & \sqrt{p_2} - \sqrt{\frac{1}{p_2}} & \cdots & \sqrt{p_k} \\ \cdots & \cdots & \cdots & \cdots \\ \sqrt{p_1} & \sqrt{p_2} & \cdots & \sqrt{p_k} - \sqrt{\frac{1}{p_k}} \end{pmatrix} \tag{2.4}$$

and $\sigma_1^2 = \cdots = \sigma_k^2 = \sigma_w^2$ *given in (2.1).*

(ii) $\mathbf{V}_n$ *has an asymptotic normal distribution with mean vector* $\mathbf{0}$ *and covariance matrix*

$$\mathbf{\Sigma}_{\mathbf{V}_w} = \mathbf{H}\,\mathbf{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2)\,\mathbf{H}', \tag{2.5}$$

where

$$\mathbf{H} = \begin{pmatrix} -\sqrt{\frac{1}{p_1}} & \sqrt{\frac{1}{p_2}} & 0 & \cdots & 0 \\ -\sqrt{\frac{1}{p_1}} & 0 & \sqrt{\frac{1}{p_3}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\sqrt{\frac{1}{p_1}} & 0 & 0 & \cdots & \sqrt{\frac{1}{p_k}} \end{pmatrix} \tag{2.6}$$

and σ_l^2 *is as given in (i).*

(iii) In addition, if

$$\max_{1 \le i \le n} E\left[\sum_{j=1}^{K_i} \left\{W_n^{(l)}(T_{K_i,j}) - W(T_{K_i,j})\right\}^2\right] \longrightarrow 0 \tag{2.7}$$

for $l = 1, \ldots, k$, *then* $\mathbf{\Sigma}_{\mathbf{U}_w}$ *and* $\mathbf{\Sigma}_{\mathbf{V}_w}$ *can be consistently estimated by*

$$\hat{\mathbf{\Sigma}}_{U_n} = \mathbf{\Gamma}_n\,\mathbf{diag}(\hat{\sigma}_1^2, \hat{\sigma}_2^2, \ldots, \hat{\sigma}_k^2)\,\mathbf{\Gamma}_n', \tag{2.8}$$

and

$$\hat{\mathbf{\Sigma}}_{V_n} = \mathbf{H}_n\,\mathbf{diag}(\hat{\sigma}_1^2, \hat{\sigma}_2^2, \cdots, \hat{\sigma}_k^2)\,\mathbf{H}_n', \tag{2.9}$$

where

$$\mathbf{\Gamma}_n = \begin{pmatrix} \sqrt{\frac{n_1}{n}} - \sqrt{\frac{n}{n_1}} & \sqrt{\frac{n_2}{n}} & \cdots & \sqrt{\frac{n_k}{n}} \\ \sqrt{\frac{n_1}{n}} & \sqrt{\frac{n_2}{n}} - \sqrt{\frac{n}{n_2}} & \cdots & \sqrt{\frac{n_k}{n}} \\ \cdots & \cdots & \cdots & \cdots \\ \sqrt{\frac{n_1}{n}} & \sqrt{\frac{n_2}{n}} & \cdots & \sqrt{\frac{n_k}{n}} - \sqrt{\frac{n}{n_k}} \end{pmatrix}, \tag{2.10}$$

$$\mathbf{H}_n = \begin{pmatrix} -\sqrt{\frac{n}{n_1}} & \sqrt{\frac{n}{n_2}} & 0 & \cdots & 0 \\ -\sqrt{\frac{n}{n_1}} & 0 & \sqrt{\frac{n}{n_3}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\sqrt{\frac{n}{n_1}} & 0 & 0 & \cdots & \sqrt{\frac{n}{n_k}} \end{pmatrix} \tag{2.11}$$

and

$$\hat{\sigma}_l^2 = \frac{1}{n} \sum_{i=1}^{n} \left[\sum_{j=1}^{K_i} W_n^{(l)}(T_{K_i,j}) \left\{ N_i(T_{K_i,j}) - \hat{\Lambda}_n(T_{K_i,j}) \right\} \right]^2 \tag{2.12}$$

for $l = 1, \ldots, k$.

Proof. (i) Let

$$G_{n_l}(t) = \frac{1}{n_l} \sum_{i \epsilon S_l} \sum_{j=1}^{K_i} I(T_{K_i,j} \leq t)$$

for $l = 1, \ldots, k$. To obtain the asymptotic distribution of $\mathbf{U}_n$, we first note that $U_n^{(l)}$ can rewritten as

$$U_n^{(l)} = U_{1n}^{(l)} - \sqrt{\frac{n}{n_l}} U_{2n}^{(l)},$$

where, for $l = 1, \ldots, k$,

$$U_{1n}^{(l)} = \sqrt{n} \int_0^{\tau} W_n^{(l)}(t) \{ \hat{\Lambda}_n(t) - \Lambda_0(t) \} dG_n(t)$$

and

$$U_{2n}^{(l)} = \sqrt{n_l} \int_0^\tau W_n^{(l)}(t)\{\hat{\Lambda}_{n_l}(t) - \Lambda_0(t)\} dG_n(t).$$

Further

$$U_{1n}^{(l)} = I_{1n}^{(l)} + I_{2n}^{(l)} + I_{3n}^{(l)},$$

where

$$\begin{aligned} I_{1n}^{(l)} &= \sqrt{n} \int_0^\tau \{W_n^{(l)}(t) - W(t)\}\{\hat{\Lambda}_n(t) - \Lambda_0(t)\}\, dG_n(t) \\ &= \sqrt{n} \int_0^\tau \{W_n^{(l)}(t) - W(t)\}\{\hat{\Lambda}_n(t) - \Lambda_0(t)\}\, dG(t) + o_p(1), \end{aligned}$$

$$I_{2n}^{(l)} = \sqrt{n}(P_n - P)\left[\sum_{j=1}^{K} W(T_{K,j})\{\hat{\Lambda}_n(T_{K,j}) - \Lambda_0(T_{K,j})\}\right],$$

and

$$I_{3n}^{(l)} = \sqrt{n} \int_0^\tau W(t)\{\hat{\Lambda}_n(t) - \Lambda_0(t)\}\, dG(t).$$

First, we show that $I_{1n}^{(l)} = o_p(1)$, $l = 1, \ldots, k$. Using Cauchy-Schwarz inequality and the proof of Theorem 2.3.1, we have

$$\begin{aligned} &\left|\sqrt{n} \int_0^\tau \{W_n^{(l)}(t) - W(t)\}\{\hat{\Lambda}_n(t) - \Lambda_0(t)\}\, dG(t)\right| 1_{\{\hat{\Lambda}_n(\tau) \le M_\varepsilon\}} \\ \le\ & \sqrt{n}\left\{\int_0^\tau (W_n^{(l)}(t) - W(t))^2\, dG(t)\right\}^{1/2} \left\{\int_0^\tau (\hat{\Lambda}_{n,\varepsilon}(t) - \Lambda_0(t))^2\, dG(t)\right\}^{1/2} \\ \longrightarrow\ & 0 \end{aligned}$$

in probability, since

$$\left\{\int_0^\tau (\hat{\Lambda}_{n,\varepsilon}(t) - \Lambda_0(t))^2\, dG(t)\right\}^{1/2} = O_p(n^{-1/3}).$$

Hence, $I_{1n}^{(l)} = o_p(1)$, $l = 1, \ldots, k$. Now, as in the proof of Theorem 2.3.1, it can be shown that $I_{2n}^{(l)} = o_p(1)$, $l = 1, \ldots, k$. Also, it follows from Theorem 2.3.1 that

$$I_{3n}^{(l)} = \sqrt{n} \int_0^\tau W(t)\{N(t) - \Lambda_0(t)\}\, dG_n(t) + o_p(1)$$

for $l = 1, \ldots, k$. Hence, we have

$$U_{1n}^{(l)} = \sqrt{n} \int_0^\tau W(t)\{N(t) - \Lambda_0(t)\}\, dG_n(t) + o_p(1)\ ,\ l = 1, \ldots, k.$$

Similarly, we can show that

$$U_{2n}^{(l)} = \sqrt{n_l} \int_0^\tau W(t)\{N(t) - \Lambda_0(t)\}\, dG_{n_l}(t) + o_p(1)\ ,\ l = 1, \ldots, k.$$

Let

$$Y_n = \sqrt{n} \int_0^\tau W(t)\{N(t) - \Lambda_0(t)\}\, dG_n(t)$$

and

$$Y_n^{(l)} = \sqrt{n_l} \int_0^\tau W(t)\{N(t) - \Lambda_0(t)\}\, dG_{n_l}(t)\ , \quad \text{for } l = 1, \ldots, k.$$

Evidently, $Y_n^{(l)}$'s are i.i.d., and $\sqrt{n}Y_n = \sum_{l=1}^k \sqrt{n_l} Y_n^{(l)}$. Then,

$$\begin{aligned} U_n^{(l)} &= Y_n - \sqrt{\frac{n}{n_l}} Y_n^{(l)} + o_p(1) \\ &= \sum_{i=1}^k \sqrt{\frac{n_i}{n}} Y_n^{(i)} - \sqrt{\frac{n}{n_l}} Y_n^{(l)} + o_p(1),\ l = 1, \ldots, k, \end{aligned}$$

and so

$$\mathbf{U}_n = \Gamma_n \mathbf{Y}_n + o_p(1) = \Gamma \mathbf{Y}_n + o_p(1),$$

where Γ_n and Γ are as given in the theorem, and

$$\mathbf{Y}_n = \left(Y_n^{(1)}, \ldots, Y_n^{(k)}\right)^T$$

converges in distribution to $\mathbf{Y}_w$ having a k-dimensional normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{diag}(\sigma_1^2, \ldots, \sigma_k^2)$. Thus, we have $\mathbf{U}_n$ converging in distribution to a random variable $\mathbf{U}_w$ that has a normal distribution $N(\mathbf{0}, \mathbf{\Sigma}_w)$, in which $\mathbf{\Sigma}_w$ is presented in part (i) in the theorem.

(ii) We note that $V_n^{(l)} = U_n^{(1,l)} - U_n^{(l)}$, $l = 2, \ldots, k$, where $U_n^{(1,l)}$ is defined as $U_n^{(1)}$ by replacing $W_n^{(1)}$ with $W_n^{(l)}$ for $l = 2, \ldots, k$. Then, (ii) follows from (i).

(iii) To show that $\hat{\sigma}_l^2 - \sigma_w^2 = o_p(1)$ for $l = 1, \ldots, k$, we set

$$\phi(\xi, \Lambda, X) = \sum_{j=1}^{K} \xi(T_{K,j}) \{N(T_{K,j}) - \Lambda(T_{K,j})\}$$

Then $\sigma_w^2 = P\phi^2(W, \Lambda_0, X)$ and $\hat{\sigma}_l^2 = P_n\phi^2(W_n^{(l)}, \hat{\Lambda}_n, X)$. Note that

$$\begin{aligned}
\hat{\sigma}_l^2 - \sigma_w^2 &= P_n\left\{\phi^2(W_n^{(l)}, \hat{\Lambda}_n, X) - \phi^2(W_n^{(l)}, \Lambda_0, X)\right\} \\
&\quad + P_n\left\{\phi^2(W_n^{(l)}, \Lambda_0, X) - \phi^2(W, \Lambda_0, X)\right\} \\
&\quad + (P_n - P)\phi^2(W, \Lambda_0, X).
\end{aligned}$$

It can be easily shown that

$$P_n\left\{\phi^2(W_n^{(l)}, \hat{\Lambda}_n, X) - \phi^2(W_n^{(l)}, \Lambda_0, X)\right\} = o_p(1)$$

and

$$(P_n - P)\phi^2(W_0, \Lambda_0, X) = o_p(1).$$

Since it follows from Conditions A and B that

$$\begin{aligned}
&\left|\phi(W_n^{(l)}, \Lambda_0, X) - \phi(W, \Lambda_0, X)\right| = \left|\phi(W_n^{(l)} - W, \Lambda_0, X)\right| \\
\leq\ & c_1\{N(\tau) + M\} \sum_{j=1}^{K} \left|W_n^{(l)}(T_{K,j}) - W(T_{K,j})\right|
\end{aligned}$$

with probability 1 for some constant c_1 and

$$\begin{aligned}
&\left|\phi(W_n^{(l)}, \Lambda_0, X) + \phi(W, \Lambda_0, X)\right| = \left|\phi(W_n^{(l)} + W, \Lambda_0, X)\right| \\
\leq\ & c_2\{N(\tau) + M\}\}
\end{aligned}$$

with probability 1 for some constant c_2, then we have from the Cauchy-Schwarz inequality, Condition C, and (2.7)

$$
\begin{aligned}
& E\left|\phi^2(W_n^{(l)}, \Lambda_0, X_i) - \phi^2(W, \Lambda_0, X_i)\right| \\
& \leq c_3 E\left[\{N_i(\tau) + M\}^2 \left\{\sum_{j=1}^{K_i} \left|W_n^{(l)}(T_{K_i,j}) - W(T_{K_i,j})\right|\right\}\right] \\
& \leq c_3 \left[E\{N_i(\tau) + M\}^4\right]^{1/2} \left[E\left\{\sum_{j=1}^{K_i} \left|W_n^{(l)}(T_{K_i,j}) - W(T_{K_i,j})\right|\right\}^2\right]^{1/2} \\
& \leq c_4 \max_{1\leq i\leq n} \left[E\left\{\sum_{j=1}^{K_i} \left|W_n^{(l)}(T_{K_i,j}) - W(T_{K_i,j})\right|^2\right\}\right]^{1/2} \\
& \longrightarrow 0
\end{aligned}
$$

where c_3 and c_4 are finite positive constants , which completes the proof of part (iii).

Hence, the proof of the theorem is completed. □

Clearly,

$$\left[\int_0^\tau \{W_n^{(l)}(t) - W(t)\}^2 \, dG(t)\right]^{1/2} = d(W_n^{(l)}, W)$$

Here, we need

$$n^{1/6} d(W_n^{(l)}, W) \rightarrow_p 0$$

and (2.7) for weight processes $W_n^{(l)}$, $l = 1, \ldots, k$. For example, $W_n^{(1,l)}$ and $W_n^{(2,l)}$ given in Section 2.2, $1 - W_n^{(2,l)}$, Y_{n_l} and $1 - Y_{n_l}$ all satisfy these conditions. The weight processes $Y_{n_1} Y_{n_l}/Y_n$ and $(1 - Y_{n_1})(1 - Y_{n_l})/(1 - Y_n)$ also satisfy these conditions.

Now, we are ready to present the nonparametric k-sample tests for panel count data. Let $\mathbf{U}_0$ denote the first $(k-1)$ components of $\mathbf{U}_n$ and $\hat{\mathbf{\Sigma}}_0$ the matrix obtained by deleting the last row and column of $\hat{\mathbf{\Sigma}}_{U_n}$. Then, using Theorem 2.3.2, two tests of the hypothesis H_0 can be carried out by means of the statistics $T_1 = \mathbf{U}_0^T \hat{\mathbf{\Sigma}}_0^{-1} \mathbf{U}_0$ and

$T_2 = \mathbf{V}_n^T \hat{\Sigma}_{V_n}^{-1} \mathbf{V}_n$, which have asymptotically a central χ^2-distribution with $(k-1)$ degrees of freedom. This can be seen readily from the proof of the theorem.

2.4 Simulation Study

To examine the finite-sample properties of the proposed test statistic T_1 and compare its power with T_2, we carry out a simulation study for the three-sample comparison problem. To generate panel count data $\{k_i, t_{ij}, n_{ij}, j = 1, \ldots, k_i, i = 1, \ldots, n\}$, we mimic medical follow-up studies such as the example discussed in the next section. We first generate the number of observation times k_i from the uniform distribution $U\{1, \ldots, 10\}$, and then, given k_i, we generate observation times t_{ij}'s from $U\{1, \ldots, 10\}$, for simplicity. To generate n_{ij}'s, we assume that N_i's are nonhomogeneous Poisson or mixed Poisson processes. In particular, let $\{\nu_i, i = 1, \ldots, n\}$ be i.i.d. random variables, and given ν_i, let $N_i(t)$ be a Poisson process with mean function $\Lambda_i(t) = \nu_i t$ for $i \in S_1$, $\Lambda_i(t) = \nu_i t \exp(\beta_1)$ for $i \in S_2$ and $\Lambda_i(t) = \nu_i t \exp(\beta_2)$ for $i \in S_3$.

We consider two cases: $\nu_i = 1$ and $\nu_i \sim Gamma(2, 1/2)$. For each case, we consider two sample sizes, $n_1 = n_2 = n_3 = 50$ and 100, respectively. As mentioned earlier in Section 2.2, we choose the three weight processes: $W_n^{(1,l)}(t) = 1$, $l = 1, \ldots, k$, $W_n^{(2,l)}(t) = Y_n(t) = \sum_{i=1}^n I(t \le t_{i,k_i})/n$, $l = 1, \ldots, k$, and $W_n^{(3,l)}(t) = 1 - Y_n(t)$. Let

$$W_n^{(j)}(t) = \left(W_n^{(j,1)}(t), \ldots, W_n^{(j,k)}(t)\right), \quad j = 1, 2, 3.$$

All the results reported here are based on 1000 Monte Carlo replications.

Tables 2.1–2.4 present the estimated sizes and powers of the tests T_1 and T_2 at significance level $\alpha = 0.05$ for different values of β and the three weight processes based

on the simulated data for the two cases, respectively. In the first case, the $N_i(t)$'s are Poisson processes. In the second case, the $N_i(t)$'s are mixed Poisson processes. Tables 2.1 and 2.3 are for the situation with the total sample size of 150 and Tables 2.2 and 2.4 are for the situation with the total sample size of 300. For the situation considered here, the tests seem to have good powers, the powers of two tests are close for the three weight processes with the weight process $W_n^{(1)}$ showing a little higher power, and T_1 is slightly powerful than T_2. As expected, the power increases when the sample size increases, and the power decreases in the presence of variability. To evaluate the asymptotic result given in Theorem 2.3.2, the quantile plots of the test statistic T_1 against the chi-square distribution with 2 degrees of freedom are constructed. Figures 2.1 and 2.2 present the plots for the cases with $W_n(t) = W_n^{(1)}(t)$ and $n = 150$ and $n = 300$, respectively, and they clearly reveal that the asymptotic approximation is quite good. Similar plots were obtained for other situations as well.

In the above simulation study, we did examine all three weight processes suggested earlier in Section 2.2, and in all situations considered here, the weight process $W_n^{(1)}$ yielded slightly higher power than the other two weight processes. This may not always be true as one can see from the next section and simulation results presented by Zhang (2006). In general, one should select appropriate weight processes based on the behavior of the mean functions to improve power. Zhang (2006) provided a detailed discussion about the roles of these weight processes through Monte Carlo simulations. In addition to the three processes considered here, some other weight processes can be found in Andersen et al. (1993), which discusses nonparametric treatment comparison based on recurrent event data. It would, therefore, be of great interest to investigate the problem of the selection of a weight process based on data.

Table 2.1: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for Poisson processes ($n_1 = n_2 = n_3 = 50$)

	T_1			T_2		
(β_1, β_2)	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$
(-0.5,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(-0.3,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(-0.3,-0.3)	0.969	0.936	0.949	0.968	0.931	0.949
(-0.1,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(-0.1,-0.3)	0.903	0.853	0.870	0.898	0.851	0.867
(-0.1,-0.1)	0.224	0.195	0.203	0.218	0.187	0.193
(0.0,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.0,-0.3)	0.974	0.950	0.959	0.965	0.946	0.953
(0.0,-0.1)	0.253	0.214	0.229	0.239	0.205	0.217
(0.0,0.0)	0.053	0.049	0.048	0.052	0.050	0.047
(0.0,0.1)	0.267	0.234	0.247	0.266	0.233	0.244
(0.0,0.3)	0.993	0.981	0.987	0.993	0.980	0.987
(0.0,0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.1,0.1)	0.273	0.223	0.236	0.270	0.221	0.236
(0.1,0.3)	0.963	0.939	0.952	0.960	0.934	0.949
(0.1,0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.3,0.3)	0.995	0.985	0.989	0.995	0.984	0.986
(0.3,0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.5,0.5)	1.000	1.000	1.000	1.000	1.000	1.000

Table 2.2: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for Poisson processes ($n_1 = n_2 = n_3 = 100$)

	T_1			T_2		
(β_1, β_2)	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$
(-0.5,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(-0.3,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(-0.3,-0.3)	1.000	1.000	1.000	1.000	1.000	1.000
(-0.1,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(-0.1,-0.3)	0.997	0.993	0.996	0.996	0.993	0.995
(-0.1,-0.1)	0.466	0.373	0.394	0.463	0.373	0.390
(0.0,-0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.0,-0.3)	1.000	1.000	1.000	1.000	1.000	1.000
(0.0,-0.1)	0.438	0.388	0.403	0.434	0.382	0.403
(0.0,0.0)	0.056	0.049	0.054	0.055	0.056	0.055
(0.0,0.1)	0.491	0.432	0.449	0.488	0.429	0.447
(0.0,0.3)	1.000	1.000	1.000	1.000	1.000	1.000
(0.0,0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.1,0.1)	0.450	0.371	0.388	0.450	0.368	0.388
(0.1,0.3)	0.999	0.999	0.999	0.999	0.999	0.999
(0.1,0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.3,0.3)	1.000	1.000	1.000	1.000	1.000	1.000
(0.3,0.5)	1.000	1.000	1.000	1.000	1.000	1.000
(0.5,0.5)	1.000	1.000	1.000	1.000	1.000	1.000

Table 2.3: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for mixed Poisson processes ($n_1 = n_2 = n_3 = 50$)

	T_1			T_2		
(β_1, β_2)	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$
(-0.5,-0.5)	0.773	0.761	0.769	0.779	0.763	0.772
(-0.3,-0.5)	0.635	0.629	0.615	0.617	0.614	0.615
(-0.3,-0.3)	0.348	0.344	0.342	0.348	0.332	0.337
(-0.1,-0.5)	0.657	0.642	0.654	0.648	0.637	0.645
(-0.1,-0.3)	0.251	0.244	0.243	0.250	0.246	0.243
(-0.1,-0.1)	0.088	0.085	0.083	0.077	0.081	0.076
(0.0,-0.5)	0.715	0.706	0.722	0.713	0.709	0.708
(0.0,-0.3)	0.331	0.326	0.334	0.327	0.321	0.314
(0.0,-0.1)	0.074	0.075	0.077	0.068	0.070	0.069
(0.0,0.0)	0.060	0.059	0.057	0.056	0.060	0.057
(0.0,0.1)	0.082	0.086	0.074	0.078	0.075	0.078
(0.0,0.3)	0.381	0.386	0.357	0.379	0.369	0.372
(0.0,0.5)	0.815	0.811	0.805	0.814	0.813	0.791
(0.1,0.1)	0.086	0.086	0.073	0.085	0.086	0.068
(0.1,0.3)	0.317	0.312	0.300	0.300	0.303	0.285
(0.1,0.5)	0.741	0.736	0.723	0.731	0.727	0.729
(0.3,0.3)	0.323	0.325	0.296	0.317	0.324	0.308
(0.3,0.5)	0.654	0.655	0.630	0.647	0.650	0.629
(0.5,0.5)	0.769	0.767	0.745	0.765	0.764	0.744

Table 2.4: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for mixed Poisson processes ($n_1 = n_2 = n_3 = 100$)

	T_1			T_2		
(β_1, β_2)	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$	$W_n^{(1)}(t)$	$W_n^{(2)}(t)$	$W_n^{(3)}(t)$
(-0.5,-0.5)	0.976	0.971	0.970	0.976	0.973	0.972
(-0.3,-0.5)	0.902	0.901	0.894	0.901	0.899	0.892
(-0.3,-0.3)	0.609	0.598	0.617	0.603	0.595	0.615
(-0.1,-0.5)	0.933	0.929	0.932	0.931	0.926	0.930
(-0.1,-0.3)	0.501	0.487	0.494	0.493	0.481	0.493
(-0.1,-0.1)	0.119	0.123	0.118	0.119	0.121	0.116
(0.0,-0.5)	0.966	0.961	0.961	0.961	0.957	0.960
(0.0,-0.3)	0.593	0.579	0.577	0.583	0.578	0.576
(0.0,-0.1)	0.105	0.106	0.099	0.097	0.097	0.096
(0.0,0.0)	0.053	0.056	0.054	0.054	0.057	0.052
(0.0,0.1)	0.109	0.108	0.109	0.103	0.105	0.104
(0.0,0.3)	0.676	0.668	0.655	0.668	0.659	0.656
(0.0,0.5)	0.989	0.987	0.982	0.988	0.985	0.982
(0.1,0.1)	0.117	0.116	0.097	0.117	0.116	0.098
(0.1,0.3)	0.539	0.534	0.513	0.535	0.533	0.491
(0.1,0.5)	0.956	0.953	0.958	0.956	0.949	0.955
(0.3,0.3)	0.606	0.600	0.575	0.602	0.599	0.573
(0.3,0.5)	0.937	0.934	0.938	0.935	0.933	0.929
(0.5,0.5)	0.976	0.975	0.976	0.979	0.975	0.977

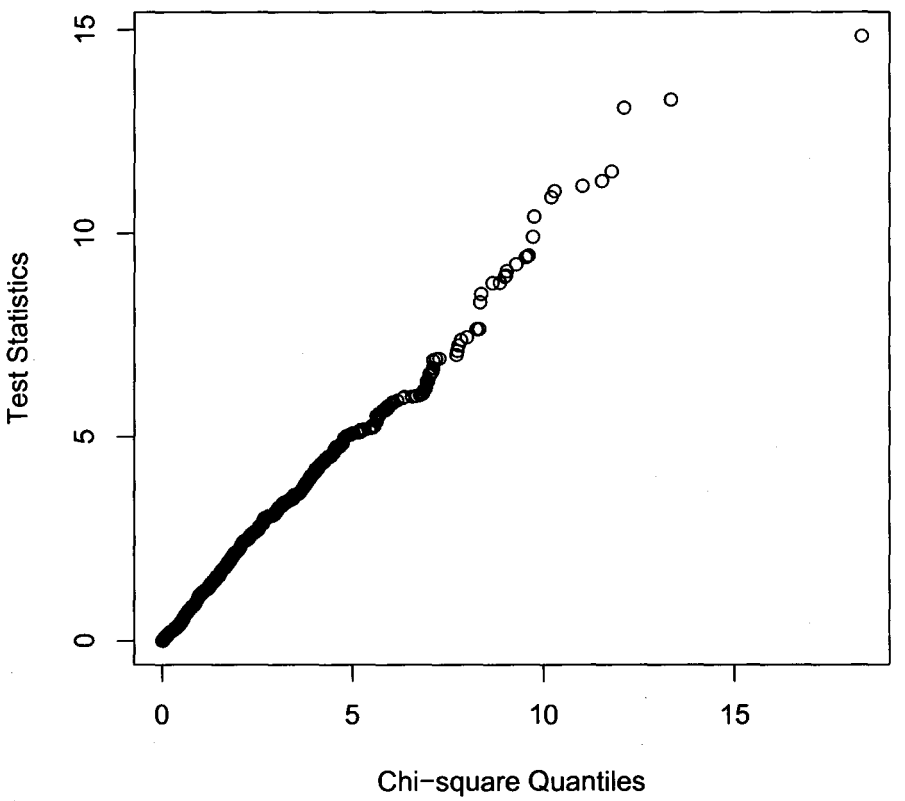

Figure 2.1: Simulation study. Chi-square quantile plot for T_1 ($n = 150$)

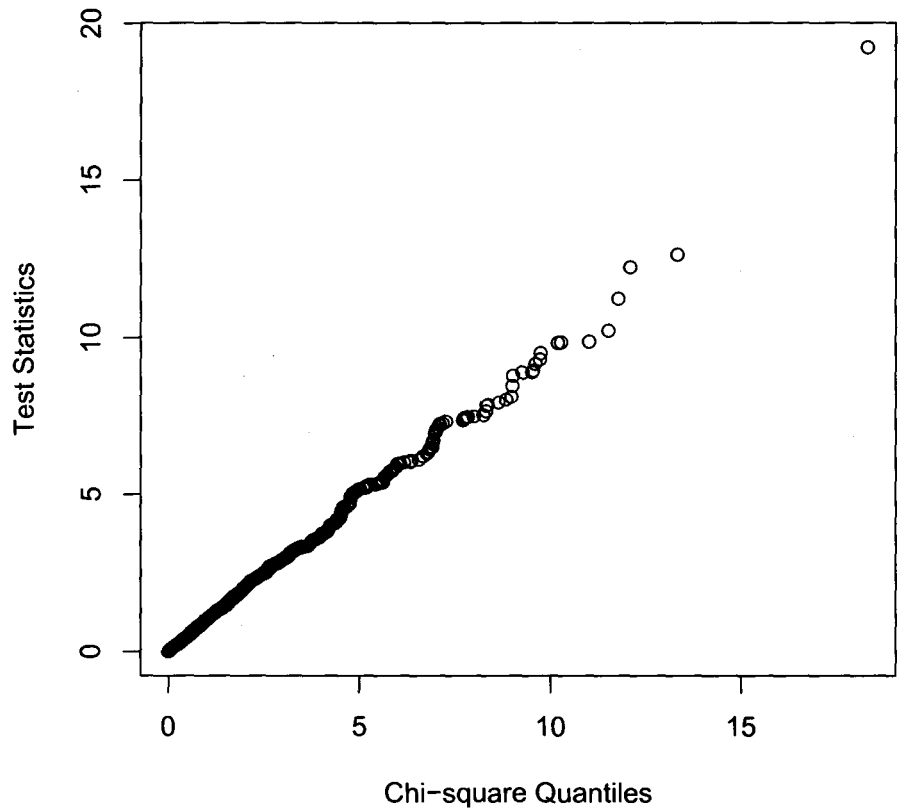

Figure 2.2: Simulation study. Chi-square quantile plot for T_1 ($n = 300$)

2.5 An Illustrative Example

To illustrate the proposed method, we consider the data from a bladder tumor study conducted by the Veterans Administration Co-operative Urological Research Group (VACURG), and the data are presented in Andrews and Herzberg (1985). For some analyses of these data, one may refer to Byar et al. (1977), Byar (1980), Wellner and Zhang (2000), Sun and Wei (2000), and Zhang (2002, 2006). The data were obtained from a randomized clinical trial. All patients had superficial bladder tumors when they entered the trial, and they were assigned randomly to one of three treatments: placebo, thiotepa and pyridoxine. The study included 116 patients, of which there were 47 in placebo group, 38 in thiotepa group and 31 in pyridoxine. At subsequent follow-up visits, any tumors noticed were removed and treatment was continued. We can get a set of panel count data $\{k_i, t_{ij}, n_{ij}, j = 1, \ldots, k_i, i = 1, \ldots, n\}$ where for the ith patient, k_i is the number of visits, t_{ij}'s are all visit times, and n_{ij} is total number of tumors until t_{ij} $(j = 1, \ldots, k_i)$. The objective of the study is to determine the effect of treatment on the frequency of tumor recurrence.

Let $G_1(t)$, $G_2(t)$ and $G_3(t)$ be as defined in Theorem 2.3.1 for the placebo, thiotepa and pyridoxine groups, respectively. We need to check $G_1(t) = G_2(t) = G_3(t)$. This can be done by the Kolmogorov-Smirnov test. Let

$$G_{n_l}(t) = \frac{1}{n_l} \sum_{i=1}^{n} \sum_{j=1}^{K_i} 1_{\{T_{K_i,j} \leq t\}},$$

which is the empirical estimator of $G_l(t)$, $l = 1, 2, 3$, where $n_1 = 48$, $n_2 = 38$, $n_3 = 31$ and $n = 116$. Let's carry out the Kolmogorov-Smirnov test to check the equality of

each pair. Define the two-sample Kolmogrov-Smirnov test statistics as follows

$$D_{n_1,n_2} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \sup_{0 \le t \le \tau} |G_{n_1}(t) - G_{n_2}(t)|,$$

$$D_{n_1,n_3} = \sqrt{\frac{n_1 n_3}{n_1 + n_3}} \sup_{0 \le t \le \tau} |G_{n_1}(t) - G_{n_3}(t)|,$$

and

$$D_{n_2,n_3} = \sqrt{\frac{n_2 n_3}{n_2 + n_3}} \sup_{0 \le t \le \tau} |G_{n_2}(t) - G_{n_3}(t)|.$$

All three pairs of the empirical functions are shown in Figure 2.3. Using two-sample bootstrap method presented by Van der Vaart and Wellner (1996, pp. 365), we obtain p-values 0.213, 0.806 and 0.385 for three Kolmogorov-Smirnov tests under the null hypotheses $G_1(t) = G_2(t)$, $G_1(t) = G_3(t)$ and $G_2(t) = G_3(t)$, respectively. These results suggest that the null hypotheses cannot be rejected. Sun and Wei (2000) and Hu et al. (2003) analyzed the data from the placebo and thiotepa groups using the regression model method and concluded that the mean function $G_l(t)$ of the counting process arising from observation times depends on the group indicator. However, their results depend on the expression of the model used and model checking is needed.

Now we can illustrate the application of our method to the bladder tumor study based on the Kolmogorov-Smirnov test results. Let $\Lambda_1(t)$, $\Lambda_2(t)$ and $\Lambda_3(t)$ be the mean functions corresponding to the three treatment groups: placebo, thiotepa and pyridoxine, respectively. The estimated mean functions from the three groups and from the pooled data are presented in Figure 2.4.

We observe from Figure 2.4 that the difference of the three groups becomes larger when the time increases. To test the null hypothesis $H_0 : \Lambda_1(t) = \Lambda_2(t) = \Lambda_3(t)$, we applied the proposed method to this panel count data and computed $T_1 = 6.139$

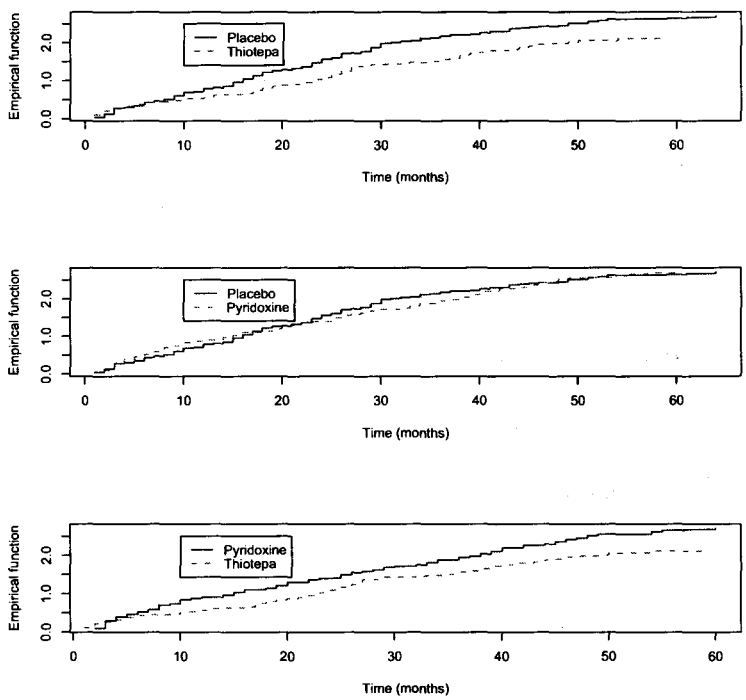

Figure 2.3: Bladder tumor study. Empirical estimates of the mean functions of counting processes from observation times

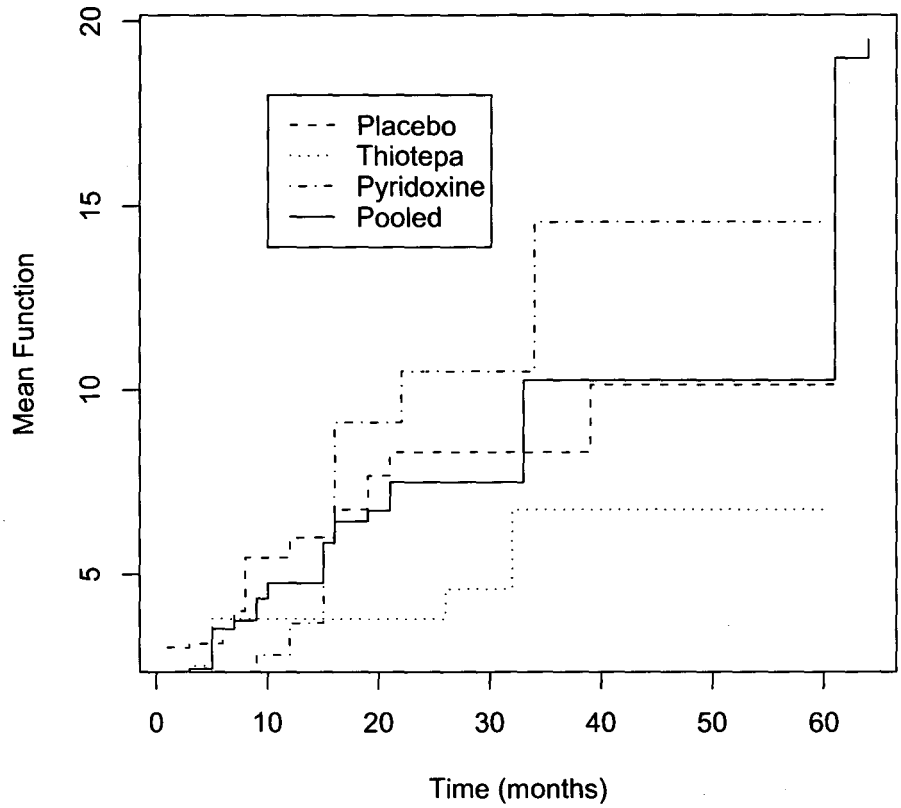

Figure 2.4: Bladder tumor study. Estimates of the mean functions

and p-value $= 0.046$ with $W_n^{(l)}(t) = W_n^{(1,l)}(t)$, $T_1 = 4.768$ and p-value $= 0.092$ with $W_n^{(l)}(t) = W_n^{(2,l)}(t)$, and $T_1 = 7.024$ and p-value $= 0.030$ with $W_n^{(l)}(t) = W_n^{(3,l)}(t) = 1-Y_{n_l}(t)$, respectively. These results suggest that the frequency of tumor recurrence are significantly different for the patients in the three groups at 10% level of significance. Incidentally, through a regression analysis of the data from two treatments, placebo and thiotepa, Sun and Wei (2000) and Zhang (2002) concluded that thiotepa effectively reduces the recurrence of tumors. Zhang (2006) obtained p-values 0.0851, 0.1445 and 0.0840 by using weight processes $W_n = 1$, $Y_n(t)$ and $1 - Y_n(t)$, respectively. If we assume that treatment indicators are independent and identically distributed random variables, then the test presented by Sun and Fang (2003) would yield p-value $= 0.082$ with the treatment indicators $z_i = -1, 1, 0$ for $i \in S_1, S_2, S_3$, p-value $= 0.696$ with the treatment indicators $z_i = -1, 0, 1$ for $i \in S_1, S_2, S_3$, p-value $= 0.064$ with the treatment indicators $z_i = 0, -1, 1$ for $i \in S_1, S_2, S_3$, p-value $= 0.139$ with the treatment indicators $z_i = 0, 1, -1$ for $i \in S_1, S_2, S_3$, p-value $= 0.628$ with the treatment indicators $z_i = 1, 0, -1$ for $i \in S_1, S_2, S_3$, and p-value $= 0.109$ with the treatment indicators $z_i = 1, -1, 0$ for $i \in S_1, S_2, S_3$. One possible reason for such a difference between these p-values is the assumption that treatment indicators are independent and identically distributed random variables, which may not be true if we look at the difference in sample sizes of the groups.

This example illustrates that different weights may result in different conclusions, and the tests with appropriate weight process could lead to better power of the test. Therefore, the selection of a suitable weight process would be important for detecting difference between groups.

2.6 Concluding Remarks

This chapter discusses the problem of the multi-sample comparison of point processes when only panel count data are available. A class of nonparametric tests are proposed for the problem and the asymptotic properties of the test statistics are derived. Simulation studies are carried out and they suggest that the proposed method works well for practical situations. The proposed method applies to more general situations than the existing methods (Thall and Lachin, 1988; Sun and Fang, 2003; Zhang, 2006).

A direction for future research is to study the properties of the test statistics under alternatives for selection of weight processes $W_n^{(l)}$'s. One can discuss the local asymptotic power of the tests and drive optimal tests along the lines of Anderson et al. (1993, pp. 372–379).

The proposed inferential procedures are established under the assumption that the observation scheme is the same for different treatment groups. This assumption may not be satisfied in many practical applications. Zhang (2006) discussed the problem and proposed an alternative test statistic which involves the estimation of $G_l'(t)$, where $G_l(t)$ is the mean function of the count process arising from observation times for group l. For the problem, we prefer to construct some test statistics which do not involve the estimation of $G_l'(t)$ and are easily computable. This is in progress.

Further research is to replace the isotonic regression estimates by maximum likelihood estimates for the mean function in the statistic $\mathbf{U}_n$. Wellner and Zhang (2000) showed that the nonparametric maximum likelihood estimator (NPMLE) of the mean function is more efficient than the nonparametric maximum pseudo-likelihood estimator (NPMPLE, the isotonic regression estimator) by means of Monte Carlo simulations.

From this, one would naturally expect that the tests based on the NPMLE could be more efficient than the proposed tests based on the NPMPLE. However, unlike the isotonic regression estimate, the maximum likelihood estimate has no closed form and its computation requires an iterative convex minorant algorithm.

Chapter 3

A New Nonparametric Test for the Equality of Counting Processes with Panel Count Data

3.1 Introduction

Consider a study that concerns some recurrent event and suppose that each subject in the study gives rise to a counting process $N(t)$, denoting the total number of occurrences of the event of interest up to time t. Also suppose that for each subject, observations include only the values of $N(t)$ at discrete observation times or the numbers of occurrences of the event between the observation times. Such data are usually referred to as *panel count data* (Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000). Our focus here will be on the situation when such a study involves k (≥ 2) groups. Let $\Lambda_l(t)$ denote the mean function of $N(t)$ corresponding to the lth group for $l = 1, \ldots, k$.

The problem of interest is then to test the hypothesis $H_0 : \Lambda_1(t) = \ldots = \Lambda_k(t)$.

A number of authors have discussed the analysis of recurrent event data when each subject in the study is observed continuously over an interval or when the exact times of occurrences of the recurrent event are known. For example, the book by Andersen et al. (1993) presents many of the commonly used statistical methods for the analysis of recurrent event data. In contrast, there exists limited research on the analysis of panel count data. Sun and Kalbfleisch (1995) and Wellner and Zhang (2000) studied estimation of the mean function of $N(t)$. Sun and Wei (2000), Zhang (2002), Hu, Sun and Wei (2003) and Sun, Tong and He (2007) discussed regression analysis for such data. To test the hypothesis H_0, Thall and Lachin (1988) suggested to transform the problem to a multivariate comparison problem and then apply a multivariate Wilcoxon-type rank test. Sun and Fang (2003) proposed a nonparametric procedure for this problem under the assumption that treatment indicators can be regarded as independent and identically distributed random variables. Park et al. (2007) proposed a class of nonparametric tests for the two-sample comparison based on the istonic regression estimator of the mean function of counting process. Zhang (2006) also presented nonparametric tests for the problem based on the nonparametric maximum pseudo-likelihood estimator that is equivalent to the istonic regression estimator (Wellner and Zhang, 2000). Also, Wellner and Zhang (2000) showed through Monte Carlo simulations that the nonparametric maximum likelihood estimator (NPMLE) of the mean function is more efficient than the nonparametric maximum pseudo-likelihood estimator (NPMPLE). However, no nonparametric tests have been discussed in the literature for panel count data based on the NPMLE since the NPMLE is more complicated both theoretically and computationally. It is, therefore, particularly important to develop

nonparametric tests based on the NPMLE for panel count data. However, unlike the isotonic regression estimate, the maximum likelihood estimate has no closed-form expression and its computation requires an iterative convex minorant algorithm. In this chapter, for simplicity, we focus on the situation considered by Sun and Fang (2003) and propose a nonparametric test using the maximum likelihood estimator and then compare its power with those of existing tests for the problem of two-sample nonparametric comparison of counting processes with simulated panel count data.

The rest of this chapter is organized as follows. Section 3.2 discusses estimation of the mean function and the existing nonparametric tests for the hypothesis H_0 when only panel count data are available. Section 3.3 presents a new nonparametric test statistic motivated by the property of the NPMLE and the idea used by Sun and Fang (2003). Also, the asymptotic normality of the test statistic is established. In Section 3.4, finite-sample property of the proposed test statistic is examined through Monte Carlo simulations. In Section 3.5, we apply the proposed method to a data from a floating gallstones study. Finally, some concluding remarks are made in Section 3.6.

3.2 Nonparametric Maximum Likelihood Estimation of Mean Function

Wellner and Zhang (2000) studied two estimators of the mean of a counting process with panel count data: the nonparametric maximum pseudo-likelihood estimator and the nonparametric maximum likelihood estimator. To describe the test statistics, we first introduce the NPMLE. Suppose that $N = \{N(t) : t \geq 0\}$ is a non-homogeneous

Poisson process with the mean function $E(N(t)) = \Lambda_0(t)$. Suppose that K is an integer-valued random variable and $T = \{T_{k,j},\ j = 1, \ldots, k,\ k = 1, 2, \ldots\}$ is a random triangular array, where $T_{k,j-1} < T_{k,j}$ and $T_{k,0} = 0$, for $j = 1, \ldots, k$ and $k = 1, 2, \ldots$. We assume that $\{(K; T_{K,1}, \ldots, T_{K,K})\}$ are independent of N. Let $X = (K, T_K, N_K)$, where T_k is the kth row of the triangular array T and $N_k = (N(T_{k,1}), \ldots, N(T_{k,k}))$. Then, $X_i = (K_i, T_{K_i}, N_{i,K_i})$, $i = 1, \ldots, n$, is a random sample of size n from the distribution of X. Let $\mathbf{X} = (X_1, \ldots, X_n)$. Then the log likelihood function for the mean function Λ is

$$l_n(\Lambda|\mathbf{X}) = \sum_{i=1}^{n} \sum_{j=1}^{K_i} (N_i(T_{K_i,j}) - N_i(T_{K_i,j-1})) \log (\Lambda(T_{K_i,j}) - \Lambda(T_{K_i,j-1})) - \sum_{i=1}^{n} \Lambda(T_{K_i,K_i})$$

after omitting the parts independent of Λ.

Let $t_1 < \ldots < t_m$ denote the ordered distinct observation time points in the set of all observation time points $\{T_{K_i,j}, j = 1, \ldots, K_i, i = 1, \ldots, n\}$. Then the NPMLE of Λ_0, $\hat{\Lambda}_n$, is defined to be the nondecreasing, non-negative step function with possible jumps only occurring at $t_\ell, \ell = 1, \ldots, m$, that maximizes $l_n(\Lambda|\mathbf{X})$. Wellner and Zhang (2000) gave the characteristic and the algorithm for computing this estimator, and studied its asymptotic properties.

The existing nonparametric tests (Park et al., 2007; Zhang, 2006) are based on the asymptotic normality of a smooth functional of the nonparametric maximum pseudo-likelihood estimator $\tilde{\Lambda}_n$ (the istonic regression estimator). However, it is unknown if the asymptotic normality of the functional of the nonparametric maximum likelihood estimator still holds because of the complexity of the NPMLE. We observe that the test presented by Sun and Fang (2003) is related to the characteristic of the $\tilde{\Lambda}_n$ given

by

$$\sum_{i=1}^{n}\sum_{j=1}^{K_i}(\tilde{\Lambda}_n(T_{K_i,j}) - N_i(T_{K_i,j})) = 0. \tag{3.1}$$

However, from equation (2.14) of Wellner and Zhang (2000), the corresponding characteristic of the NPMLE can be written as

$$\sum_{i=1}^{n}\left[\sum_{j=1}^{K_i-1}\hat{\Lambda}_n(T_{K_i,j})\left\{\frac{\Delta N_i(T_{K_i,j+1})}{\Delta\hat{\Lambda}_n(T_{K_i,j+1})} - \frac{\Delta N_i(T_{K_i,j})}{\Delta\hat{\Lambda}_n(T_{K_i,j})}\right\}\right.$$
$$\left.+\hat{\Lambda}_n(T_{K_i,K_i})\left\{1 - \frac{\Delta N_i(T_{K_i,K_i})}{\Delta\hat{\Lambda}_n(T_{K_i,K_i})}\right\}\right] = 0, \tag{3.2}$$

where $\Delta\Lambda(T_{K,j}) = \Lambda(T_{K,j}) - \Lambda(T_{K,j-1})$ and $\Delta N(T_{K,j}) = N(T_{K,j}) - N(T_{K,j-1})$. Clearly, the structure of (3.2) is different from that of (3.1) and considerably more complicated. Therefore, we need to develop a new form of test statistic when the NPMLE is used to estimate the mean function of counting process with panel count data.

3.3 A Nonparametric Test with Panel Count Data

Consider a longitudinal study that is concerned with some recurrent event and involves n independent subjects from k different groups. Let Z_i denote the group indicator of subject i ($i = 1, \ldots, n$) and assume that group indicator is a scalar variable. Let $N_i(t)$ denote the counting process arising from subject i and $\Lambda_l(t)$ ($l = 1, \ldots, k$) be defined as before, for $i = 1, \ldots, n$. Suppose that each subject is observed only at discrete time points $0 < T_{K_i,1} < \ldots < T_{K_i,K_i}$ and that no information is available about $N_i(t)$ between observation times; that is, only panel count data are available. Also assume that N_i and $(K_i, T_{K_i,1}, \ldots, T_{K_i,K_i})$ are independent of Z_i. For simplicity, assume that H_0 is true, and let $\Lambda_0(t)$ denote the common mean function of the $N_i(t)$'s.

Let $\hat{\Lambda}_n$ be the nonparametric maximum likelihood estimate of Λ_0 based on the combined data. To test the hypothesis H_0, motivated by the characteristic of the NPMLE and an idea used in Sun and Fang (2003), we propose the statistic

$$\begin{aligned} U_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i \left[\sum_{j=1}^{K_i-1} \hat{\Lambda}_n(T_{K_i,j}) \left\{ \frac{\Delta N_i(T_{K_i,j+1})}{\Delta \hat{\Lambda}_n(T_{K_i,j+1})} - \frac{\Delta N_i(T_{K_i,j})}{\Delta \hat{\Lambda}_n(T_{K_i,j})} \right\} \right. \\ &\quad \left. + \hat{\Lambda}_n(T_{K_i,K_i}) \left\{ 1 - \frac{\Delta N_i(T_{K_i,K_i})}{\Delta \hat{\Lambda}_n(T_{K_i,K_i})} \right\} \right]. \end{aligned}$$

Let $\mathcal{B}$ denote the collection of Borel sets in $\mathcal{R}$, and let $\mathcal{B}_{[0,\tau]} = \{B \cap [0,\tau], B \in \mathcal{B}\}$. On $([0,\tau], \mathcal{B}_{[0,\tau]})$, set

$$\mu(B) = \sum_{k=1}^{\infty} P(K = k) \sum_{j=1}^{k} P(T_{k,j} \in B | K = k).$$

For establishing asymptotic result on U_n, we need the following regularity conditions:

Condition 1. There exists a constant K_0 such that $\mathrm{pr}\{K \leq K_0\} = 1$ and that the random variables $T_{k,j}$'s take values in a bounded set $[\tau_0, \tau]$, where $\tau_0, \tau \in (0, \infty)$;

Condition 2. The mean function Λ_0 is continuous such that $\Lambda_0(\tau_0) > 0$ and $\Lambda_0(\tau) \leq M$ for some constant $M \in (0, \infty)$;

Condition 3. There exists a constant L_0 such that

$$P\{\min_{1 \leq j \leq K} (\Lambda_0(T_{K,j+1}) - \Lambda_0(T_{K,j})) \geq L_0\} = 1;$$

Condition 4. $E\{N(t)\}^2 \leq M_1$ for all $t \leq \tau$ where M_1 is a constant;

Condition 5. $\mu(\{\tau_0\}) > 0$ and for all $\tau_0 < \tau_1 < \tau_2 < \tau$, $\Lambda_0(\tau_1) < \Lambda_0(\tau_2) < \Lambda_0(\tau)$ implies $\mu((\tau_1, \tau_2)) > 0$.

Condition 3 holds if Λ_0 is differentiable, Λ_0' has a positive lower bound in $[\tau_0, \tau]$, and $P\{\min_{1\le j\le K}(T_{K,j} - T_{K,j-1}) \ge s_0\}$ for some fixed time s_0, where s_0 can be considered as the smallest length of consecutive observation times. Condition 5 holds if Λ_0 is strictly increasing, $P\{T_{K,1} = \tau_0\} > 0$ and $\mu'(t) > 0$ for $t \in (\tau_0, \tau)$. The asymptotic distribution of U_n is as presented in the following theorem.

Theorem 3.3.1 *Suppose that Conditions 1-5 hold. Also suppose that Z_i's are independent and identically distributed random variables. Then as $n \to \infty$,*

$$U_n \longrightarrow U$$

in distribution, where U has a normal distribution with mean zero and variance σ^2 with

$$\begin{aligned}\sigma^2 &= E\left[(Z - E(Z))\left\{\sum_{j=1}^{K-1}\Lambda_0(T_{K,j})\left(\frac{\Delta N(T_{K,j+1})}{\Delta\Lambda_0(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right)\right.\right.\\ &\qquad\left.\left.+\Lambda_0(T_{K,K})\left(1 - \frac{\Delta N(T_{K,K})}{\Delta\Lambda_0(T_{K,K})}\right)\right\}\right]^2\end{aligned}$$

which can be consistently estimated by

$$\begin{aligned}\hat\sigma^2 &= \frac{1}{n}\sum_{i=1}^{n}\left[(Z_i - \bar Z)\left\{\sum_{j=1}^{K_i-1}\hat\Lambda_n(T_{K_i,j})\left(\frac{\Delta N_i(T_{K_i,j+1})}{\Delta\hat\Lambda_n(T_{K_i,j+1})} - \frac{\Delta N_i(T_{K_i,j})}{\Delta\hat\Lambda_n(T_{K_i,j})}\right)\right.\right.\\ &\qquad\left.\left.+\hat\Lambda_n(T_{K_i,K_i})\left(1 - \frac{\Delta N_i(T_{K_i,K_i})}{\Delta\hat\Lambda_n(T_{K_i,K_i})}\right)\right\}\right]^2,\end{aligned}$$

where $\bar Z = \sum_{i=1}^n Z_i/n$.

Proof. Let

$$\mathcal{F} = \{\Lambda : [0,\tau] \to [0,\infty) | \Lambda \text{ is nondecreasing, } \Lambda(0) = 0\},$$

and let d be the $L_2(\mu)$ metric on $\mathcal{F}$. Then for $\Lambda_1, \Lambda_2 \in \mathcal{F}$,

$$\begin{aligned} d^2(\Lambda_1, \Lambda_2) &= \int |\Lambda_1(t) - \Lambda_2(t)|^2 d\mu(t) \\ &= E\left[\sum_{j=1}^{K} \{\Lambda_1(T_{K,j}) - \Lambda_2(T_{K,j})\}^2\right]. \end{aligned}$$

Wellner and Zhang (2000) showed

$$d(\hat{\Lambda}_n, \Lambda_0) \xrightarrow{a.s.} 0$$

and hence the uniform consistency of $\hat{\Lambda}_n$ can be shown by using the similar arguments to Proposition 5 of Schick and Yu (2000); that is, Conditions 1-2, 4 and 5 imply that

$$\sup_{t \in [\tau_0, \tau]} |\hat{\Lambda}_n(t) - \Lambda_0(t)| \xrightarrow{a.s.} 0.$$

Note that the uniform consistency of $\hat{\Lambda}_n$ implies for every $0 < \delta_0 < \min(L_0/2, \Lambda_0(\tau_0))$ and any $\varepsilon > 0$, there exists a positive integer N_ε such that

$$\sup_{n > N_\varepsilon} P\{\sup_{t \in [\tau_0, \tau]} |\hat{\Lambda}_n(t) - \Lambda_0(t)| > \delta_0\} < \varepsilon.$$

Here, we fix δ_0. Let

$$\mathcal{F}_0 = \{\Lambda : \Lambda \in \mathcal{F}, \sup_{t \in [\tau_0, \tau]} |\Lambda(t) - \Lambda_0(t)| \leq \delta_0\}.$$

Define $\hat{\Lambda}_n^*$ as

$$\hat{\Lambda}_n^* = \text{argmax}_{\Lambda \in \Omega \cap \mathcal{F}_0} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{K_i} (\Delta N_i(T_{K_i,j}) \log(\Delta\Lambda(T_{K_i,j})) - \Delta\Lambda(T_{K_i,j})) \right\},$$

where Ω is the class of nondecreasing step functions with possible jumps only at the observation time points $\{T_{K_i,j},\ j = 1, \ldots, K_i,\ i = 1, \ldots, n\}$. Clearly, we have

$$\sup_{n > N_\varepsilon} P(\hat{\Lambda}_n \neq \hat{\Lambda}_n^*) \leq \sup_{n > N_\varepsilon} P\{\sup_{t \in [\tau_0, \tau]} |\hat{\Lambda}_n(t) - \Lambda_0(t)| > \delta_0\} < \varepsilon.$$

Let

$$h(X,\Lambda) = \sum_{j=1}^{K-1} \Lambda(T_{K,j}) \left\{ \frac{\Delta N(T_{K,j+1})}{\Delta\Lambda(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\Lambda(T_{K_i,j})} \right\} + \Lambda(T_{K,K}) \left\{ 1 - \frac{\Delta N(T_{K,K})}{\Delta\Lambda(T_{K,K})} \right\}.$$

From equation (3.2), we have

$$\sum_{i=1}^{n} h(X_i, \hat{\Lambda}_n) = 0.$$

Now U_n can be expressed as

$$U_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{Z_i - E(Z_i)\} h(X_i, \hat{\Lambda}_n) = V_n + \Delta_n,$$

where

$$V_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{Z_i - E(Z_i)\} h(X_i, \Lambda_0)$$

and

$$\Delta_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{Z_i - E(Z_i)\} \{h(X_i, \hat{\Lambda}_n) - h(X_i, \Lambda_0)\}$$

It is easy to see that V_n is a U-statistic and has an asymptotic normal distribution with mean zero and variance σ^2 that can be consistently estimated by $\hat{\sigma}^2$ presented earlier in Theorem 3.3.1. Hence, it is sufficient to show that Δ_n converges in probability to zero.

Let Δ_n^* denote the version of Δ_n obtained by replacing $\hat{\Lambda}_n$ with $\hat{\Lambda}_n^*$. Then, to prove that Δ_n converges to zero in probability, it is sufficient to show that $\Delta_n^* = o_p(1)$ since $P\{\hat{\Lambda}_n \neq \hat{\Lambda}_n^*\} < \varepsilon$. From the assumption that Z is independent of (N, K, T), we have

$$E\left\{(\Delta_n^*)^2 \,|X_1, \ldots, X_n\right\} = \frac{\sigma_z^2}{n} \sum_{i=1}^{n} \left\{ h(X_i, \hat{\Lambda}_n^*) - h(X_i, \Lambda_0) \right\}^2$$

where $\sigma_z^2 = E\{Z - E(Z)\}^2 < \infty$. Also from the definition of $\hat{\Lambda}_n^*$ and Conditions 1-3, we have

$$0 < \Lambda_0(\tau_0) - \delta_0 \leq \Lambda_0(t) - \delta_0 \leq \hat{\Lambda}_n^*(t) \leq \Lambda_0(t) + \delta_0 \leq \Lambda_0(\tau) + \delta_0 \leq M + \delta_0$$

for $t \in [\tau_0, \tau]$ and

$$0 < L_0 - 2\delta_0 \leq \Delta\Lambda_0(T_{K_i,j}) - 2\delta_0 \leq \Delta\hat{\Lambda}_n^*(T_{K_i,j}) \leq \Delta\Lambda_0(T_{K_i,j}) + 2\delta_0 \leq 2M + 2\delta_0$$

for $j = 1, ..., K_i$, $i = 1, \ldots, n$ with probability 1. Hence, we have

$$\begin{aligned}
&\left|h(X_i, \hat{\Lambda}_n^*) - h(X_i, \Lambda_0)\right| \\
\leq\ & c_1 \sum_{j=1}^{K_i-1} \Delta N_i(T_{K_i,j+1}) \left\{|\hat{\Lambda}_n^*(T_{K_i,j}) - \Lambda_0(T_{K_i,j})| + |\hat{\Lambda}_n^*(T_{K_i,j+1}) - \Lambda_0(T_{K_i,j+1})|\right\} \\
&+c_2 \sum_{j=1}^{K_i-1} \Delta N_i(T_{K_i,j}) \left\{|\hat{\Lambda}_n^*(T_{K_i,j}) - \Lambda_0(T_{K_i,j})| + |\hat{\Lambda}_n^*(T_{K_i,j-1}) - \Lambda_0(T_{K_i,j-1})|\right\} \\
&+|\hat{\Lambda}_n^*(T_{K_i,K_i}) - \Lambda_0(T_{K_i,K_i})| \\
&+c_3 \sum_{j=1}^{K_i-1} \Delta N_i(T_{K_i,K_i}) \left\{|\hat{\Lambda}_n^*(T_{K_i,K_i}) - \Lambda_0(T_{K_i,K_i})| + |\hat{\Lambda}_n^*(T_{K_i,K_i-1}) - \Lambda_0(T_{K_i,K_i-1})|\right\} \\
\leq\ & c_4 \left\{1 + \sum_{j=1}^{K_i} \Delta N_i(T_{K_i,j})\right\} \sup_{t\in[\tau_0,\tau]} |\hat{\Lambda}_n^*(t) - \lambda_0(t)| \\
=\ & c_4 \left\{1 + N_i(T_{K_i,K_i})\right\} \sup_{t\in[\tau_0,\tau]} |\hat{\Lambda}_n^*(t) - \lambda_0(t)|,
\end{aligned}$$

for some constants c_1, c_2, c_3 and c_4, and so

$$\frac{1}{n}\sum_{i=1}^{n} \left\{h(X_i, \hat{\Lambda}_n^*) - h(X_i, \Lambda_0)\right\}^2 \leq c_4 \sup_{t\in[\tau_0,\tau]} |\hat{\Lambda}_n^*(t) - \Lambda_0(t)|^2 \frac{1}{n}\sum_{i=1}^{n} \left\{1 + N_i(T_{K_i,K_i})\right\}^2.$$

Thus, $\Delta_n^* = o_p(1)$. This completes the proof of the theorem. □

Remark. Note that the asymptotic result requires Z_i's as independent and identically distributed random variables. For example, this assumption is satisfied in randomized clinical trials, where all patients in the study are randomly assigned to one of the treatments.

3.4 Simulation Study

To examine the finite-sample property of the proposed test statistic and compare it with those of the tests presented by Sun and Fang (2003), Park et al. (2007) and Zhang (2006), we carry out a simulation study for the two-sample comparison problem. When $k = 2$, then the null hypothesis can be tested by $T = U_n/\hat{\sigma}$ which has asymptotic standard normal distribution, where U_n and $\hat{\sigma}$ are as presented in Theorem 3.3.1. Let T_{SF} denote the test proposed by Sun and Fang (2003), and let T_i (i=1, 2, 3) denote the tests presented by Park et al. (2007) and Zhang (2006) with three different wight processes: $W_n^{(1)}(t) = 1$, $W_n^{(2)}(t) = Y_n(t) = \sum_{i=1}^n I(t \leq T_{K_i,K_i})/n$, and $W_n^{(3)}(t) = \{Y_{n_1}Y_{n_2}\}/Y_n(t)$, where $Y_{n_l} = \sum_{i \in S_l} I(t \leq T_{K_i,K_i})/n_l$, S_l denotes the set of indices for subjects in group l and n_l is the number of subjects in group l, $l = 1, 2$. To generate panel count data $\{K_i, T_{K_i,j}, N_i(T_{K_i,j}), j = 1, \ldots, K_i, i = 1, \ldots, n\}$, we mimic medical follow-up studies similar to the example discussed in the next section. We first generate the number of observation times K_i from the uniform distribution $U\{1, \ldots, 10\}$, and then, given K_i, we generate observation times $T_{K_i,j}$'s from $U\{1, \ldots, 10\}$, for simplicity. To generate $N_i(T_{K_i,j})$'s, we assume that N_i's are nonhomogeneous Poisson or mixed Poisson processes. In particular, let $\{\nu_i, i = 1, \ldots, n\}$ be independent and identically distributed random variables, and given ν_i, let $N_i(t)$ be a Poisson process with mean function $\Lambda_i(t) = \nu_i t$ for $i \in S_1$, $\Lambda_i(t) = \nu_i t \exp(\beta)$ for $i \in S_2$. Here, it is assumed that $Z_i = 0$ for $i \in S_1$ and $Z_i = 1$ for $i \in S_2$.

We consider two cases: $\nu_i = 1$ and $\nu_i \sim$ Gamma$(2, 1/2)$. For each case, we consider two sample sizes, $n_1 = n_2 = 50$ and 100, respectively. The NPMLE $\hat{\Lambda}_n$ is computed by using the modified iterative convex minorant algorithm (MICM); see Wellner and

Zhang (2000). All the results reported are based on 1000 Monte Carlo replications using R software.

Table 3.1: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for Poisson processes

β	T	T_{SF}	T_1	T_2	T_3
		$n_1 = n_2 = 50$			
0.0	0.051	0.047	0.053	0.055	0.055
0.1	0.298	0.207	0.214	0.200	0.200
0.2	0.855	0.693	0.697	0.667	0.665
0.3	1.000	0.979	0.981	0.974	0.974
		$n_1 = n_2 = 100$			
0.0	0.049	0.041	0.043	0.047	0.047
0.1	0.553	0.422	0.423	0.405	0.405
0.2	0.990	0.957	0.958	0.948	0.947
0.3	1.000	1.000	1.000	1.000	1.000

Table 3.2: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for mixed Poisson processes

β	T	T_{SF}	T_1	T_2	T_3
		$n_1 = n_2 = 50$			
0.0	0.046	0.044	0.045	0.047	0.047
0.1	0.098	0.083	0.084	0.085	0.085
0.2	0.223	0.183	0.185	0.184	0.184
0.3	0.450	0.370	0.380	0.375	0.375
		$n_1 = n_2 = 100$			
0.0	0.043	0.046	0.048	0.045	0.045
0.1	0.141	0.111	0.114	0.111	0.111
0.2	0.411	0.316	0.317	0.307	0.307
0.3	0.710	0.590	0.596	0.592	0.592

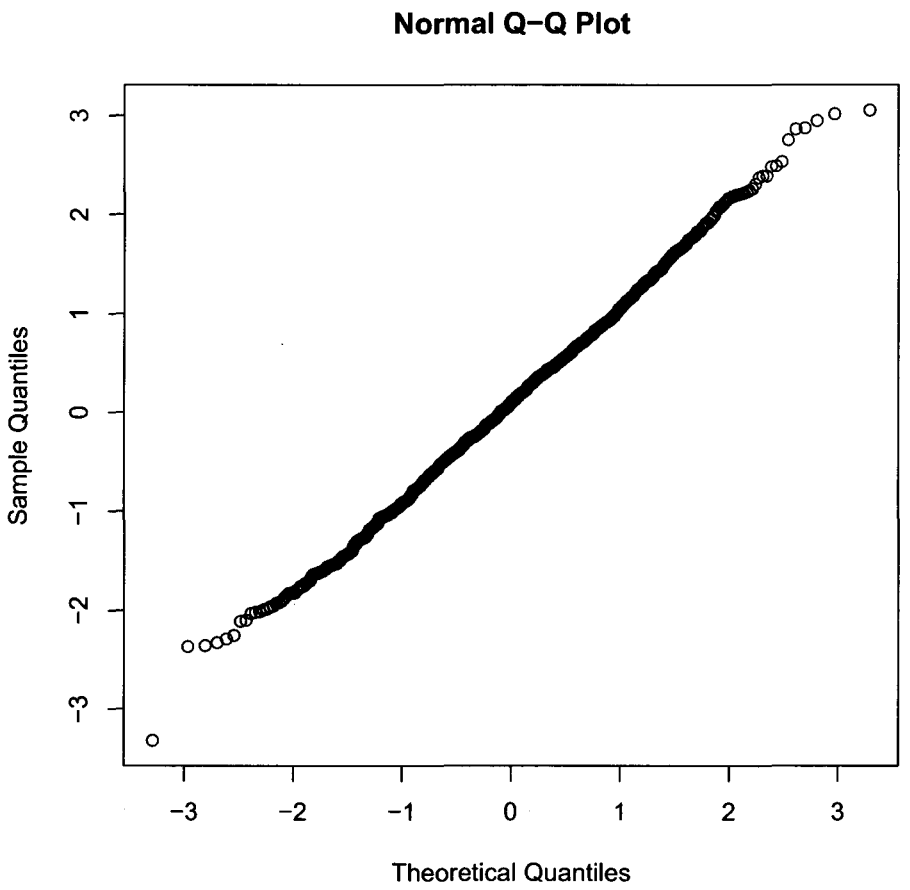

Figure 3.1: Simulation study. Normal quantile plot ($n = 100$).

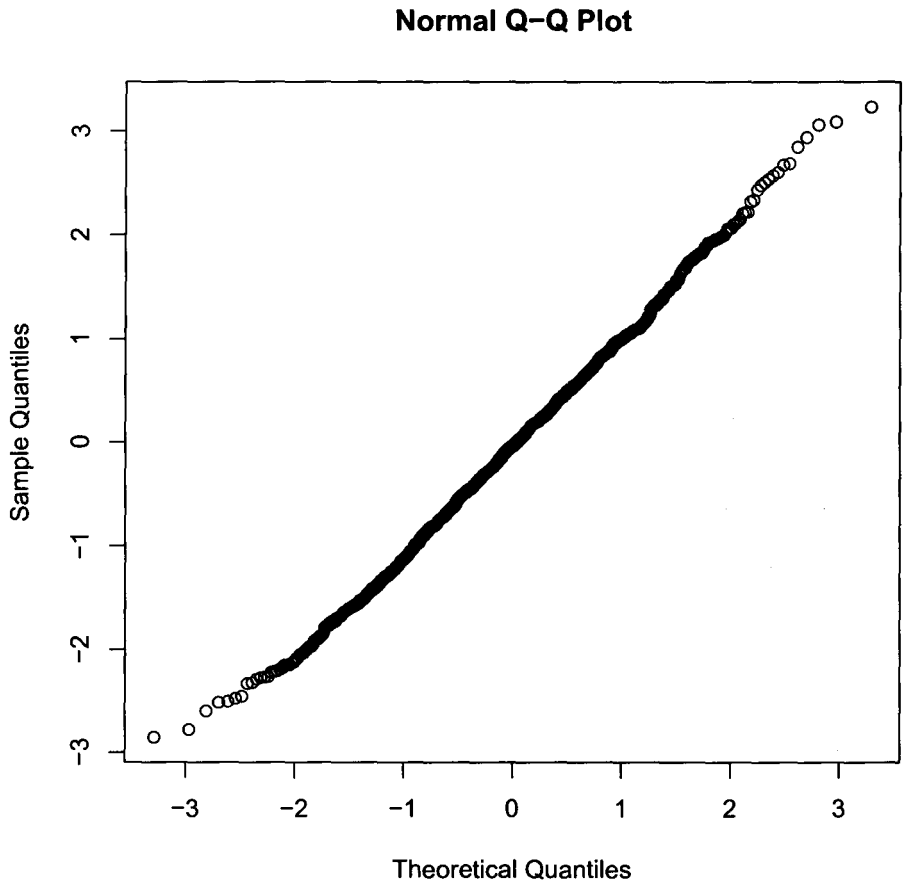

Figure 3.2: Simulation study. Normal quantile plot ($n = 200$).

Tables 3.1 and 3.2 present the estimated sizes and powers of the proposed test statistic T, the test T_{SF} (Sun and Fang, 2003) and the test statistics T_i's (Park et al., 2007 and Zhang, 2006) at significance level $\alpha = 0.05$ for different values of β for the two cases, respectively. In the first case, the $N_i(t)$'s are Poisson processes. In the second case, the $N_i(t)$'s are mixed Poisson processes. The first part of the table is for the situation with the total sample size of 100 and the second part is for the situation with the total sample size of 200. For the situation considered here, the proposed test displays the highest power. As expected, the power increases when the sample size increases, and the power decreases in the presence of more variability. To evaluate the asymptotic result presented in Theorem 3.3.1, the quantile plots of the proposed test statistic T against the standard normal distribution are constructed. Figures 3.1 and 3.2 present the plots for $n = 100$ and $n = 200$, respectively, and they clearly reveal that the asymptotic approximation is quite good. From Tables 3.1 and 3.2, we conclude that the proposed test based on the NPMLE is more powerful than the existing tests based on NPMPLE.

3.5 An Illustrative Example

To illustrate the proposed method, we consider a floating gallstones study presented by Thall and Lachin (1988). The data comprise the first year follow-up of the patients in two study groups, placebo (48) and high-dose chenodiol (65), from the National Cooperative Gallstone Study. The data include the successive visit-times in study weeks and the associated counts of episodes of nausea. The whole study consists of 916 patients who were randomized to placebo, low dose, or high dose group and followed

for up to two years and one of the objectives of the study is to test the difference of the two treatments with respect to the incidence rate of nausea.

During the study, patients were scheduled to return for clinical visits at 1, 2, 3, 6, 9, and 12 months. In reality, most of the patients visited about six times within the first year, but actual visit times differed from patient to patient. Some patients had only one visit and some had 9 visits. As pointed out by Thall and Lachin (1988), there is no evidence that the number of observations and actual observation times are related to the incidence of nausea, and so it seems reasonable to assume that conditions required for the asymptotic result are satisfied.

Define $Z_i = 1$ for patients in the placebo group and $Z_i = 0$ otherwise. To test the difference between the two groups, we apply the proposed method to the data from 113 gallstone patients in the two groups and obtain $T = 0.264$ which yields a p-value of 0.792 for testing H_0 based on the asymptotic result in Theorem 3.3.1. This result suggests that the incidence rates of nausea were not significantly different for the patients in the two groups, which agrees with the findings of Schoenfield et al. (1981), Sun and Fang (2003) and Park et al. (2007).

3.6 Concluding Remarks

This chapter discusses the problem of the nonparametric comparison of counting processes when only panel count data are available. The nonparametric maximum likelihood estimators are used to estimate the mean functions of counting processes. A new nonparametric test is proposed for the problem and the asymptotic property of the test statistic is derived. Simulation studies are carried out which suggest that the

proposed method works well for practical situations, and is more powerful than the existing tests based on the nonparametric maximum pseudo-likelihood estimators of the mean functions.

Note that the proposed procedure depends on the assumption that treatment indicators can be regarded as independent and identically distributed random variables. Next chapter will be to develop a class of tests applicable to general situations by using the nonparametric maximum likelihood estimates instead of the nonparametric maximum pseudo-likelihood estimates for the mean functions.

Chapter 4

New Nonparametric Tests for Panel Count Data Based on Likelihood Approach

4.1 Introduction

Consider a study that concerns some recurrent event and suppose that each subject in the study gives rise to a counting process $N(t)$, denoting the total number of occurrences of the event of interest up to time t. Also suppose that for each subject, observations include only the values of $N(t)$ at discrete observation times or the numbers of occurrences of the event between the observation times. Such data are usually referred to as *panel count data* (Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000). Our focus here will be on the situation when such a study involves k (≥ 2) groups. Let $\Lambda_l(t)$ denote the mean function of $N(t)$ corresponding to the lth group for $l = 1, \ldots, k$.

The problem of interest is then to test the hypothesis $H_0 : \Lambda_1(t) = \cdots = \Lambda_k(t)$.

A number of authors have discussed the analysis of recurrent event data when each subject in the study is observed continuously over an interval or when the exact times of occurrences of the recurrent event are known. For example, the book by Andersen et al. (1993) presents many of the commonly used statistical methods for the analysis of recurrent event data. In contrast, there exists limited research on the analysis of panel count data. Sun and Kalbfleisch (1995) and Wellner and Zhang (2000) studied estimation of the mean function of $N(t)$. Sun and Wei (2000) and Zhang (2002) discussed regression analysis for such data. To test the hypothesis H_0, Thall and Lachin (1988) suggested to transform the problem to a multivariate comparison problem and then apply a multivariate Wilcoxon-type rank test. Sun and Fang (2003) proposed a nonparametric procedure for this problem, but their procedure depends on the assumption that treatment indicators can be regarded as independent and identically distributed random variables, which may not be the case in practice. Park et al. (2007) proposed a class of nonparametric tests for the two-sample comparison based on the istonic regression estimator of the mean function of counting process. Zhang (2006) also presented nonparametric tests for the problem based on the nonparametric maximum pseudo-likelihood estimator which is equivalent to the istonic regression estimator (Wellner and Zhang, 2000). Also, Wellner and Zhang (2000) showed through Monte Carlo simulations that the nonparametric maximum likelihood estimator (NPMLE) of the mean function is more efficient than the nonparametric maximum pseudo-likelihood estimator (NPMPLE). However, no nonparametric tests have been discussed in the literature for panel count data based on the NPMLE since the NPMLE is more complicated both theoretically and computationally. It is, therefore, particularly important to develop

nonparametric tests based on the NPMLE for panel count data. One would naturally expect the tests based on the NPMLE to be more powerful than the tests based on the NPMPLE. However, unlike the isotonic regression estimate, the maximum likelihood estimate has no closed-form expression and its computation requires an iterative convex minorant algorithm. In this chapter, we propose some nonparametric tests based on the maximum likelihood estimator and then compare them with the existing tests for the problem of multi-sample nonparametric comparison of counting processes with panel count data.

The rest of this chapter is organized as follows. Section 4.2 discusses estimation of the mean function and the existing nonparametric tests for the hypothesis H_0 when only panel count data are available. The asymptotic normality of the functional of the NPMLE is established, while its proof is presented in Section 4.6. Section 4.3 presents two classes of nonparametric test statistics. The statistics, motivated by the property of the NPMLE and the idea used in survival analysis, are formulated as the integrated weighted difference between the rates of increase of the estimated mean functions corresponding to the pooled data and each group or two groups. The asymptotic normality of these test statistics is also established, while proofs are given in Section 4.6. In Section 4.4, finite-sample properties of the proposed test statistics are examined through Monte Carlo simulations. In Section 4.5, we apply the proposed methods to two data from a floating gallstones study and a bladder tumor study, respectively.

4.2 Nonparametric Maximum Likelihood Estimation of Mean Function

Wellner and Zhang (2000) studied two estimators of the mean of a counting process with panel count data: the nonparametric maximum pseudo-likelihood estimator and the nonparametric maximum likelihood estimator. To describe the test statistics, we introduce first the NPMLE. Suppose that $N = \{N(t) : t \geq 0\}$ is a non-homogeneous Poisson process with the mean function $E(N(t)) = \Lambda_0(t)$. Suppose that K is an integer-valued random variable and $T = \{T_{k,j},\ j = 1, \ldots, k,\ k = 1, 2, \ldots\}$ is a random triangular array, where $T_{k,j-1} < T_{k,j}$ and $T_{k,0} = 0$, for $j = 1, \ldots, k$ and $k = 1, 2, \ldots$. We assume that $\{(K; T_{K,1}, \ldots, T_{K,K})\}$ are independent of N. Let $X = (K, T_K, N_K)$, where T_k is the kth row of the triangular array T and $N_k = (N(T_{k,1}), \ldots, N(T_{k,k}))$. Then, $X_i = (K_i, T_{K_i}, N_{i,K_i})$, $i = 1, \ldots, n$, is a random sample of size n from the distribution of X. Let $\mathbf{X} = (X_1, \ldots, X_n)$. Then, the log-likelihood function for the mean function Λ is

$$\begin{aligned} l_n(\Lambda|\mathbf{X}) &= \sum_{i=1}^{n}\sum_{j=1}^{K_i} (N_i(T_{K_i,j}) - N_i(T_{K_i,j-1})) \log (\Lambda(T_{K_i,j}) - \Lambda(T_{K_i,j-1})) \\ &\quad - \sum_{i=1}^{n} \Lambda(T_{K_i,K_i}) \end{aligned}$$

after omitting the parts independent of Λ.

Let $t_1 < \cdots < t_m$ denote the ordered distinct observation time points in the set of all observation time points $\{T_{K_i,j}, j = 1, \ldots, K_i, i = 1, \ldots, n\}$. Then the NPMLE of Λ_0, $\hat{\Lambda}_n$, is defined to be the nondecreasing, non-negative step function with possible jumps only occurring at $t_\ell, \ell = 1, \ldots, m$, that maximizes $l_n(\Lambda|\mathbf{X})$. Wellner and Zhang (2000) gave the characteristic and the algorithm for computing this estimator, and

studied its asymptotic properties.

Next, we need some more notation, some of which were introduced by Schick and Yu (2000) and Wellner and Zhang (2000). Let $\mathcal{B}$ denote the collection of Borel sets in $\mathcal{R}$, and let $\mathcal{B}_{[0,\tau]} = \{B \cap [0,\tau] : B \in \mathcal{B}\}$. Define measures μ_1, μ_2, μ_3 and ν on $([0,\tau], \mathcal{B}_{[0,\tau]})$ by

$$\mu_1(B) = \sum_{k=1}^{\infty} P(K=k) \sum_{j=1}^{k} P(T_{k,j} \in B | K=k),$$

$$\mu_2(B_1 \times B_2) = \sum_{k=1}^{\infty} P(K=k) \sum_{j=1}^{k} P(T_{k,j-1} \in B_1, T_{k,j} \in B_2 | K=k),$$

$$\begin{aligned}&\mu_3(B_1 \times B_2 \times B_3)\\ &= \sum_{k=2}^{\infty} P(K=k) \sum_{j=1}^{k-1} P(T_{k,j-1} \in B_1, T_{k,j} \in B_2, T_{k,j+1} \in B_3 | K=k),\end{aligned}$$

and

$$\nu(B_1 \times B_2) = \sum_{k=1}^{\infty} P(K=k) P(T_{k,k-1} \in B_1, T_{k,k} \in B_2 | K=k)$$

for $B, B_1, B_2, B_3 \in \mathcal{B}_{[0,\tau]}$.

The existing nonparametric tests (Park et al., 2007; Zhang, 2006) are based on the asymptotic normality of a smooth functional of the nonparametric maximum pseudo-likelihood estimator (the istonic regression estimator) $\tilde{\Lambda}_n$,

$$\int_0^{\tau} W(t)\{\tilde{\Lambda}_n(t) - \Lambda_0(t)\} d\mu_1(t) = P\left[\sum_{j=1}^{K} W(T_{K,j})\{\tilde{\Lambda}_n(T_{K,j}) - \Lambda_0(T_{K,j})\}\right],$$

where $W(t)$ is a weight function, and P is the probability measure of X, $Pf = \int f dP$. However, it is unknown if the asymptotic normality of the functional of the nonparametric maximum likelihood estimator, $\int_0^{\tau} W(t)\{\hat{\Lambda}_n(t) - \Lambda_0(t)\} d\mu_1(t)$, still holds. We

observe a key to the proof of such asymptotic normality is to use an important characteristic of the $\tilde{\Lambda}_n$ given by

$$\sum_{i=1}^{n}\sum_{j=1}^{K_i}\varphi(\tilde{\Lambda}_n(T_{K_i,j}))(\tilde{\Lambda}_n(T_{K_i,j})-N_i(T_{K_i,j}))=0 \tag{4.1}$$

for any real function φ. However, from (2.13) of Wellner and Zhang (2000), the corresponding characteristic of the NPMLE can be written as

$$\begin{aligned}\sum_{i=1}^{n}\Bigg[\sum_{j=1}^{K_i-1}\hat{\Lambda}_n(T_{K_i,j})\left\{\frac{\Delta N_i(T_{K_i,j+1})}{\Delta\hat{\Lambda}_n(T_{K_i,j+1})}-\frac{\Delta N_i(T_{K_i,j})}{\Delta\hat{\Lambda}_n(T_{K_i,j})}\right\}\\ +\hat{\Lambda}_n(T_{K_i,K_i})\left\{1-\frac{\Delta N_i(T_{K_i,K_i})}{\Delta\hat{\Lambda}_n(T_{K_i,K_i})}\right\}\Bigg]=0,\end{aligned} \tag{4.2}$$

where

$$\Delta\Lambda(T_{K,j})=\Lambda(T_{K,j})-\Lambda(T_{K,j-1})$$

and

$$\Delta N(T_{K,j})=N(T_{K,j})-N(T_{K,j-1}).$$

(2.2) can be extended to the form

$$\begin{aligned}\sum_{i=1}^{n}\Bigg[\sum_{j=1}^{K_i-1}\varphi(\hat{\Lambda}_n(T_{K_i,j}))\hat{\Lambda}_n(T_{K_i,j})\left\{\frac{\Delta N_i(T_{K_i,j+1})}{\Delta\hat{\Lambda}_n(T_{K_i,j+1})}-\frac{\Delta N_i(T_{K_i,j})}{\Delta\hat{\Lambda}_n(T_{K_i,j})}\right\}\\ +\varphi(\hat{\Lambda}_n(T_{K_i,K_i}))\hat{\Lambda}_n(T_{K_i,K_i})\left\{1-\frac{\Delta N_i(T_{K_i,K_i})}{\Delta\hat{\Lambda}_n(T_{K_i,K_i})}\right\}\Bigg]=0,\end{aligned} \tag{4.3}$$

which will be shown in Lemma 4.6.1. Clearly, the structure of (4.3) is different from that of (4.1) and is much more complicated. This is why the derivation of the asymptotic property of $\int_0^\tau W(t)\{\hat{\Lambda}_n(t)-\Lambda_0(t)\}\,d\mu_1(t)$ has not been done yet. So, we need to develop a new form of the test statistic when the NPMLE is used to estimate the mean function of counting process with panel count data. Motivated by such characteristic

of the NPMLE, we define

$$\begin{aligned} f_\Lambda(X) &= \sum_{j=1}^{K-1} W(T_{K,j})\Lambda(T_{K,j}) \left\{ \frac{\Delta\Lambda_0(T_{K,j+1})}{\Delta\Lambda(T_{K,j+1})} - \frac{\Delta\Lambda_0(T_{K,j})}{\Delta\Lambda(T_{K,j})} \right\} \\ &\quad + W(T_{K,K})\Lambda(T_{K,K}) \left\{ 1 - \frac{\Delta\Lambda_0(T_{K,K})}{\Delta\Lambda(T_{K,K})} \right\}. \end{aligned} \tag{4.4}$$

It is easy to see that $Pf_\Lambda(X)$ can be expressed as

$$\begin{aligned} & Pf_\Lambda(X) \\ &= \int\int\int W(u)\Lambda(u) \left\{ \frac{\Lambda_0(v) - \Lambda_0(u)}{\Lambda(v) - \Lambda(u)} - \frac{\Lambda_0(u) - \Lambda_0(t)}{\Lambda(u) - \Lambda(t)} \right\} d\mu_3(t, u, v) \\ &\quad + \int\int W(u)\Lambda(u) \left\{ 1 - \frac{\Lambda_0(u) - \Lambda_0(t)}{\Lambda(u) - \Lambda(t)} \right\} d\nu(t, u). \end{aligned}$$

For establishing asymptotic results on $Pf_{\hat{\Lambda}_n}(X)$, we need the following regularity conditions:

A. There exists a constant K_0 such that $P\{K \leq K_0\} = 1$ and that the random variables $T_{k,j}$'s take values in a bounded set $[\tau_0, \tau]$, where $\tau_0, \tau \in (0, \infty)$;

B. The mean function Λ_0 is strictly increasing such that $\Lambda_0(\tau_0) > 0$ and $\Lambda_0(\tau) \leq M$ for some constant $M \in (0, \infty)$;

C. There exists a constant L_0 such that

$$P\left\{ \min_{1 \leq j \leq K} (\Lambda_0(T_{K,j}) - \Lambda_0(T_{K,j-1})) \geq L_0 \right\} = 1;$$

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

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

Condition C holds if Λ_0 is differentiable, Λ_0' has a positive lower bound in $[\tau_0, \tau]$, and $P\{\min_{1 \leq j \leq K}(T_{K,j} - T_{K,j-1}) \geq s_0\} = 1$ for some fixed time s_0, where s_0 can be

considered as the smallest length of consecutive observation times. Condition E holds if $P\{T_{K,1} = \tau_0\} > 0$ and $\mu_1'(t) > 0$ for $t \in (\tau_0, \tau)$.

Now let Λ_0^{-1} denote the inverse function of Λ_0, and let $W \circ \Lambda_0^{-1}$ denote composition of two functions W and Λ_0^{-1}. Zhang (2006) established the asymptotic normality of $\int_0^\tau W(t)\{\tilde{\Lambda}_n(t) - \Lambda_0(t)\}d\mu_1(t)$ when $W \circ \Lambda_0^{-1}$ is not only bounded Lipschitz but also monotone. However, the assumption that $W \circ \Lambda_0^{-1}$ is monotone is not required for the tests with interval-censored data as a special case of panel count data; see Huang and Wellner (1995) and Zhang et al. (2001). Here, we do not need this monotonicity condition for $W \circ \Lambda_0^{-1}$.

Theorem 4.2.1 *Suppose that Conditions A, B, C, D and E hold. Further, suppose that $W(t)$ is a bounded weight process such that $W \circ \Lambda_0^{-1}$ is a bounded Lipschitz function. Then as $n \to \infty$,*

$$\sqrt{n} P f_{\hat{\Lambda}_n}(X) \longrightarrow U_w \tag{4.5}$$

in distribution, where U_w has a normal distribution with mean zero and variance σ_w^2 with

$$\begin{aligned} \sigma_w^2 &= E\left[\sum_{j=1}^{K-1} W(T_{K,j})\Lambda_0(T_{K,j})\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\Lambda_0(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right\}\right. \\ &\quad \left. + W(T_{K,K})\Lambda_0(T_{K,K})\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\Lambda_0(T_{K,K})}\right\}\right]^2 \end{aligned} \tag{4.6}$$

which can be consistently estimated by

$$\begin{aligned} \hat{\sigma}_w^2 &= \frac{1}{n}\sum_{i=1}^{n}\left[\sum_{j=1}^{K_i-1} W(T_{K_i,j})\hat{\Lambda}_n(T_{K_i,j})\left\{\frac{\Delta N_i(T_{K_i,j+1})}{\Delta\hat{\Lambda}_n(T_{K_i,j+1})} - \frac{\Delta N_i(T_{K_i,j})}{\Delta\hat{\Lambda}_n(T_{K_i,j})}\right\}\right. \\ &\quad \left. + W(T_{K_i,K_i})\hat{\Lambda}_n(T_{K_i,K_i})\left\{1 - \frac{\Delta N_i(T_{K_i,K_i})}{\Delta\hat{\Lambda}_n(T_{K_i,K_i})}\right\}\right]^2. \end{aligned} \tag{4.7}$$

4.3 Nonparametric Tests

Consider a longitudinal study that is concerned with some recurrent event and involves n independent subjects, n_l in the lth group with $n_1 + \cdots + n_k = n$ and $k \geq 2$. Let $N_i(t)$ denote the counting process arising from subject i and $\Lambda_l(t)$ $(l = 1, \ldots, k)$ be as defined before, for $i = 1, \ldots, n$. Suppose that each subject is observed only at discrete time points $0 < T_{K_i,1} < \cdots < T_{K_i,K_i}$ and that no information is available about $N_i(t)$ between observation times; that is, only panel count data are available. For simplicity, assume that H_0 is true, and let $\Lambda_0(t)$ denote the common mean function of the $N_i(t)$'s.

Let $\hat{\Lambda}_{n_l}$ denote the nonparametric maximum likelihood estimate of Λ_l based on samples from all the subjects in the lth group, and $\hat{\Lambda}_n$ based on the pooled data. To test the hypothesis H_0, motivated by our asymptotic results in Section 4.2 and an idea commonly used in survival analysis (Andersen et al., 1993 , Pepe and Fleming, 1989; Petroni and Wolfe, 1994; Cook et al., 1996; Zhang et al., 2001; Park et al., 2007; Zhang, 2006), we propose the statistics

$$\begin{aligned} U_n^{(l)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\sum_{j=1}^{K_i-1} W_n^{(l)}(T_{K_i,j}) \hat{\Lambda}_n(T_{K_i,j}) \right. \\ &\qquad \times \left\{ \frac{\Delta \hat{\Lambda}_{n_l}(T_{K_i,j+1})}{\Delta \hat{\Lambda}_n(T_{K_i,j+1})} - \frac{\Delta \hat{\Lambda}_{n_l}(T_{K_i,j})}{\Delta \hat{\Lambda}_n(T_{K_i,j})} \right\} \\ &\qquad \left. + W_n^{(l)}(T_{K_i,K_i}) \hat{\Lambda}_n(T_{K_i,K_i}) \left\{ 1 - \frac{\Delta \hat{\Lambda}_{n_l}(T_{K_i,K_i})}{\Delta \hat{\Lambda}_n(T_{K_i,K_i})} \right\} \right] \end{aligned} \tag{4.8}$$

(for $l = 1, \ldots, k$) and

$$\begin{aligned} V_n^{(l)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\sum_{j=1}^{K_i-1} W_n^{(l)}(T_{K_i,j}) \hat{\Lambda}_n(T_{K_i,j}) \right. \\ &\qquad \times \left\{ \left(\frac{\Delta \hat{\Lambda}_{n_1}(T_{K_i,j+1})}{\Delta \hat{\Lambda}_n(T_{K_i,j+1})} - \frac{\Delta \hat{\Lambda}_{n_1}(T_{K_i,j})}{\Delta \hat{\Lambda}_n(T_{K_i,j})} \right) \right. \end{aligned}$$

$$
\begin{aligned}
&\quad - \left(\frac{\Delta \hat{\Lambda}_{n_l}(T_{K_i,j+1})}{\Delta \hat{\Lambda}_n(T_{K_i,j+1})} - \frac{\Delta \hat{\Lambda}_{n_l}(T_{K_i,j})}{\Delta \hat{\Lambda}_n(T_{K_i,j})} \right) \Bigg\} \\
&+ W_n^{(l)}(T_{K_i,K_i}) \hat{\Lambda}_n(T_{K_i,K_i}) \\
&\times \left\{ \left(1 - \frac{\Delta \hat{\Lambda}_{n_1}(T_{K_i,K_i})}{\Delta \hat{\Lambda}_n(T_{K_i,K_i})} \right) - \left(1 - \frac{\Delta \hat{\Lambda}_{n_l}(T_{K_i,K_i})}{\Delta \hat{\Lambda}_n(T_{K_i,K_i})} \right) \right\} \Bigg] \qquad (4.9)
\end{aligned}
$$

(for $l = 2, \ldots, k$), and $W_n^{(l)}(t)$'s are bounded weight processes. The statistic $U_n^{(l)}$ is the integrated weighted difference between the rates of increase of $\hat{\Lambda}_n$ and $\hat{\Lambda}_{n_l}$ over the observation times and the statistic $V_n^{(l)}$ has a similar meaning. For the selection of the weight process $W_n^{(l)}(t)$, a simple and natural choice is $W_n^{(1,l)}(t) = 1$, $l = 1, \ldots, k$. Another natural choice is $W_n^{(2,l)}(t) = Y_n(t) = \sum_{i=1}^n I(t \leq T_{K_i,K_i}) / n$, $l = 1, \ldots, k$, in which case weights are proportional to the number of subjects under observation. Based on groups, one may choose the weight process $W_n^{(l)}(t)$ as

$$
W_n^{(3,l)}(t) = Y_{n_l}(t) \quad \text{or} \quad \frac{Y_{n_l}(t)}{Y_n(t)} \quad \text{or} \quad \frac{Y_{n_1}(t) Y_{n_l}(t)}{Y_n(t)},
$$

where $Y_{n_l}(t)$ $(l = 1, \ldots, k)$ are defined as $Y_n(t)$ with the summation being only over subjects in the lth group. Some weight processes similar to $W_n^{(3)}$ have been used when recurrent event data are observed; see Andersen et al. (1993). In addition, $\sum_{i=1}^n I(t > T_{K_i,K_i}) / n$ is also chosen as another weight process by Zhang (2006). Some other possible choices are:

$$
1 - Y_{n_l}(t), \quad \frac{1 - Y_{n_l}(t)}{1 - Y_n(t)}, \quad \frac{(1 - Y_{n_1}(t))(1 - Y_{n_l}(t))}{1 - Y_n(t)}.
$$

Now, we state the asymptotic distribution of $\mathbf{U}_n = (U_n^{(1)}, \ldots, U_n^{(k)})^T$ and $\mathbf{V}_n = (V_n^{(2)}, \ldots, V_n^{(k)})^T$.

Theorem 4.3.1 *Suppose that Conditions A, B, C, D and E hold. Further, suppose that $W_n^{(l)}(t)$'s are bounded weight processes and that there exists a bounded function*

$W(t)$ such that $W \circ \Lambda_0^{-1}$ is a bounded Lipschitz function, and

$$\left[\int_0^\tau \{W_n^{(l)}(t) - W(t)\}^2 \, d\mu_1(t)\right]^{1/2} = o_p(n^{-1/6}) \ , \ l = 1, \ldots, k. \tag{4.10}$$

Also suppose that $n_l/n \to p_l$ as $n \to \infty$, where $0 < p_l < 1$, $l = 1, \ldots, k$, and $p_1 + \cdots + p_k = 1$. Then, under $H_0 : \Lambda_1 = \cdots = \Lambda_k = \Lambda_0$,

(i) $\mathbf{U}_n$ *has an asymptotic normal distribution with mean vector* $\mathbf{0}$ *and covariance matrix*

$$\mathbf{\Sigma}_{\mathbf{U}_w} = \mathbf{\Gamma}\,\mathbf{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2)\,\mathbf{\Gamma}', \tag{4.11}$$

where

$$\mathbf{\Gamma} = \begin{pmatrix} \sqrt{p_1} - \sqrt{\frac{1}{p_1}} & \sqrt{p_2} & \cdots & \sqrt{p_k} \\ \sqrt{p_1} & \sqrt{p_2} - \sqrt{\frac{1}{p_2}} & \cdots & \sqrt{p_k} \\ \cdots & \cdots & \cdots & \cdots \\ \sqrt{p_1} & \sqrt{p_2} & \cdots & \sqrt{p_k} - \sqrt{\frac{1}{p_k}} \end{pmatrix}$$

and $\sigma_1^2 = \cdots = \sigma_k^2 = \sigma_w^2$ given in (4.6).

(ii) $\mathbf{V}_n$ *has an asymptotic normal distribution with mean vector* $\mathbf{0}$ *and covariance matrix*

$$\mathbf{\Sigma}_{\mathbf{V}_w} = \mathbf{H}\,\mathbf{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2)\,\mathbf{H}', \tag{4.12}$$

where

$$\mathbf{H} = \begin{pmatrix} -\sqrt{\frac{1}{p_1}} & \sqrt{\frac{1}{p_2}} & 0 & \cdots & 0 \\ -\sqrt{\frac{1}{p_1}} & 0 & \sqrt{\frac{1}{p_3}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\sqrt{\frac{1}{p_1}} & 0 & 0 & \cdots & \sqrt{\frac{1}{p_k}} \end{pmatrix}$$

and σ_l^2 is as given in (i).

(iii) In addition, if

$$\max_{1\le i\le n} E\left[\sum_{j=1}^{K_i}\left\{W_n^{(l)}(T_{K_i,j})-W(T_{K_i,j})\right\}^2\right]\longrightarrow 0 \tag{4.13}$$

for $l=1,\ldots,k$, then $\mathbf{\Sigma}_{\mathbf{U}_w}$ and $\mathbf{\Sigma}_{\mathbf{V}_w}$ can be consistently estimated by

$$\hat{\mathbf{\Sigma}}_{U_n}=\mathbf{\Gamma}_n\,\mathbf{diag}(\hat\sigma_1^2,\hat\sigma_2^2,\ldots,\hat\sigma_k^2)\,\mathbf{\Gamma}_n', \tag{4.14}$$

and

$$\hat{\mathbf{\Sigma}}_{V_n}=\mathbf{H}_n\,\mathbf{diag}(\hat\sigma_1^2,\hat\sigma_2^2,\cdots,\hat\sigma_k^2)\,\mathbf{H}_n', \tag{4.15}$$

where

$$\mathbf{\Gamma}_n=\begin{pmatrix}\sqrt{\frac{n_1}{n}}-\sqrt{\frac{n}{n_1}} & \sqrt{\frac{n_2}{n}} & \cdots & \sqrt{\frac{n_k}{n}}\\ \sqrt{\frac{n_1}{n}} & \sqrt{\frac{n_2}{n}}-\sqrt{\frac{n}{n_2}} & \cdots & \sqrt{\frac{n_k}{n}}\\ \cdots & \cdots & \cdots & \cdots\\ \sqrt{\frac{n_1}{n}} & \sqrt{\frac{n_2}{n}} & \cdots & \sqrt{\frac{n_k}{n}}-\sqrt{\frac{n}{n_k}}\end{pmatrix},$$

$$\mathbf{H}_n=\begin{pmatrix}-\sqrt{\frac{n}{n_1}} & \sqrt{\frac{n}{n_2}} & 0 & \cdots & 0\\ -\sqrt{\frac{n}{n_1}} & 0 & \sqrt{\frac{n}{n_3}} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ -\sqrt{\frac{n}{n_1}} & 0 & 0 & \cdots & \sqrt{\frac{n}{n_k}}\end{pmatrix}$$

and

$$\begin{aligned}\hat\sigma_l^2 \;=\; \frac{1}{n}\sum_{i=1}^{n}\Bigg[&\sum_{j=1}^{K_i-1}W_n^{(l)}(T_{K_i,j})\hat\Lambda_n(T_{K_i,j})\\ &\times\left\{\frac{\Delta N_i(T_{K_i,j+1})}{\Delta\hat\Lambda_n(T_{K_i,j+1})}-\frac{\Delta N_i(T_{K_i,j})}{\Delta\hat\Lambda_n(T_{K_i,j})}\right\}\\ &+W_n^{(l)}(T_{K_i,K_i})\hat\Lambda_n(T_{K_i,K_i})\left\{1-\frac{\Delta N_i(T_{K_i,K_i})}{\Delta\hat\Lambda_n(T_{K_i,K_i})}\right\}\Bigg]^2\end{aligned} \tag{4.16}$$

for $l=1,\ldots,k$.

Let $\mathbf{U}_0$ denote the first $(k-1)$ components of $\mathbf{U}_n$ and $\hat{\mathbf{\Sigma}}_0$ the matrix obtained by deleting the last row and column of $\hat{\mathbf{\Sigma}}_{U_n}$. Then, using Theorem 4.3.1, two tests can be carried out for testing H_0 by means of the statistic $\chi_0^2 = \mathbf{U}_0^T \hat{\mathbf{\Sigma}}_0^{-1} \mathbf{U}_0$ and $\mathbf{V}_n^T \hat{\mathbf{\Sigma}}_{V_n}^{-1} \mathbf{V}_n$, which have asymptotically a central χ^2-distribution with $(k-1)$ degrees of freedom. This can be seen readily from the proof of the theorem.

REMARK 1. If the weight process $W_n^{(l)}$ is symmetric about $X_1, \ldots, X_n$, then (4.13) is equivalent to

$$E\left[\sum_{j=1}^{K_1} \left\{W_n^{(l)}(T_{K_1,j}) - W(T_{K_1,j})\right\}^2\right] \longrightarrow 0.$$

REMARK 2. For selection of weight processes, Zhang (2006) required that $W_n(t)$, $W(t)$ and $W \circ \Lambda_0^{-1}$ are monotone. These monotonicity assumptions restrict availability of weight processes. For example, the weight process $\frac{Y_{n_1}(t)Y_{n_2}(t)}{Y_{n_1}(t)+Y_{n_2}(t)}$ is often used in survival analysis, but it is not monotone. In addition, the monotonicity assumption on the weight process is not appropriate for deriving optimal tests under alternatives. In the above theorem, we have removed these assumptions. Therefore, compared to those stated in Zhang (2006), more weight processes are available here. It can be easily shown that the weight processes mentioned earlier all satisfy the conditions required by the theorem.

4.4 Simulation Study

To examine the finite-sample properties of the proposed test statistics and compare them with those of the tests presented by Sun and Fang (2003), Park et al. (2007) and Zhang (2006), we carry out a simulation study for the two-sample comparison

problem. When $k = 2$, then the null hypothesis can be tested by $T_1 = U_n^{(1)}/\hat{\sigma}_U$ and $T_2 = V_n^{(2)}/\hat{\sigma}_V$ which have asymptotic standard normal distribution, where

$$\hat{\sigma}_U = \left\{ \left(\sqrt{\frac{n_1}{n}} - \sqrt{\frac{n}{n_1}} \right)^2 \hat{\sigma}_1^2 + \frac{n_2}{n} \hat{\sigma}_2^2 \right\}^{1/2},$$

$$\hat{\sigma}_V = \left\{ \frac{n}{n_1} \hat{\sigma}_1^2 + \frac{n}{n_2} \hat{\sigma}_2^2 \right\}^{1/2},$$

and $U_n^{(1)}$, $V_n^{(2)}$ and $\hat{\sigma}_l$ are as given in (4.8), (4.9) and (4.16), respectively. Let T_{SF}, T_{PSZ} and T_Z denote the tests presented by Sun and Fang (2003), Park et al. (2007) and Zhang (2006), respectively. Here, we focus on evaluating the performance of T_1 and T_2 and comparing them to those of T_{PSZ}, T_Z and T_{SF}. Note that $T_Z = T_{PSZ}$ for $k = 2$. To generate panel count data $\{k_i, t_{ij}, n_{ij}, j = 1, \ldots, k_i, i = 1, \ldots, n\}$, we mimic medical follow-up studies such as the examples discussed in the next section. We first generate the number of observation times k_i from the uniform distribution $U\{1, \ldots, 10\}$, and then, given k_i, we generate observation times t_{ij}'s from $U\{1, \ldots, 10\}$, for simplicity. To generate n_{ij}'s, we assume that N_i's are nonhomogeneous Poisson or mixed Poisson processes. In particular, let $\{\nu_i, i = 1, \ldots, n\}$ be independent and identically distributed random variables, and given ν_i, let $N_i(t)$ be a Poisson process with mean function $\Lambda_i(t|\nu_i) = E(N_i(t)|\nu_i)$. Let S_l denote the set of indices for subjects in group l, $l = 1, 2$. For the objective of the study, we consider two cases as follows:

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

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

Figures 4.1-4.2 display the graphs of the true mean functions for two cases with $\nu = 1$ and different values of β. It can be seen that the two mean functions do not

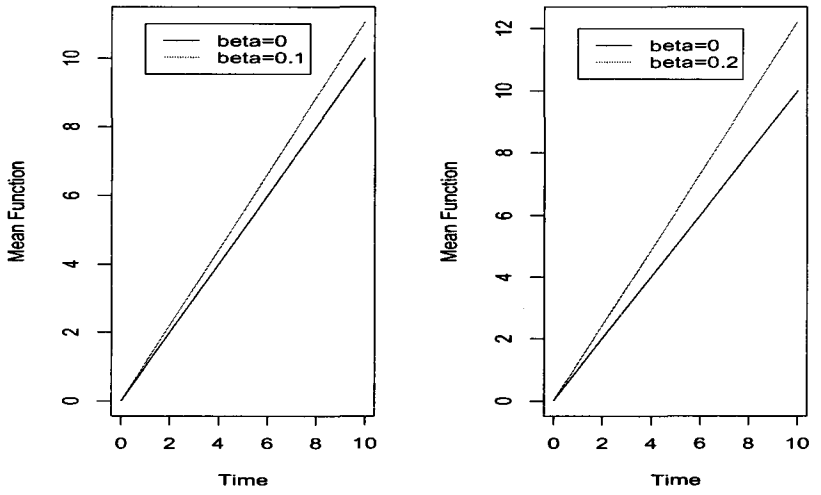

Figure 4.1: True mean functions for Case 1 with $\nu = 1$ and $\beta = 0.1, 0.2$.

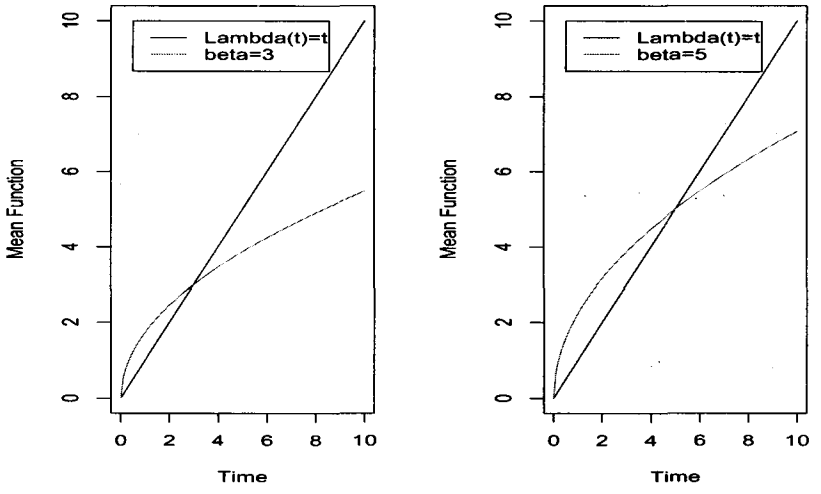

Figure 4.2: True mean functions for Case 2 with $\nu = 1$ and $\beta = 3, 5$.

overlap in Case 1 and they cross over in Case 2.

For each case, we consider $\nu_i = 1$ and $\nu_i \sim$ Gamma$(2, 1/2)$ corresponding to Poisson and mixed Poisson processes, respectively. For each setting, we consider two sample sizes, $n_1 = n_2 = 50$ and 100, respectively. As mentioned earlier in Section 4.3, we choose the four weight processes:

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

$$W_n^{(3)}(t) = \frac{Y_{n_1}(t)Y_{n_2}(t)}{Y_n(t)}, \text{ and } W_n^{(4)}(t) = 1 - Y_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(t > t_{k_i,k_i}).$$

The NPMLEs $\hat{\Lambda}_n$ and $\hat{\Lambda}_{n_l}$ are computed by using the modified iterative convex minorant algorithm (MICM); see Wellner and Zhang (2000). All the results reported here are based on 1000 Monte Carlo replications using R software.

Tables 4.1-4.4 present the estimated sizes and powers of the proposed test statistics T_1 and T_2 and those of the test statistics T_{PSZ}, T_Z and T_{SF} (Park et al., 2007; Zhang, 2006; Sun and Fang, 2003) at significance level $\alpha = 0.05$ for different values of β and the four weight processes based on the simulated data for the two cases with $\nu_i = 1$ and $\nu_i \sim$ Gamma$(2, 1/2)$, respectively. When $\nu_i = 1$, the $N_i(t)$'s are Poisson processes; when $\nu_i \sim$ Gamma$(2, 1/2)$, the $N_i(t)$'s are mixed Poisson processes. The first part of the table is for the situation with the total sample size of 100 and the second part is for the situation with the total sample size of 200. For Case 1 considered here, the proposed tests display good power properties and the powers are close for the four weight processes. As expected, the power increases when the sample size increases, and the power decreases in the presence of more variability. As seen in Tables 4.1 and 4.2, the proposed tests with $W_n^{(1)}(t)$ have the best power performance, and the proposed tests based on the NPMLE are more powerful than the tests based on NPMPLE when

more variability exists, as one would expect. For Case 2 considered here, the proposed tests also display good power properties, but the powers rely on choices of weight processes. As seen in Tables 4.3 and 4.4, the proposed tests with $W_n^{(4)}$ have the best power performance, and the proposed tests with appropriate weights based on NPMLE are much more powerful and more robust than those based on NPMPLE in this case. For example, when $\beta = 5, 8$ for mixed Poisson processes, the new tests with $W_n^{(4)}$ have good powers, but the tests T_{PSZ} & T_Z (Park et al., 2007; Zhang, 2006) with four weights and T_{SF} (Sun and Fang, 2003) have very poor powers. For all situations considered here, the performance of T_1 and T_2 are the same.

Note that the tests with different weights have different powers for Case 2. Let's explain why these results are reasonable. In this case, two true mean functions cross over at time $t = \beta$, the differences before this time point and after this time point have different signs, the difference after this point seems to dominate the difference before this point for the cases of $\beta = 3, 5$ and seems to be dominated by the difference before this point for the case of $\beta = 8$. When $\beta = 3, 5$, the tests with $W_n^{(1)}$ and $W_n^{(4)}$ have better powers than those with $W_n^{(2)}$ and $W_n^{(3)}$, and the test with $W_n^{(4)}$ has the largest power since it weights the difference at later times more than those with $W_n^{(1)}$, $W_n^{(2)}$ and $W_n^{(3)}$. In particular, when $\beta = 5$, the powers of the tests with $W_n^{(2)}$ and $W_n^{(3)}$ are very poor. This is because the small difference with large weights before this point and the large difference with small weights after this point seem to cancel each other. When $\beta = 8$, the tests with $W_n^{(2)}$, $W_n^{(3)}$ and $W_n^{(4)}$ perform better than the test with $W_n^{(1)}$. When $\beta = 8$, the biggest difference between two mean functions occurs at earlier time so that the tests with $W_n^{(2)}$ and $W_n^{(3)}$ have reasonable powers. But the test with $W_n^{(1)} = 1$ has a poor power though the difference at earlier times seems to dominate

Table 4.1: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for Poisson processes in Case 1

β	T_2				T_{PSZ} & T_Z				T_{SF}
	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	
				$n_1 = n_2 = 50$					
0.0	0.060	0.058	0.058	0.056	0.063	0.061	0.061	0.063	0.061
0.1	0.298	0.210	0.209	0.194	0.214	0.200	0.200	0.204	0.207
0.2	0.858	0.747	0.748	0.790	0.697	0.667	0.665	0.695	0.693
0.3	1.000	0.987	0.983	0.986	0.981	0.974	0.974	0.968	0.979
				$n_1 = n_2 = 100$					
0.0	0.047	0.047	0.047	0.049	0.044	0.046	0.046	0.045	0.043
0.1	0.542	0.472	0.471	0.489	0.423	0.405	0.405	0.411	0.422
0.2	0.993	0.967	0.964	0.991	0.958	0.948	0.947	0.952	0.950
0.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	T_1				T_1				
	$n_1 = n_2 = 50$				$n_1 = n_2 = 100$				
0.0	0.052	0.051	0.051	0.053	0.051	0.049	0.049	0.050	
0.1	0.340	0.218	0.218	0.220	0.548	0.479	0.474	0.492	
0.2	0.868	0.787	0.764	0.798	0.996	0.976	0.974	0.993	
0.3	1.000	0.997	0.995	0.991	1.000	1.000	1.000	1.000	

Table 4.2: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for mixed Poisson processes in Case 1

β	T_2				T_{PSZ} & T_Z				T_{SF}
	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	
				$n_1 = n_2 = 50$					
0.0	0.043	0.040	0.042	0.045	0.037	0.040	0.040	0.042	0.035
0.1	0.100	0.097	0.097	0.099	0.084	0.085	0.085	0.086	0.083
0.2	0.221	0.205	0.207	0.204	0.185	0.184	0.184	0.180	0.183
0.3	0.458	0.407	0.408	0.415	0.380	0.375	0.375	0.379	0.370
				$n_1 = n_2 = 100$					
0.0	0.043	0.041	0.041	0.046	0.048	0.045	0.045	0.044	0.046
0.1	0.140	0.125	0.125	0.138	0.114	0.111	0.111	0.112	0.111
0.2	0.410	0.364	0.362	0.368	0.317	0.307	0.307	0.314	0.316
0.3	0.708	0.663	0.662	0.672	0.596	0.592	0.592	0.593	0.590
	T_1				T_1				
	$n_1 = n_2 = 50$				$n_1 = n_2 = 100$				
0.0	0.054	0.048	0.048	0.046	0.048	0.047	0.047	0.051	
0.1	0.108	0.102	0.102	0.100	0.142	0.126	0.123	0.137	
0.2	0.216	0.205	0.206	0.207	0.412	0.388	0.390	0.391	
0.3	0.474	0.404	0.402	0.437	0.710	0.672	0.670	0.671	

Table 4.3: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for Poisson processes in Case 2

β	T_2				T_{PSZ} & T_Z				T_{SF}
	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	
				$n_1 = n_2 = 50$					
3	1.000	0.787	0.766	1.000	0.956	0.900	0.899	1.000	0.955
5	0.969	0.080	0.077	1.000	0.189	0.113	0.111	0.880	0.188
8	0.127	0.674	0.688	0.993	0.403	0.559	0.562	0.069	0.400
				$n_1 = n_2 = 100$					
3	1.000	0.964	0.960	1.000	0.999	0.993	0.993	1.000	0.999
5	1.000	0.078	0.079	1.000	0.290	0.140	0.139	0.988	0.284
8	0.222	0.935	0.939	1.000	0.670	0.843	0.846	0.082	0.667
	T_1				T_1				
	$n_1 = n_2 = 50$				$n_1 = n_2 = 100$				
3	1.000	0.784	0.769	1.000	1.000	0.958	0.955	1.000	
5	0.969	0.088	0.086	1.000	1.000	0.084	0.083	1.000	
8	0.130	0.675	0.689	0.995	0.232	0.932	0.935	1.000	

Table 4.4: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for mixed Poisson processes in Case 2

β	T_2				T_{PSZ} & T_Z				T_{SF}
	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(4)}$	
				$n_1 = n_2 = 50$					
3	0.858	0.301	0.294	0.992	0.386	0.321	0.318	0.708	0.380
5	0.424	0.078	0.078	0.943	0.089	0.071	0.071	0.289	0.086
8	0.062	0.255	0.263	0.771	0.117	0.158	0.158	0.039	0.111
				$n_1 = n_2 = 100$					
3	0.992	0.534	0.530	1.000	0.695	0.594	0.594	0.949	0.691
5	0.677	0.071	0.072	1.000	0.100	0.065	0.065	0.473	0.095
8	0.096	0.434	0.437	0.961	0.185	0.280	0.280	0.067	0.182
	T_1				T_1				
	$n_1 = n_2 = 50$				$n_1 = n_2 = 100$				
3	0.858	0.299	0.289	0.991	0.993	0.533	0.529	1.000	
5	0.396	0.074	0.074	0.942	0.685	0.068	0.066	1.000	
8	0.063	0.259	0.268	0.771	0.094	0.432	0.438	0.960	

the difference at later times. This can be understood from the expressions of the test statistics $\mathbf{U}_n$ and $\mathbf{V}_n$ where the differences with different signs multiplied by the value of the mean function may cancel each other since the mean function takes small values at earlier times and large values at later times. When $\beta = 8$, the test with $W_n^{(4)}$ still perform well. This is because it puts zero weight at earlier times and heavily weight at later times. Similar situations happened in real examples considered in the next section.

To evaluate the asymptotic result given in Theorem 4.3.1, the quantile plots of the test statistic T_2 against the standard normal distribution are constructed. Figures 4.3 and 4.4 present the plots for the cases with $W_n(t) = W_n^{(1)}(t)$ and $n = 100$ and $n = 200$, respectively, and they clearly reveal that the asymptotic approximation is very good. Similar plots were obtained for test statistic T_1 and other situations as well.

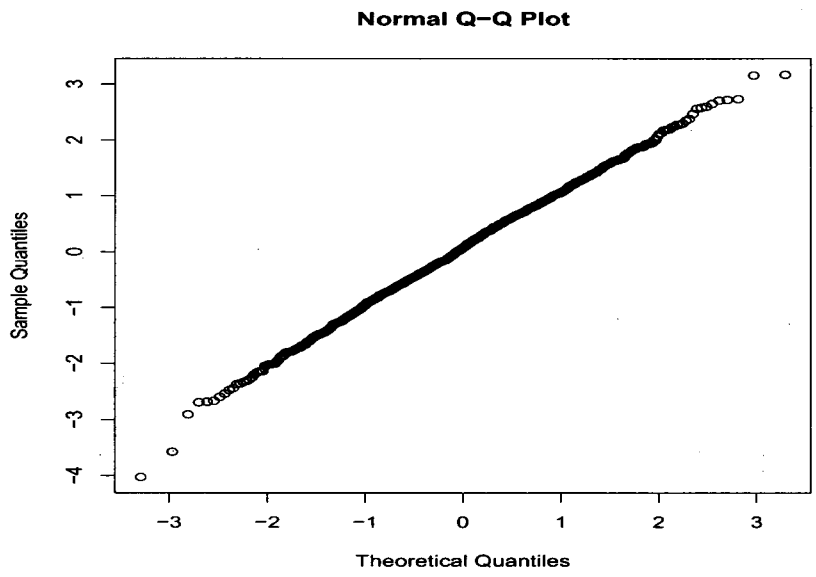

Figure 4.3: Simulation study. Normal quantile plot for T_1 ($n = 100$).

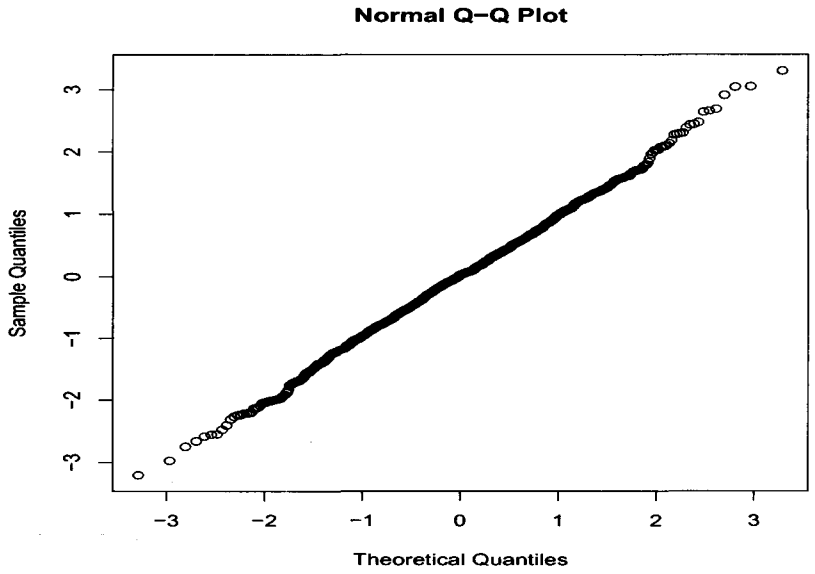

Figure 4.4: Simulation study. Normal quantile plot for T_2 ($n = 200$).

In the above simulation study, we did examine all four weight processes suggested earlier in Section 4.3, and in Case 1, the weight process $W_n^{(1)}$ yielded slightly higher power than the other three weight processes, and in Case 2, the weight process $W_n^{(4)}$ yielded the largest power. These simulation results suggest that, when the mean functions do not cross over, the test with the equal weight has a good power; otherwise, the test with the unequal and appropriate weight will also have a good power. In general, one can choose the weight process based on the behavior of the NPMLEs of the mean functions to improve power since the true mean functions are unknown. When the difference of mean functions at earlier times dominate the difference at later times, the tests with $W_n^{(2)}$ and $W_n^{(3)}$ tend to have good powers; when the difference of mean functions at later times dominate the difference at earlier times, the test with $W_n^{(4)}$ tends to have a good power. In addition to the four processes considered here, some other weight processes can be found in Andersen et al. (1993), which discusses non-

parametric treatment comparison based on recurrent event data. It would, therefore, be of great interest to investigate the problem of the selection of a weight process based on data.

The new tests based on the NPMLE are more powerful and more robust than the existing tests based on the NPMPLE. One possible reason is that the NPMLE is more efficient than the NPMPLE. The main drawback of the NPMPLE is that the dependence of events within a subject is ignored. Another reason is that the structure of new test statistic is more reasonable since it is based on the characteristic of the NPMLE.

4.5 Illustrative Examples

To illustrate the proposed method, we consider here two examples: a floating gallstones study and a bladder tumor study.

4.5.1 A Floating Gallstones Study

Thall and Lachin (1988) described a follow-up study on patients with floating gallstones. The data consist of the first year follow-up of the patients in two study groups, placebo (48) and high-dose chenodiol (65), from the National Cooperative Gallstone Study. The observed data include the successive visit-times in study weeks and the associated counts of episodes of nausea for patients in different treatment groups; see Table 1 of Thall and Lachin (1988). The whole study consists of 916 patients who were randomized to placebo, low dose, or high dose group and followed for up to two years.

During the study, patients were scheduled to return for clinical visits at 1, 2, 3, 6, 9, and 12 months. In reality, most of the patients visited about six times within the first year, but actual visit times differ from patient to patient. Some patients had only one visit and some had 9 visits. As pointed out by Thall and Lachin (1988), there is no evidence that the number of observations and actual observation times are related to the incidence of nausea, and so it seems reasonable to assume that conditions required for the asymptotic results hold in this case. The problem of interest here is to compare the two treatment groups in terms of the incidence rates of nausea.

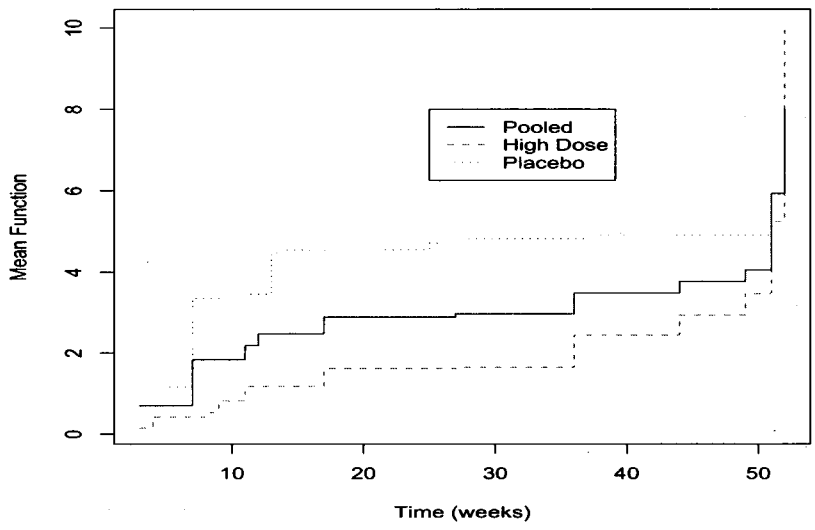

Figure 4.5: Floating gallstone study. Estimates of the mean functions.

To test the difference between the two groups, we treated the placebo group as group 1 ($\Lambda_1(t)$) and the high-dose chenodiol group as group 2 ($\Lambda_2(t)$) and applied the proposed method to the data from 113 gallstone patients in the two groups to test the null hypothesis $H_0 : \Lambda_1(t) = \Lambda_2(t)$. The nonparametric maximum likelihood estimators of the incidence rates of nausea and the increments of the estimators are

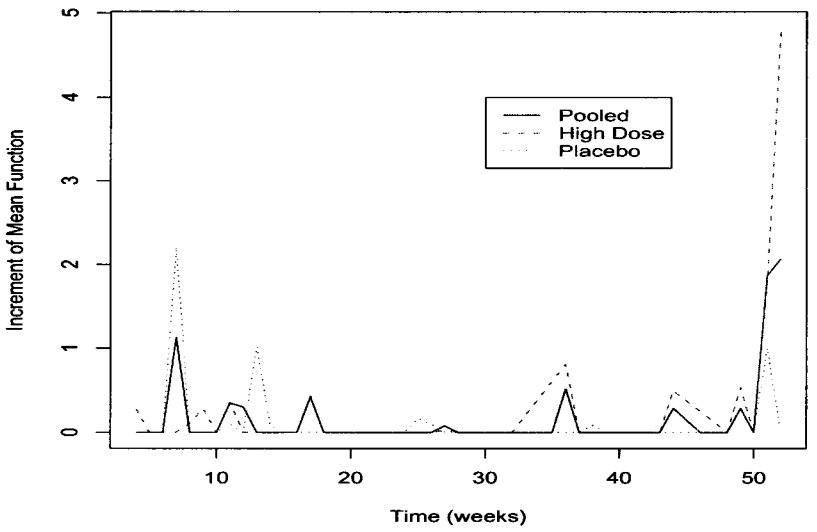

Figure 4.6: Floating gallstone study. Increments of the estimated mean functions.

shown in Figures 4.5 and 4.6. We obtained $T_1 = 0.175$ and $T_2 = 0.206$ with $W_n(t) = W_n^{(1)}(t)$, giving p-values of 0.861 and 0.837 based on the standard normal distribution, and $T_1 = -337.221$, -494.571 and 241.159 and $T_2 = -193.238$, -283.739 and 138.311 with $W_n(t) = W_n^{(2)}(t)$, $W_n^{(3)}(t)$ and $1 - W_n^{(2)}(t)$ which correspond to p-values $\ll 0.0001$. The proposed tests with appropriate weights suggest that the incidence rates of nausea were significantly different for the patients in the two groups and this agrees with the results given in Thall and Lachin (1988); the proposed unweighted test fails to reject H_0. This can be easily understood by looking at the behavior of increments of the estimators. From Figure 4.6, we can see clearly that the increment of the mean event rate in the placebo group is higher than that in the high dose group at earlier times and in contrast, the increment of the mean event rate in the high dose group is higher than that in the placebo group at later times in the year. So, the test with equal weights couldn't detect the difference between two groups. In comparison, the use of

the approach in Sun and Fang (2003) gave a p-value of 0.1428, Park et al. (2007) gave p-values: 0.454, 0.417 and 0.413 with three weights, respectively, and the tests presented by Zhang (2006) would give the same result as above. Thus, none of the existing tests based on NPMPLE can detect the difference of two treatments, and the proposed tests with suitable weights have detected successfully that, as we expected. One possible reason for this is that the nonparametric maximum likelihood estimator is more efficient than the nonparametric pseudo-likelihood estimator.

4.5.2 A Bladder Tumor Study

We consider a bladder tumor study conducted by the Veterans Administration Cooperative Urological Research Group (VACURG), and the data are presented in Andrews and Herzberg (1985). For some earlier analyses of these data, one may refer to Byar, Blackard and The VACURG (1977), Byar (1980), Wellner and Zhang (2000), Sun and Wei (2000), and Zhang (2002, 2006). The data were obtained from a randomized clinical trial. All patients had superficial bladder tumors when they entered the trial, and they were assigned randomly to one of three treatments: placebo, thiotepa and pyridoxine. At subsequent follow-up visits, any tumors noticed were removed and treatment was continued. The study included 116 patients, of which there were 47 in placebo group, 38 in thiotepa group and 31 in pyridoxine. We can get a set of panel count data $\{k_i, t_{ij}, n_{ij}, j = 1, \ldots, k_i, i = 1, \ldots, n\}$ where for the ith patient, k_i is the number of visits, t_{ij}'s are all visit times, and n_{ij} is total number of tumors until t_{ij} $(j = 1, \ldots, k_i)$. The objective of the study is to determine the effect of treatment on the frequency of tumor recurrence.

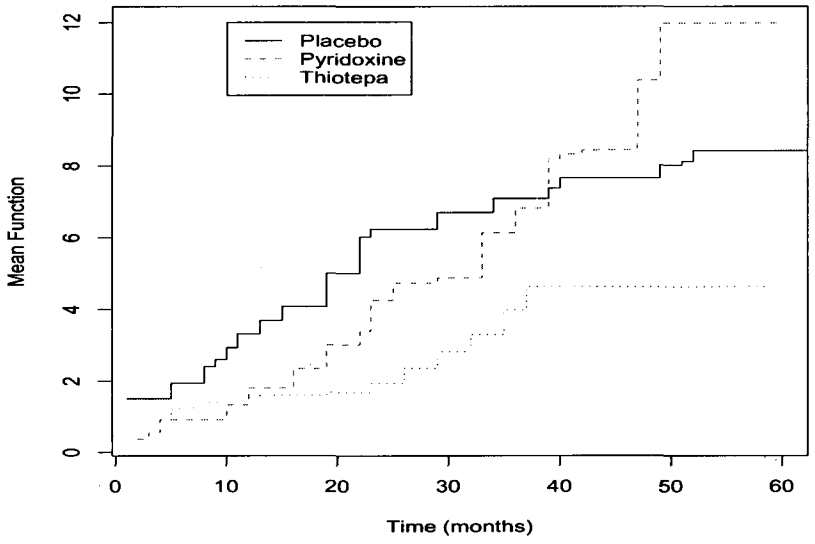

Figure 4.7: Bladder tumor study. Estimates of the mean functions.

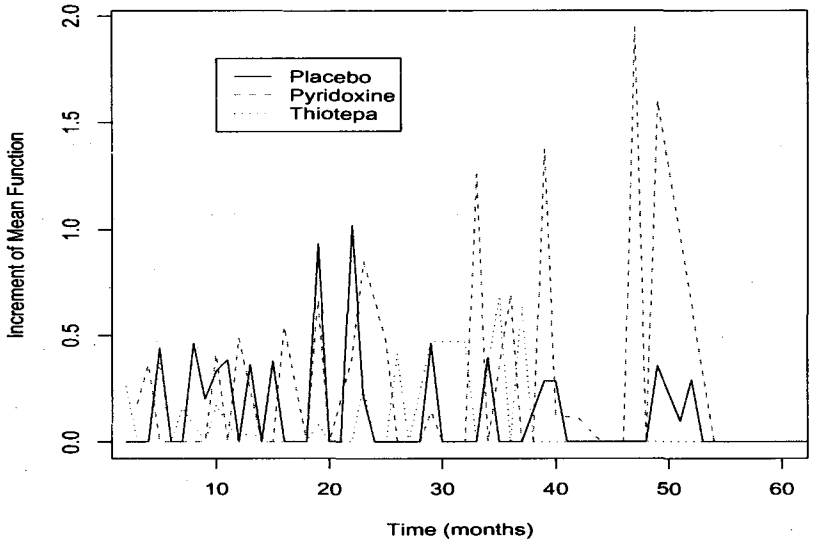

Figure 4.8: Bladder tumor study. Increments of the estimated mean functions.

Let $\Lambda_1(t), \Lambda_2(t)$ and $\Lambda_3(t)$ be the mean functions corresponding to the three treatment groups: placebo, thiotepa and pyridoxine, respectively. The nonparametric maximum likelihood estimators of mean functions and their increments from the three groups are presented in Figures 4.7 and 4.8, respectively. We observe from Figure 7 that the difference of the three groups becomes larger when the time increases. To test the null hypothesis $H_0 : \Lambda_1(t) = \Lambda_2(t) = \Lambda_3(t)$, we applied the proposed method to this panel count data. We obtained $\chi_0^2 = 3.617, 3.269$ and p-value $= 0.164, 0.195$ with $W_n(t) = 1$, $\chi_0^2 = 1196123, 300179.2$ and p-values $< 10^{-8}$ with $W_n(t) = Y_n(t)$, and $\chi_0^2 = 489000.4, 121908.1$ and p-values $< 10^{-8}$ with $W_n(t) = 1 - Y_n(t)$, based on the asymptotic distributions for test statistics $\mathbf{U}_n$ and $\mathbf{V}_n$ given in Theorem 3.1, respectively. The proposed tests having weights suggest that the frequency of tumor recurrence are significantly different for the patients in the three groups at 0.01 level of significance, while the proposed unweighted test fails to detect the difference. This can also be understood from the behavior of the increments of the estimated mean functions shown in Figure 4.8. Incidentally, through a regression analysis of the data from two treatments, placebo and thiotepa, Sun and Wei (2000) and Zhang (2002) concluded that thiotepa effectively reduces the recurrence of tumors. However, the existing test procedures (Sun and Fang, 2003; Zhang, 2006) based on NPMPLE fail to reject the null hypothesis at level 0.05.

These examples illustrate that different weights may result in different conclusions, and the tests with appropriate weight process could lead to better power of the test. Therefore, the selection of a suitable weight process would be important for detecting difference between groups.

4.6 Proofs

In this section we present the proofs of Theorems 4.2.1 and 4.3.1.

4.6.1 Proof of Theorem 4.2.1

We begin with some preliminary results. For convenience, let us first recall some notation given in Wellner and Zhang (2000). Set

$$\mathcal{F} = \{\Lambda : [0, \tau] \to [0, \infty) | \Lambda \text{ is nondecreasing, } \Lambda(0) = 0\}.$$

Let $t_1 < t_2 < \cdots < t_m$ denote the ordered distinct observation time points in the set of all observation time points $\{T_{K_i,j}, j = 1, \ldots, K_i, i = 1, \ldots, n\}$. Also let $\Omega = \{\mathbf{u} = (u_1, u_2, \ldots, u_m) : 0 \leq u_1 \leq \cdots \leq u_m < \infty\}$ and the map $\mathcal{A} : \mathcal{F} \to \Omega$ be defined by

$$\mathbf{u} = \mathcal{A}(\Lambda) = (\Lambda(t_1), \Lambda(t_2), \ldots, \Lambda(t_m)) \qquad \text{for all } \Lambda \in \mathcal{F}.$$

We also define a rank function R: $\{T_{K_i,j} : j = 1, 2, \ldots, K_i; i = 1, 2, \ldots, n\} \to \{1, 2, \ldots, m\}$ such that

$$R(T_{K_i,j}) = s \quad \text{if} \quad T_{K_i,j} = t_s.$$

Then, the log-likelihood function can be rewritten as

$$\begin{aligned} \phi(\mathbf{u}|\mathbf{X}) \;=\; \sum_{i=1}^{n} \Bigg[\sum_{j=1}^{K_i} \{N_i(T_{K_i,j}) - N_i(T_{K_i,j-1})\} \\ \times \log \left\{ u_{R(T_{K_i,j})} - u_{R(T_{K_i,j-1})} \right\} - u_{R(T_{K_i,K_i})} \Bigg] \end{aligned}$$

and the NPMLE $\hat{\Lambda}_n$ of Λ_0 is then given by

$$(\hat{\Lambda}_n(t_1), \hat{\Lambda}(t_2), \ldots, \hat{\Lambda}_n(t_m)) = \hat{\mathbf{u}}_n = \text{argmax}_{\mathbf{u} \in \Omega} \phi(\mathbf{u}|\mathbf{X}).$$

Set

$$\phi_\ell(\mathbf{u}) = \frac{\partial \phi(\mathbf{u}|\mathbf{X})}{\partial u_\ell} = \sum_{i=1}^{n} \phi_{i,\ell}(\mathbf{u}) \quad \text{for } \ell = 1, 2, ..., m,$$

where

$$\begin{aligned}\phi_{i,\ell}(\mathbf{u}) \quad = \quad & \sum_{j=1}^{K_i-1} \left\{ \frac{N_i(T_{K_i,j}) - N_i(T_{K_i,j-1})}{u_{R(T_{K_i,j})} - u_{R(T_{K_i,j-1})}} \right. \\ & \left. - \frac{N_i(T_{K_i,j+1}) - N_i(T_{K_i,j})}{u_{R(T_{K_i,j+1})} - u_{R(T_{K_i,j})}} \right\} 1_{\{T_{K_i,j}=t_\ell\}} \\ & + \left\{ \frac{N_i(T_{K_i,K_i}) - N_i(T_{K_i,K_i-1})}{u_{R(T_{K_i,K_i})} - u_{R(T_{K_i,K_i-1})}} - 1 \right\} 1_{\{T_{K_i,K_i}=t_\ell\}}.\end{aligned}$$

Lemma 4.6.1 *Let φ be any real function. Then*

$$\sum_{\ell=1}^{m} \varphi(\hat{u}_\ell) \left\{ \sum_{i=1}^{n} \phi_{i,\ell}(\hat{\mathbf{u}}) \right\} = 0. \tag{4.17}$$

Proof. Let $\alpha_j = \hat{\Lambda}_n(t_j) - \hat{\Lambda}_n(t_{j-1})$, $j = 1, \ldots, m$. Using arguments similar to Proposition 2.1 of Groenebom (1996), we have

$$\sum_{j=i}^{m} \frac{\partial \phi(\hat{\mathbf{u}})}{\partial u_j} = 0 \quad \text{if } \alpha_i > 0 \text{ or } i = 1.$$

Let $t_{k_1} < t_{k_2} < \cdots < t_{k_p}$ be jump points of $\hat{\Lambda}_n$. Then

$$\sum_{\ell=k_j}^{m} \sum_{i=1}^{n} \phi_{i,\ell}(\hat{\mathbf{u}}) = 0, \quad j = 1, \ldots, p,$$

and so

$$\sum_{k_j \le \ell < k_{j+1}} \sum_{i=1}^{n} \phi_{i,\ell}(\hat{\mathbf{u}}) = 0, \quad j = 1, \ldots, p-1.$$

Thus

$$\sum_{k_j \le \ell < k_{j+1}} \varphi(\hat{u}_\ell) \sum_{i=1}^{n} \phi_{i,\ell}(\hat{\mathbf{u}}) = 0, \quad j = 1, \cdots, p-1,$$

since $\hat{u}_\ell = \hat{\Lambda}(t_\ell)$ is a constant for $k_j \le \ell < k_{j+1}$. Therefore, we conclude that

$$\sum_{\ell=1}^{m} \varphi(\hat{u}_\ell) \sum_{i=1}^{n} \phi_{i,\ell}(\hat{\mathbf{u}}) = 0.$$

Hence, the lemma follows.

Now, let μ_i be as defined in Section 4.2, and let d_i be the $L_2(\mu_i)$ metric on $\mathcal{F}$, $i = 1, 2$. Then for $\Lambda_1, \Lambda_2 \in \mathcal{F}$,

$$\begin{aligned} d_1^2(\Lambda_1, \Lambda_2) &= \int |\Lambda_1(t) - \Lambda_2(t)|^2 d\mu_1(t) \\ &= E\left[\sum_{j=1}^{K} \{\Lambda_1(T_{K,j}) - \Lambda_2(T_{K,j})\}^2\right] \end{aligned} \tag{4.18}$$

and

$$\begin{aligned} & d_2^2(\Lambda_1, \Lambda_2) \\ &= \int\int |(\Lambda_1(s) - \Lambda_1(t)) - (\Lambda_2(s) - \Lambda_2(t))|^2 d\mu_2(s,t) \\ &= E\left[\sum_{j=1}^{K} \{(\Lambda_1(T_{K,j}) - \Lambda_1(T_{K,j-1})) - (\Lambda_2(T_{K,j}) - \Lambda_2(T_{K,j-1}))\}^2\right]. \end{aligned} \tag{4.19}$$

If $P(K \le K_0) = 1$ for some constant K_0, then we have

$$\frac{1}{2} d_2(\Lambda_1, \Lambda_2) \le d_1(\Lambda_1, \Lambda_2) \le K_0 d_2(\Lambda_1, \Lambda_2). \tag{4.20}$$

Wellner and Zhang (2000) showed that

$$d_1(\hat{\Lambda}_n, \Lambda_0) \xrightarrow{a.s.} 0 \tag{4.21}$$

and hence that the uniform consistency of $\hat{\Lambda}_n$ can be shown by using arguments similar to Proposition 5 of Schick and Yu (2000) under Conditions A, B, D and E; that is,

$$\sup_{t \in [\tau_0, \tau]} |\hat{\Lambda}_n(t) - \Lambda_0(t)| \xrightarrow{a.s.} 0. \tag{4.22}$$

Note that the uniform consistency of $\hat{\Lambda}_n$ implies for every $0 < \delta_0 < \min\{L_0/2, \Lambda_0(\tau_0)\}$ and any $\varepsilon > 0$, there exists a positive integer N_ε such that

$$\sup_{n>N_\varepsilon} P\left\{\sup_{t\in[\tau_0,\tau]} |\hat{\Lambda}_n(t) - \Lambda_0(t)| > \delta_0\right\} < \varepsilon. \tag{4.23}$$

Here, we fix δ_0. Let

$$\mathcal{F}_0 = \{\Lambda : \Lambda \in \mathcal{F}, \sup_{t\in[\tau_0,\tau]} |\Lambda(t) - \Lambda_0(t)| \leq \delta_0\}. \tag{4.24}$$

Define $\hat{\Lambda}_n^*$ as

$$\hat{\Lambda}_n^* = \text{argmax}_{\Lambda\in\Omega\cap\mathcal{F}_0}\left\{\sum_{i=1}^{n}\sum_{j=1}^{K_i}(\Delta N_i(T_{K_i,j})\log(\Delta\Lambda(T_{K_i,j})) - \Delta\Lambda(T_{K_i,j}))\right\},$$

where Ω is the class of nondecreasing step functions with possible jumps only at the observation time points $\{T_{K_i,j},\ j = 1, ..., K_i,\ i = 1, ..., n\}$. Clearly, we have

$$\sup_{n>N_\varepsilon} P(\hat{\Lambda}_n \neq \hat{\Lambda}_n^*) \leq \sup_{n>N_\varepsilon} P\left\{\sup_{t\in[\tau_0,\tau]} |\hat{\Lambda}_n(t) - \Lambda_0(t)| > \delta_0\right\} < \varepsilon. \tag{4.25}$$

Lemma 4.6.2 *We have* $d_1(\hat{\Lambda}_n^*, \Lambda_0) = O_p(n^{-1/3})$.

Proof. To establish the rate of convergence for $\hat{\Lambda}_n^*$, we shall apply Theorem 3.2.5 of Van der Vaart and Wellner (1996). Define

$$\begin{aligned} m_\Lambda(X) &= \sum_{j=1}^{K}[(N(T_{K,j}) - N(T_{K,j-1}))\log\{\Lambda(T_{K,j}) - \Lambda(T_{K,j-1})\} \\ &\quad -\{\Lambda(T_{K,j}) - \Lambda(T_{K,j-1})\}] \end{aligned} \tag{4.26}$$

and

$$\mathbb{M}(\Lambda) = Pm_\Lambda(X). \tag{4.27}$$

Let $h(x) = x(\log(x) - 1) + 1$. Then, $h(x) \geq \frac{1}{5}(x-1)^2$ for x in a neighbourhood of $x = 1$. Thus, in a neighbourhood of Λ_0,

$$
\begin{aligned}
&\mathbb{M}(\Lambda_0) - \mathbb{M}(\Lambda) \\
&= \mathrm{P}\left[\sum_{j=1}^{K}\{(\Lambda_0(T_{K,j}) - \Lambda_0(T_{K,j-1}))\log\left(\frac{\Lambda_0(T_{K,j}) - \Lambda_0(T_{K,j-1})}{\Lambda(T_{K,j}) - \Lambda(T_{K,j-1})}\right)\right. \\
&\qquad \left. -(\Lambda_0(T_{K,j}) - \Lambda_0(T_{K,j-1})) + (\Lambda(T_{K,j}) - \Lambda(T_{K,j-1}))\}\right] \\
&= \mathrm{P}\left[\sum_{j=1}^{K}\{\Lambda(T_{K,j}) - \Lambda(T_{K,j-1})\}h\left(\frac{\Lambda_0(T_{K,j}) - \Lambda_0(T_{K,j-1})}{\Lambda(T_{K,j}) - \Lambda(T_{K,j-1})}\right)\right] \\
&= \int\int\{\Lambda(u) - \Lambda(v)\}h\left(\frac{\Lambda_0(u) - \Lambda_0(v)}{\Lambda(u) - \Lambda(v)}\right)d\mu_2(u,v) \\
&\geq \frac{1}{5}\int\int\{\Lambda(u) - \Lambda(v)\}\left\{\frac{\Lambda_0(u) - \Lambda_0(v)}{\Lambda(u) - \Lambda(v)} - 1\right\}^2 d\mu_2(u,v) \\
&= \frac{1}{5}\int\int\frac{\{(\Lambda_0(u) - \Lambda(u)) - (\Lambda_0(v) - \Lambda(v))\}^2}{\Lambda(u) - \Lambda(v)}d\mu_2(u,v) \\
&\geq c_1\, d_1^2(\Lambda, \Lambda_0)
\end{aligned}
$$

for some constant c_1, and hence the separation condition of the theorem is satisfied. Also, let

$$\mathcal{F}_\delta = \{\Lambda : d_1(\Lambda, \Lambda_0) \leq \delta, \Lambda \in \mathcal{F}_0\} \quad (\delta > 0) \tag{4.28}$$

and

$$\mathcal{M}_\delta = \{m_\Lambda(X) - m_{\Lambda_0}(X) : \Lambda \in \mathcal{F}_\delta\}. \tag{4.29}$$

Note that $\mathcal{F}_\delta$ is a class of monotone nondecreasing functions. Then, it follows from Theorem 2.7.5 of Van der Vaart and Wellner (1996) that for any $\eta > 0$, there exists a set of brackets $\{[\Lambda_i^L, \Lambda_i^R] : i = 1, \cdots, J\}$ where $J \leq e^{c_2/\eta}$ for some constant c_2 and $d_1(\Lambda_i^L, \Lambda_i^R) \leq \eta$ such that for any $\Lambda \in \mathcal{F}_\delta$, $\Lambda_i^L \leq \Lambda \leq \Lambda_i^R$ for some i with $1 \leq i \leq J$. Note that Λ_i^L, Λ_i^R $(i = 1, \cdots, J)$ may not belong to $\mathcal{F}_\delta$ and so they may not have a uniform positive lower bound and a uniform finite upper bound in $[\tau_0, \tau]$. Also note

that for any $\Lambda \in \mathcal{F}_0$, we have from Conditions A, B and C that

$$0 < \Lambda_0(\tau_0) - \delta_0 \leq \Lambda_0(t) - \delta_0 \leq \Lambda(t) \leq \Lambda_0(t) + \delta_0 \leq \Lambda_0(\tau) + \delta_0 \leq M + \delta_0$$

for $t \in [\tau_0, \tau]$ and

$$0 < L_0 - 2\delta_0 \leq \Delta\Lambda_0(T_{K,j}) - 2\delta_0 \leq \Delta\Lambda(T_{K,j}) \leq \Delta\Lambda_0(T_{K,j}) + 2\delta_0 \leq 2M + 2\delta_0$$

for $j = 1, ..., K$ with probability 1. Hence, for $\mathcal{M}_\delta$, we can construct a set of brackets $\{[M_i^L(X), M_i^R(X)] : i = 1, \cdots, J\}$ as follows:

$$\begin{aligned} M_i^L(X) &= \sum_{j=1}^{K} [\Delta N(T_{K,j}) \\ &\quad \times \log\left\{\max\left(\Lambda_i^L(T_{K,j}) - \Lambda_i^R(T_{K,j-1}),\ \Delta\Lambda_0(T_{K,j}) - 2\delta_0\right)\right\} \\ &\quad - \left\{\Lambda_i^R(T_{K,j}) - \Lambda_i^L(T_{K,j-1})\right\}] - m_{\Lambda_0}(X) \end{aligned}$$

and

$$\begin{aligned} M_i^R(X) &= \sum_{j=1}^{K} [\Delta N(T_{K,j}) \log\left\{\Lambda_i^R(T_{K,j}) - \Lambda_i^L(T_{K,j-1})\right\} \\ &\quad - \left\{\Lambda_i^L(T_{K,j}) - \Lambda_i^R(T_{K,j-1})\right\}] - m_{\Lambda_0}(X) \end{aligned}$$

Set $\|\cdot\|_{P,B}$ be the Bernstein norm as defined in Van der Vaart and Wellner (1996) and $N_{[]}$ the bracking number for the class M_δ. Then, it follows from Condition D that

$$\| M_i^R(X) - M_i^L(X) \|_{P,B}^2 \leq c_3 d_1^2(\Lambda_i^L, \Lambda_i^R) \leq c_3\eta^2, \quad i = 1, ..., J$$

for some constant c_3 and for any $\Lambda \in \mathcal{F}_\delta$,

$$\| m_\Lambda(X) - m_{\Lambda_0}(X) \|_{P,B}^2 \leq c_4 d_1^2(\Lambda, \Lambda_0) \leq c_4\delta^2$$

for some constant c_4. So

$$\log N_{[]}(\eta, \mathcal{M}_\delta, \|\cdot\|_{P,B}) \leq c_5\eta^{-1}$$

for some constant c_5. Hence, by applying Lemma 3.4.3 of Van der Vaart and Wellner (1996), we have

$$E^*||\sqrt{n}(P_n - P)||_{\mathcal{M}_\delta} \leq c_6\phi_n(\delta)$$

for some constant c_6, where E^* denotes the outer expectation, P_n is the empirical measure corresponding to X, $P_nf = \sum_{i=1}^n f(X_i)/n$, and $\phi_n(\delta) = \delta^{\frac{1}{2}} + \delta^{-1}n^{-\frac{1}{2}}$. Now, upon using Theorem 3.2.5 of Van der Vaart and Wellner (1996), $d_1(\hat{\Lambda}_n^*, \Lambda_0)$ converges in probability to zero of order at least $n^{-\frac{1}{3}}$. This completes the proof of the lemma.

Now we turn to the proof of Theorem 4.2.1. First, note that

$$\sqrt{n}\mathrm{P}f_{\hat{\Lambda}_n}(X) = -I_{1n} + I_{2n} + I_{3n}, \tag{4.30}$$

where

$$I_{1n} = \sqrt{n}(P_n - P)f_{\hat{\Lambda}_n}(X),$$

$$\begin{aligned} I_{2n} &= \sqrt{n}P_n\left[\sum_{j=1}^{K-1} W(T_{K,j})\hat{\Lambda}_n(T_{K,j})\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})}\right\}\right. \\ &\qquad \left. + W(T_{K,K})\hat{\Lambda}_n(T_{K,K})\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})}\right\}\right] \end{aligned}$$

and

$$\begin{aligned} I_{3n} &= \sqrt{n}P_n\left[\sum_{j=1}^{K-1} W(T_{K,j})\hat{\Lambda}_n(T_{K,j})\left\{\frac{\Delta\Lambda_0(T_{K,j+1}) - \Delta N(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})}\right.\right. \\ &\qquad \left. - \frac{\Delta\Lambda_0(T_{K,j}) - \Delta N(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})}\right\} \\ &\qquad \left. + W(T_{K,K})\hat{\Lambda}_n(T_{K,K})\frac{\Delta N(T_{K,K}) - \Delta\Lambda_0(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})}\right]. \end{aligned}$$

Let

$$g_\Lambda(X) = \sum_{j=1}^{K-1} W(T_{K,j})\Lambda(T_{K,j})\left\{\frac{\Delta\Lambda_0(T_{K,j+1}) - \Delta N(T_{K,j+1})}{\Delta\Lambda(T_{K,j+1})}\right.$$

$$-\frac{\Delta\Lambda_0(T_{K,j}) - \Delta N(T_{K,j})}{\Delta\Lambda(T_{K,j})}\Bigg\}$$
$$+W(T_{K,K})\Lambda(T_{K,K})\frac{\Delta N(T_{K,K}) - \Delta\Lambda_0(T_{K,K})}{\Delta\Lambda(T_{K,K})}.$$

Note that

$$I_{3n} = \sqrt{n}(P_n - P)g_{\hat{\Lambda}_n}(X) = I_{4n} + I_{5n},$$

where

$$I_{4n} = \sqrt{n}(P_n - P)\left\{g_{\hat{\Lambda}_n}(X) - g_{\Lambda_0}(X)\right\}$$

and

$$I_{5n} = \sqrt{n}(P_n - P)g_{\Lambda_0}(X)$$

It is easy to see that I_{5n} is a U-statistic and has an asymptotic normal distribution with mean zero and variance σ_w^2 that can be consistently estimated by $\hat{\sigma}_w^2$ as given in the statement of the theorem. Hence, it is sufficient to show that I_{1n}, I_{2n} and I_{4n} all converge in probability to zero.

We will show the convergence of I_{1n} first. Let I_{1n}^* denote the version of I_{1n} obtained by replacing $\hat{\Lambda}_n$ with $\hat{\Lambda}_n^*$. Then, to prove that I_{1n} converges to zero in probability, it is sufficient to show that $I_{1n}^* = o_p(1)$ since $P\{\hat{\Lambda}_n \neq \hat{\Lambda}_n^*\} < \varepsilon$. Let

$$\mathcal{F}_1 = \{f_\Lambda(X) : \Lambda \in \mathcal{F}_0\}.$$

Also let $\{[\Lambda_i^L, \Lambda_i^R] : i = 1, \cdots, J\}$ be a set of η-brackets for covering $\mathcal{F}_0$ with $J \leq e^{c/\eta}$ for some constant c by Theorem 2.7.5 of Van der Vaart and Wellner (1996). Then, for $\mathcal{F}_1$, we can construct a set of brackets $\{[f_i^L(X), f_i^R(X)] : i = 1, \cdots, J\}$ as follows:

$$\begin{aligned} f_i^L(X) &= \sum_{j=1}^{K-1} W(T_{K,j})\left[\frac{\Lambda_i^L(T_{K,j})\Delta\Lambda_0(T_{K,j+1})}{\Lambda_i^R(T_{K,j+1}) - \Lambda_i^L(T_{K,j})}\right. \\ &\qquad \left. -\frac{\Lambda_i^R(T_{K,j})\Delta\Lambda_0(T_{K,j})}{\max\left\{\Lambda_i^L(T_{K,j}) - \Lambda_i^R(T_{K,j-1}),\ \Delta\Lambda_0(T_{K,j}) - 2\delta_0\right\}}\right] \end{aligned}$$

$$+W(T_{K,K})\left[\Lambda_i^L(T_{K,K}) - \frac{\Lambda_i^R(T_{K,K})\Lambda_0(T_{K,K})}{\max\{\Lambda_i^L(T_{K,K}) - \Lambda_i^R(T_{K,K-1}),\ \Delta\Lambda_0(T_{K,K}) - 2\delta_0\}}\right]$$

and

$$\begin{aligned} f_i^R(X) &= \sum_{j=1}^{K-1} W(T_{K,j}) \\ &\quad \times \left[\frac{\Lambda_i^R(T_{K,j})\Delta\Lambda_0(T_{K,j+1})}{\max\{\Lambda_i^L(T_{K,j+1}) - \Lambda_i^R(T_{K,j}),\ \Delta\Lambda_0(T_{K,j+1}) - 2\delta_0\}} - \frac{\Lambda_i^L(T_{K,j})\Delta\Lambda_0(T_{K,j})}{\Lambda_i^R(T_{K,j}) - \Lambda_i^L(T_{K,j-1})}\right] \\ &\quad +W(T_{K,K})\left[\Lambda_i^R(T_{K,K}) - \frac{\Lambda_i^L(T_{K,K})\Lambda_0(T_{K,K})}{\Lambda_i^R(T_{K,K}) - \Lambda_i^L(T_{K,K-1})}\right]. \end{aligned}$$

It can be shown that

$$P\{f_i^R(X) - f_i^L(X)\}^2 \le c_1 d_1^2(\Lambda_i^R, \Lambda_i^L)$$

for some constant c_1 and for any $\Lambda \in \mathcal{F}_0$, $Pf_\Lambda^2(X) \le c_2 d_1^2(\Lambda, \Lambda_0)$ for some constant c_2. Hence, $\mathcal{F}_1$ is a P-Donsker class and it follows from our Lemma 4.6.2 and Corollary 2.3.12 of Van der Vaart and Wellner (1996) that $I_{1n}^* = o_p(1)$.

Next, we show the convergence of I_{2n}. Set $W_0 = W \circ \Lambda_0^{-1}$. Then, from Lemma 4.6.1, we can rewrite I_{2n} as

$$\begin{aligned} I_{2n} &= \sqrt{n}P_n\left[\sum_{j=1}^{K-1}\{W_0(\Lambda_0(T_{K,j})) - W_0(\hat{\Lambda}_n(T_{K,j}))\}\hat{\Lambda}_n(T_{K,j}) \right. \\ &\qquad \times\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})}\right\} \\ &\qquad +\{W_0(\Lambda_0(T_{K,K})) - W_0(\hat{\Lambda}_n(T_{K,K}))\}\hat{\Lambda}_n(T_{K,K}) \\ &\qquad \left. \times\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})}\right\}\right] \\ &= \Delta_{1n} + \Delta_{2n}, \end{aligned}$$

where

$$\begin{aligned}\Delta_{1n} &= \sqrt{n}(P_n - P)\left[\sum_{j=1}^{K-1}\{W_0(\Lambda_0(T_{K,j})) - W_0(\hat{\Lambda}_n(T_{K,j}))\}\hat{\Lambda}_n(T_{K,j})\right.\\ &\qquad \times\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})}\right\}\\ &\qquad +\{W_0(\Lambda_0(T_{K,K})) - W_0(\hat{\Lambda}_n(T_{K,K}))\}\hat{\Lambda}_n(T_{K,K})\\ &\qquad \left.\times\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})}\right\}\right]\end{aligned}$$

and

$$\begin{aligned}\Delta_{2n} &= \sqrt{n}P\left[\sum_{j=1}^{K-1}\{W_0(\Lambda_0(T_{K,j})) - W_0(\hat{\Lambda}_n(T_{K,j}))\}\hat{\Lambda}_n(T_{K,j})\right.\\ &\qquad \times\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})}\right\}\\ &\qquad +\{W_0(\Lambda_0(T_{K,K})) - W_0(\hat{\Lambda}_n(T_{K,K}))\}\hat{\Lambda}_n(T_{K,K})\\ &\qquad \left.\times\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})}\right\}\right].\end{aligned}$$

Let Δ_{1n}^* and Δ_{2n}^* denote the versions of Δ_{1n} and Δ_{2n} obtained by replacing $\hat{\Lambda}_n$ with $\hat{\Lambda}_n^*$, respectively. Set

$$\begin{aligned}h_\Lambda(X) &= \sum_{j=1}^{K-1}\{W_0(\Lambda_0(T_{K,j})) - W_0(\Lambda(T_{K,j}))\}\Lambda(T_{K,j})\\ &\qquad \times\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\Lambda(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\Lambda(T_{K,j})}\right\}\\ &\qquad +\{W_0(\Lambda_0(T_{K,K})) - W_0(\Lambda(T_{K,K}))\}\Lambda(T_{K,K})\\ &\qquad \times\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\Lambda(T_{K,K})}\right\}\end{aligned}$$

and

$$\mathcal{F}_2 = \{h_\Lambda(X) : \Lambda \in \mathcal{F}_0\}.$$

Note that the uniform covering entropy for $\mathcal{F}_0$ is bounded by c/η for some constant c from Theorem 2.7.5 of Van der Vaart and Wellner (1996). Since W_0 is a bounded Lipschitz function, it can be shown that for $\Lambda_1, \Lambda_2 \in \mathcal{F}_0$,

$$P\{(h_{\Lambda_1}(X) - h_{\Lambda_2}(X))^2\} \leq c_3 d_1^2(\Lambda_1, \Lambda_2)$$

for some constant c_3 and for any $\Lambda \in \mathcal{F}_0$,

$$P(h_\Lambda^2(X)) \leq c_4 d_1^2(\Lambda, \Lambda_0)$$

for some constant c_4. Hence, the uniform entropy for $\mathcal{F}_2$ is bounded by c/η, and then $\mathcal{F}_2$ is a P-Donsker class from Theorem 2.5.2 of Van der Vaart and Wellner (1996). Since $d_1(\hat{\Lambda}_n^*, \Lambda_0) \to_p 0$, it follows from the uniform asymptotic equicontinuity of the empirical process (Van der Vaart and Wellner, 1996, pp. 168-171) that $\Delta_{1n}^* = o_p(1)$. Then, we have $\Delta_{1n} = o_p(1)$ since $P\{\Delta_{1n} \neq \Delta_{1n}^*\} < \varepsilon$.

For Δ_{2n}^*, since W_0 is a bounded Lipschitz function, it follows that

$$|\Delta_{2n}^*| \leq c_5 \sqrt{n}\, d_1^2(\hat{\Lambda}_n^*, \Lambda_0),$$

where c_5 is a constant. This shows, from Lemma 4.6.2 and $P(\hat{\Lambda}_n \neq \hat{\Lambda}_n^*) < \varepsilon$, that $\Delta_{2n} = o_p(1)$.

For I_{4n}, we let I_{4n}^* denote the version of I_{4n} obtained by replacing $\hat{\Lambda}_n$ with $\hat{\Lambda}_n^*$, and let

$$\mathcal{F}_3 = \{g_\Lambda(X) - g_{\Lambda_0}(X) : \Lambda \in \mathcal{F}_0\}.$$

We can use the same techniques as those used for proving the convergence of I_{1n} to show that $\mathcal{F}_3$ is P-Donsker and $P\{g_\Lambda(X) - g_{\Lambda_0}(X)\}^2 \leq c_6 d_1^2(\Lambda, \Lambda_0)$ for some constant c_6, and hence $I_{4n}^* = o_p(1)$ which completes the proof of the theorem. □

4.6.2 Proof of Theorem 4.3.1

(i) To obtain the asymptotic distribution of $\mathbf{U}_n$, we first note that $U_n^{(l)}$ can rewritten as

$$U_n^{(l)} = U_{1n}^{(l)} - \sqrt{\frac{n}{n_l}} U_{2n}^{(l)} + U_{3n}^{(l)} + U_{4n}^{(l)} + U_{5n}^{(l)} + U_{6n}^{(l)}, \tag{4.31}$$

where for $l = 1, \ldots, k$,

$$\begin{aligned} U_{1n}^{(l)} &= \sqrt{n} P \left[\sum_{j=1}^{K-1} W(T_{K,j}) \hat{\Lambda}_n(T_{K,j}) \left\{ \frac{\Delta\Lambda_0(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{\Delta\Lambda_0(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})} \right\} \right. \\ &\quad \left. + W(T_{K,K}) \hat{\Lambda}_n(T_{K,K}) \left\{ 1 - \frac{\Delta\Lambda_0(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})} \right\} \right], \end{aligned}$$

$$\begin{aligned} U_{2n}^{(l)} &= \sqrt{n_l} P \left[\sum_{j=1}^{K-1} W(T_{K,j}) \hat{\Lambda}_{n_l}(T_{K,j}) \left\{ \frac{\Delta\Lambda_0(T_{K,j+1})}{\Delta\hat{\Lambda}_{n_l}(T_{K,j+1})} - \frac{\Delta\Lambda_0(T_{K,j})}{\Delta\hat{\Lambda}_{n_l}(T_{K,j})} \right\} \right. \\ &\quad \left. + W(T_{K,K}) \hat{\Lambda}_{n_l}(T_{K,K}) \left\{ 1 - \frac{\Delta\Lambda_0(T_{K,K})}{\Delta\hat{\Lambda}_{n_l}(T_{K,K})} \right\} \right], \end{aligned}$$

$$\begin{aligned} U_{3n}^{(l)} &= \sqrt{n}(P_n - P) \left[\sum_{j=1}^{K-1} W_n^{(l)}(T_{K,j}) \hat{\Lambda}_n(T_{K,j}) \right. \\ &\quad \times \left\{ \frac{\Delta\hat{\Lambda}_{n_l}(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{\Delta\hat{\Lambda}_{n_l}(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})} \right\} \\ &\quad \left. + W(T_{K,K}) \hat{\Lambda}_n(T_{K,K}) \times \left\{ 1 - \frac{\Delta\hat{\Lambda}_{n_l}(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})} \right\} \right], \end{aligned}$$

$$\begin{aligned} U_{4n}^{(l)} &= \sqrt{n} P \left[\sum_{j=1}^{K-1} \left\{ W_n^{(l)}(T_{K,j}) - W(T_{K,j}) \right\} \hat{\Lambda}_n(T_{K,j}) \right. \\ &\quad \times \left\{ \frac{\Delta\hat{\Lambda}_{n_l}(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{\Delta\hat{\Lambda}_{n_l}(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})} \right\} \end{aligned}$$

$$+\left\{W_n^{(l)}(T_{K,K}) - W(T_{K,K})\right\}\hat{\Lambda}_n(T_{K,K})$$
$$\times\left\{1 - \frac{\Delta\hat{\Lambda}_{n_l}(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})}\right\}\Bigg],$$

$$\begin{aligned}
U_{5n}^{(l)} &= \sqrt{n}P\Bigg[\sum_{j=1}^{K-1} W(T_{K,j})\left\{\hat{\Lambda}_n(T_{K,j}) - \hat{\Lambda}_{n_l}(T_{K,j})\right\} \\
&\qquad \times\left\{\frac{\Delta\hat{\Lambda}_{n_l}(T_{K,j+1}) - \Delta\Lambda_0(T_{K,j+1})}{\Delta\hat{\Lambda}_n(T_{K,j+1})}\right. \\
&\qquad\qquad \left. - \frac{\Delta\hat{\Lambda}_{n_l}(T_{K,j}) - \Delta\Lambda_0(T_{K,j})}{\Delta\hat{\Lambda}_n(T_{K,j})}\right\} \\
&\qquad + W(T_{K,K})\hat{\Lambda}_n(T_{K,K})\left\{-\frac{\Delta\hat{\Lambda}_{n_l}(T_{K,K}) - \Delta\Lambda_0(T_{K,K})}{\Delta\hat{\Lambda}_n(T_{K,K})}\right\}\Bigg],
\end{aligned}$$

and

$$\begin{aligned}
U_{6n}^{(l)} &= \sqrt{n}P\Bigg[\sum_{j=1}^{K-1} W(T_{K,j})\hat{\Lambda}_{n_l}(T_{K,j})\left\{(\Delta\hat{\Lambda}_{n_l}(T_{K,j+1}) - \Delta\hat{\Lambda}_0(T_{K,j+1}))\right. \\
&\qquad \times\left(\frac{1}{\Delta\hat{\Lambda}_n(T_{K,j+1})} - \frac{1}{\Delta\hat{\Lambda}_{n_l}(T_{K,j+1})}\right) \\
&\qquad -(\Delta\hat{\Lambda}_{n_l}(T_{K,j}) - \Delta\hat{\Lambda}_0(T_{K,j})) \\
&\qquad \left. \times\left(\frac{1}{\Delta\hat{\Lambda}_n(T_{K,j})} - \frac{1}{\Delta\hat{\Lambda}_{n_l}(T_{K,j})}\right)\right\} \\
&\qquad - W(T_{K,K})\hat{\Lambda}_{n_l}(T_{K,K})\left\{\Delta\hat{\Lambda}_{n_l}(T_{K,K}) - \Delta\Lambda_0(T_{K,K})\right\} \\
&\qquad \times\left\{\frac{1}{\Delta\hat{\Lambda}_n(T_{K,K})} - \frac{1}{\Delta\hat{\Lambda}_{n_l}(T_{K,K})}\right\}\Bigg].
\end{aligned}$$

From the proof of Theorems 4.2.1, we have for $l = 1, \ldots, k$,

$$U_{1n}^{(l)} = Y_n + o_p(1)$$

and

$$U_{2n}^{(l)} = Y_n^{(l)} + o_p(1),$$

where

$$Y_n = \sqrt{n}(P_n - P)\left[\sum_{j=1}^{K-1} W(T_{K,j})\Lambda_0(T_{K,j})\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\Lambda_0(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right\}\right.$$
$$\left. + W(T_{K,K})\Lambda_0(T_{K,K})\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\Lambda_0(T_{K,K})}\right\}\right]$$

and

$$Y_n^{(l)} = \sqrt{n_l}(P_{n_l} - P)\left[\sum_{j=1}^{K-1} W(T_{K,j})\Lambda_0(T_{K,j})\left\{\frac{\Delta N(T_{K,j+1})}{\Delta\Lambda_0(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right\}\right.$$
$$\left. + W(T_{K,K})\Lambda_0(T_{K,K})\left\{1 - \frac{\Delta N(T_{K,K})}{\Delta\Lambda_0(T_{K,K})}\right\}\right],$$

where $P_{n_l}f = \frac{1}{n_l}\sum_{i\in S_l} f(X_i)$ and S_l denotes the set of indices for subjects in group l, $l = 1, \ldots, k$. Evidently, $Y_n^{(l)}$'s are independent and identically distributed, and $\sqrt{n}Y_n = \sum_{l=1}^{k}\sqrt{n_l}Y_n^{(l)}$. Set $Z_n^{(l)} = Y_n - \sqrt{\frac{n}{n_l}}Y_n^{(l)}$, $l = 1, \ldots, k$ and $\mathbf{Z}_n = (Z_n^{(1)}, \ldots, Z_n^{(k)})^T$. Then,

$$Z_n^{(l)} = \sum_{i=1}^{k}\sqrt{\frac{n_i}{n}}Y_n^{(i)} - \sqrt{\frac{n}{n_l}}Y_n^{(l)},\ l = 1, \ldots, k,$$

and so

$$\mathbf{Z}_n = \mathbf{\Gamma}_n\mathbf{Y}_n = \mathbf{\Gamma}\mathbf{Y}_n + o_p(1),$$

where $\mathbf{\Gamma}_n$ and Γ are as given in Theorem 4.3.1, and

$$\mathbf{Y}_n = \left(Y_n^{(1)}, \ldots, Y_n^{(k)}\right)^T$$

converges in distribution to $\mathbf{Y}_w$ having a k-dimensional normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{diag}(\sigma_1^2, \ldots, \sigma_k^2)$, where σ_l^2's are given in the statement of the theorem. Thus, we have $\mathbf{Z}_n$ converging in distribution to a random variable $\mathbf{U}_w$ that has a normal distribution $N(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{U}_w})$, where $\mathbf{\Sigma}_{\mathbf{U}_w}$ is given in (4.11) of Theorem 4.3.1.

Now, we need to show that $U_{3n}^{(l)}$, $U_{4n}^{(l)}$, $U_{5n}^{(l)}$, and $U_{6n}^{(l)}$ all converge in probability to 0, $l = 1, \ldots, k$. Let $U_{3n}^{(l)*}$, $U_{4n}^{(l)*}$, $U_{5n}^{(l)*}$, and $U_{6n}^{(l)*}$ denote the version of $U_{3n}^{(l)}$, $U_{4n}^{(l)}$, $U_{5n}^{(l)}$, and $U_{6n}^{(l)}$ obtained by replacing $\hat{\Lambda}_n$ with $\hat{\Lambda}_n^*$ and $\hat{\Lambda}_{n_l}$ with $\hat{\Lambda}_{n_l}^*$, respectively. Then, to prove that $U_{3n}^{(l)}$, $U_{4n}^{(l)}$, $U_{5n}^{(l)}$, and $U_{6n}^{(l)}$ all converge in probability to 0, $l = 1, \ldots, k$, it is sufficient to show that $U_{3n}^{(l)*}$, $U_{4n}^{(l)*}$, $U_{5n}^{(l)*}$, and $U_{6n}^{(l)*}$ all converge in probability to 0, $l = 1, \ldots, k$.

For $U_{3n}^{(l)*}$, set

$$\mathcal{G} = \{\xi : [0, \tau] \to [0, b]\},$$

where b is the uniform upper bound of weight process $W_n^{(l)}$ $(l = 1, ..., k)$,

$$\begin{aligned}\psi_{\Lambda_1,\Lambda_2}(\xi, X) &= \sum_{j=1}^{K-1} \xi(T_{K,j})\Lambda_1(T_{K,j}) \left\{ \frac{\Delta\Lambda_2(T_{K,j+1})}{\Delta\Lambda_1(T_{K,j+1})} - \frac{\Delta\Lambda_2(T_{K,j})}{\Delta\Lambda_1(T_{K,j})} \right\} \\ &\quad + \xi(T_{K,K})\Lambda_1(T_{K,K}) \left\{ 1 - \frac{\Delta\Lambda_2(T_{K,K})}{\Delta\Lambda_1(T_{K,K})} \right\},\end{aligned}$$

and for $\xi \in \mathcal{G}$,

$$\Psi_\delta(\xi) = \{\psi_{\Lambda_1,\Lambda_2}(\xi, X) : \Lambda_1, \Lambda_2 \in \mathcal{F}_\delta\}.$$

Note that it follows from Theorem 2.7.5 of Van der Vaart and Wellner (1996) that

$$N_{[]}(\eta, \mathcal{F}_\delta, L_2(P)) \leq e^{c_1/\eta}$$

for some constant c_1. Then, we have

$$N_{[]}(\eta, \Psi_\delta(\xi), L_2(P)) \leq e^{2c_1/\eta}.$$

It can be easily shown that $|\psi_{\Lambda_1,\Lambda_2}(\xi, X)| \leq \psi(X)$, where $\psi(X) = c_2 K$ and $P\psi^2_{\Lambda_1,\Lambda_2}(\xi, X) \leq c_3\delta^2$, where c_2 and c_3 are universal constants for ξ. Thus,

$$J_{[]}(\delta, \Psi_\delta(\xi), L_2(P)) = \int_0^\delta \sqrt{1 + \log N_{[]}(\eta ||\psi||_{P,2}, \Psi_\delta(\xi), L_2(P))} d\eta \leq c_4\delta^{1/2}$$

for some constant universal c_4. Hence, from Theorem 2.14.2 of Van der Vaart and Wellner (1996), we have

$$\begin{aligned} & E^* \left\{ \sup_{\psi_{\Lambda_1,\Lambda_2}(\xi,X)\in\Psi_\delta(\xi)} |\sqrt{n}(P_n - P)\psi_{\Lambda_1,\Lambda_2}(\xi, X)| \right\} \\ \leq \; & c_5 \left[J_{[]}(\delta, \Psi_\delta(\xi), L_2(P)) + \sqrt{n} P\psi\{\psi > \sqrt{n}a(\delta)\} \right], \end{aligned}$$

where c_5 is an universal constant and

$$a(\delta) = \delta ||\psi||_{P,2} / \sqrt{1 + \log N_{[]}(\delta ||\psi||_{P,2}, \Psi_\delta(\xi), L_2(P))}.$$

Then, it can be easily shown that

$$\limsup_{n\to\infty} E^* \left\{ \sup_{\psi_{\Lambda_1,\Lambda_2}(\xi,X)\in\Psi_\delta(\xi)} |\sqrt{n}(P_n - P)\psi_{\Lambda_1,\Lambda_2}(\xi, X)| \right\} \leq c_6 \delta^{1/2}$$

for some universal constant c_6. It follows from $d_1(\hat{\Lambda}_n, \hat{\Lambda}_{n_l}) \xrightarrow{a.s.} 0$ that

$$\limsup_{n\to\infty} E|\sqrt{n}(P_n - P)\psi_{\hat{\Lambda}_n^*, \hat{\Lambda}_{n_l}^*}(W_n^{(l)}, X)| \leq c_6 \delta^{1/2}.$$

Letting $\delta \to 0$, we have

$$\lim_{n\to\infty} E|\sqrt{n}(P_n - P)\psi_{\hat{\Lambda}_n^*, \hat{\Lambda}_{n_l}^*}(W_n^{(l)}, X)| = 0$$

which yields $U_{3n}^{(l)*} = o_p(1)$.

For $U_{4n}^{(l)*}$, we note that

$$\begin{aligned} \left| U_{4n}^{(l)*} \right| \leq \; & c_7 \Bigg[\sqrt{n} P \left\{ \sum_{j=1}^{K} |W_n^{(l)}(T_{K,j-1}) - W(T_{K,j-1})| \right. \\ & \qquad \left. \times |\hat{\Lambda}_n^*(T_{K,j}) - \hat{\Lambda}_{n_l}^*(T_{K,j})| \right\} \\ & + \sqrt{n} P \left\{ \sum_{j=1}^{K} |W_n^{(l)}(T_{K,j}) - W(T_{K,j})| |\hat{\Lambda}_n^*(T_{K,j}) - \hat{\Lambda}_{n_l}^*(T_{K,j})| \right\} \\ & + \sqrt{n} P \left\{ \sum_{j=1}^{K} |W_n^{(l)}(T_{K,j}) - W(T_{K,j})| \right. \\ & \qquad \left. \times |\hat{\Lambda}_n^*(T_{K,j-1}) - \hat{\Lambda}_{n_l}^*(T_{K,j-1})| \right\} \Bigg] \\ = \; & c_7 (A_{1n}^{(l)} + A_{2n}^{(l)} + A_{3n}^{(l)}) \end{aligned}$$

for some constant c_7, where

$$\begin{aligned}
A_{1n}^{(l)} &= \sqrt{n}\int\int |W_n^{(l)}(u) - W(u)||\hat{\Lambda}_n^*(v) - \hat{\Lambda}_{n_l}^*(v)|\, d\mu_2(u,v) \\
&\leq \sqrt{n}\int\int |W_n^{(l)}(u) - W(u)||\hat{\Lambda}_n^*(v) - \Lambda_0(v)|\, d\mu_2(u,v) \\
&\quad + \sqrt{n}\int\int |W_n^{(l)}(u) - W(u)||\hat{\Lambda}_{n_l}^*(v) - \Lambda_0(v)|\, d\mu_2(u,v),
\end{aligned}$$

$$\begin{aligned}
A_{2n}^{(l)} &= \sqrt{n}\int_0^\tau |W_n^{(l)}(t) - W(t)||\hat{\Lambda}_n^*(t) - \hat{\Lambda}_{n_l}^*(t)|\, d\mu_1(t) \\
&\leq \sqrt{n}\int_0^\tau |W_n^{(l)}(t) - W(t)||\hat{\Lambda}_n^*(t) - \Lambda_0(t)|\, d\mu_1(t) \\
&\quad + \sqrt{n}\int_0^\tau |W_n^{(l)}(t) - W(t)||\hat{\Lambda}_{n_l}^*(t) - \Lambda_0(t)|\, d\mu_1(t)
\end{aligned}$$

and

$$\begin{aligned}
A_{3n}^{(l)} &= \sqrt{n}\int\int |W_n^{(l)}(v) - W(v)||\hat{\Lambda}_n^*(u) - \hat{\Lambda}_{n_l}^*(u)|\, d\mu_2(u,v) \\
&\leq \sqrt{n}\int\int |W_n^{(l)}(v) - W(v)||\hat{\Lambda}_n^*(u) - \Lambda_0(u)|\, d\mu_2(u,v) \\
&\quad + \sqrt{n}\int\int |W_n^{(l)}(v) - W(v)||\hat{\Lambda}_{n_l}^*(u) - \Lambda_0(u)|\, d\mu_2(u,v).
\end{aligned}$$

Using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\sqrt{n}\int\int |W_n^{(l)}(u) - W(u)||\hat{\Lambda}_n^*(v) - \Lambda_0(v)|\, d\mu_2(u,v) \\
\leq\ & c_8\sqrt{n}\left\{\int_0^\tau (W_n^{(l)}(t) - W(t))^2\, d\mu_1(t)\right\}^{1/2} \\
&\quad \times \left\{\int_0^\tau (\hat{\Lambda}_n^*(t) - \Lambda_0(t))^2\, d\mu_1(t)\right\}^{1/2} \\
\longrightarrow\ & 0
\end{aligned}$$

in probability, where c_8 is a constant, since

$$\left[\int_0^\tau \left\{\hat{\Lambda}_n^*(t) - \Lambda_0(t)\right\}^2 d\mu_1(t)\right]^{1/2} = O_p(n^{-1/3}).$$

Similarly, we have

$$\sqrt{n}\int\int |W_n^{(l)}(u) - W(u)||\hat{\Lambda}_{n_l}^*(v) - \Lambda_0(v)|\, d\mu_2(u,v) = o_p(1).$$

Thus, $A_{1n}^{(l)} = o_p(1)$. Similarly, we have $A_{2n}^{(l)} = o_p(1)$ and $A_{3n}^{(l)} = o_p(1)$. Hence, $U_{4n}^{(l)*} = o_p(1)$, $l = 1, \ldots, k$.

For $U_{5n}^{(l)*}$ and $U_{6n}^{(l)*}$, we note that

$$\begin{aligned}\left|U_{5n}^{(l)*}\right| &\leq c_9\left[\sqrt{n}P\left\{\sum_{j=1}^{K}|\hat{\Lambda}_n^*(T_{K,j}) - \Lambda_0(T_{K,j})|^2\right\}\right.\\ &\qquad\left.+\sqrt{n}P\left\{\sum_{j=1}^{K}|\hat{\Lambda}_{n_l}^*(T_{K,j}) - \Lambda_0(T_{K,j})|^2\right\}\right]\\ &= c_9\{\sqrt{n}d_1^2(\hat{\Lambda}_n^*, \Lambda_0) + \sqrt{n}d_1^2(\hat{\Lambda}_{n_l}^*, \Lambda_0)\}\end{aligned}$$

and

$$\left|U_{6n}^{(l)*}\right| \leq c_{10}\{\sqrt{n}d_1^2(\hat{\Lambda}_n^*, \Lambda_0) + \sqrt{n}d_1^2(\hat{\Lambda}_{n_l}^*, \Lambda_0)\}$$

for some constants c_9 and c_{10}. Hence, $U_{5n}^{(l)*} = o_p(1)$ and $U_{6n}^{(l)*} = o_p(1)$, $l = 1, \ldots, k$. Therefore, the proof of part (i) is completed.

(ii) We note that $V_n^{(l)} = U_n^{(1,l)} - U_n^{(l)}$, $l = 2, \ldots, k$, where $U_n^{(1,l)}$ is defined as $U_n^{(1)}$ by replacing $W_n^{(1)}$ with $W_n^{(l)}$ for $l = 2, \ldots, k$. Then, it follows from (i) that

$$V_n^{(l)} = -\sqrt{\frac{n}{n_1}}Y_n^{(1)} + \sqrt{\frac{n}{n_l}}Y_n^{(l)} + o_p(1)$$

for $l = 2, \ldots, k$ and so

$$\mathbf{V}_n = \mathbf{H}_n\mathbf{Y}_n + o_p(1) = \mathbf{H}\mathbf{Y}_n + o_p(1),$$

where $\mathbf{H}_n$ and $\mathbf{H}$ are given in the theorem. This completes the proof of part (ii).

(iii) To show that $\hat{\sigma}_l^2 - \sigma_w^2 = o_p(1)$ for $l = 1, \ldots, k$, we set

$$\begin{aligned}\phi(\xi, \Lambda, X) &= \sum_{j=1}^{K-1} \xi(T_{K,j})\Lambda(T_{K,j}) \left\{ \frac{\Delta N(T_{K,j+1})}{\Delta\Lambda(T_{K,j+1})} - \frac{\Delta N(T_{K,j})}{\Delta\Lambda(T_{K,j})} \right\} \\ &\quad + \xi(T_{K,K})\Lambda(T_{K,K}) \left\{ 1 - \frac{\Delta N(T_{K,K})}{\Delta\Lambda(T_{K,K})} \right\}.\end{aligned}$$

Then $\sigma_w^2 = P\phi^2(W, \Lambda_0, X)$ and $\hat{\sigma}_l^2 = P_n\phi^2(W_n^{(l)}, \hat{\Lambda}_n, X)$. Note that

$$\begin{aligned}\hat{\sigma}_l^2 - \sigma_w^2 &= P_n \left\{ \phi^2(W_n^{(l)}, \hat{\Lambda}_n, X) - \phi^2(W_n^{(l)}, \Lambda_0, X) \right\} \\ &\quad + P_n \left\{ \phi^2(W_n^{(l)}, \Lambda_0, X) - \phi^2(W, \Lambda_0, X) \right\} \\ &\quad + (P_n - P)\phi^2(W, \Lambda_0, X).\end{aligned}$$

It can be easily shown that

$$P_n \left\{ \phi^2(W_n^{(l)}, \hat{\Lambda}_n, X) - \phi^2(W_n^{(l)}, \Lambda_0, X) \right\} = o_p(1)$$

and

$$(P_n - P)\phi^2(W_0, \Lambda_0, X) = o_p(1).$$

Since it follows from Condition C that

$$\begin{aligned}&\left|\phi(W_n^{(l)}, \Lambda_0, X) - \phi(W, \Lambda_0, X)\right| = \left|\phi(W_n^{(l)} - W, \Lambda_0, X)\right| \\ \leq\ & b_1\{1 + N(T_{K,K})\} \sum_{j=1}^{K} \left|W_n^{(l)}(T_{K,j}) - W(T_{K,j})\right|\end{aligned}$$

with probability 1 for some constant b_1 and

$$\begin{aligned}&\left|\phi(W_n^{(l)}, \Lambda_0, X) + \phi(W, \Lambda_0, X)\right| = \left|\phi(W_n^{(l)} + W, \Lambda_0, X)\right| \\ \leq\ & b_2 K\{1 + N(T_{K,K})\}\end{aligned}$$

with probability 1 for some constant b_2, then we have from the Cauchy-Schwarz inequality, Conditions $P\{K \leq K_0\} = 1$ and D, and (4.13)

$$E\left|\phi^2(W_n^{(l)}, \Lambda_0, X_i) - \phi^2(W, \Lambda_0, X_i)\right|$$

$$
\begin{aligned}
&\leq b_3 E\left[\{1+N_i(T_{K_i,K_i})\}^2\left\{\sum_{j=1}^{K_i}\left|W_n^{(l)}(T_{K_i,j})-W(T_{K_i,j})\right|\right\}\right] \\
&\leq b_3\left[E\{1+N_i(T_{K_i,K_i})\}^4\right]^{1/2}\left[E\left\{\sum_{j=1}^{K_i}\left|W_n^{(l)}(T_{K_i,j})-W(T_{K_i,j})\right|\right\}^2\right]^{1/2} \\
&\leq b_4 \max_{1\leq i\leq n}\left[E\left\{\sum_{j=1}^{K_i}\left|W_n^{(l)}(T_{K_i,j})-W(T_{K_i,j})\right|^2\right\}\right]^{1/2} \\
&\longrightarrow 0
\end{aligned}
$$

where b_3 and b_4 are finite positive constants , which completes the proof of part (iii). □

REMARK 3. The monotonicity assumption of the weight process required by Zhang (2006) can be removed by using the same techniques as those used here.

Chapter 5

Nonparametric Behrens-Fisher Hypothesis Testing for Case 1 Interval-Censored Data

5.1 Introduction

This chapter discusses nonparametric comparison of distribution functions based on incomplete survival data: case 1 interval-censored or current status data (see, for example, Peto and Peto, 1972; Keiding, 1991; Groemeneboom and Wellner, 1992; Huang and Wellner, 1995; Sun, 1999). By current status data, we mean that for each subject, the event occurrence time is unknown but we know whether the event has occurred before the observation time. Current status data often occur in tumorigenicity experiments; see Dinse and Lagakos (1983), Dewanji and Kalbfleisch (1986), and Dinse (1994). In these studies, the time until the onset of a tumor is usually of interest and

the comparison of different treatments with respect to the tumor onset distributions is often required. However, the tumor onset time is not directly observable, and only the death time of animals under study and the status of tumor onset at the death time are observed. In this situation, the animals' death time serves as an observation or censoring time and could depend on the treatments. As pointed out by Dinse and Lagakos (1983), a comparison not accounting for differences in death times could overestimate or underestimate the treatment difference. As discussed in Sun (1999), tests assuming the same death distribution could overestimate the tumor rate difference when animals in one group have longer survival times and higher tumor rates than animals in the other group.

Survival comparison is usually one of the main goals in survival studies. For the case of right-censored failure time data, there are a number of well-established methods (see, Fleming and Harrington, 1991; Kalbfleisch and Prentice, 2002). For the case of interval-censored failure time data, several authors have discussed the problem; see Peto and Peto (1972), Finkelstein (1986), Sun (1996, 1998), Zhang et al. (2001), Sun *et al.* (2005), Yuen et al. (2006), and Zhu et al. (2008). Most of the existing research have focused on testing the hypothesis that specifies the two distributions to be identical for censored or interval-censored data assuming observation or censoring times have the same distribution. For example, Peto and Peto (1972) considered the two-sample comparison problem under the Lehmann-type alternatives $G_2(t) = G_1^{\theta}(t)$, where G_1 and G_2 are survival functions corresponding to the two different samples and θ is a parameter. In this case, the comparison problem reduces to testing $\theta = 1$ and they suggested to use the score test, which they referred to as the log-rank test. Assuming the proportional hazards model, a special case of Lehmann-type alternatives,

Finkelstein (1986) investigated the general k-sample comparison problem. For the problem, she also suggested to apply the score test for testing regression parameters equal to zero. Following Finkelstein (1986), Sun (1996) studied the same problem without assuming the proportional hazards model and developed a nonparametric test using the idea of the log-rank test for right-censored data (Kalbfleisch and Prentice, 2002). Sun *et al.* (2005) presented a class of generalized log rank tests for this type of survival data. In contrast, there exists limited research for the analysis of censored or interval-censored data when the distributions of censoring or observation times may be different for different groups. Sun (1999) discussed current status or case 1 interval-censored data with unequal censoring based on the proportional hazards model. More recently, Troendle and Yu (2006) considered another null hypothesis that specifies the probability an observation from the first distribution exceeds an observation from the second distribution equals the probability of the opposite ordering. This hypothesis is referred to as the nonparametric Behrens-Fisher hypothesis (NBFH) (see Brunner and Munzel, 2000). They developed a nonparametric test procedure by the using nonparametric likelihood approach and approximating the null distribution of the test statistic based on right-censored data.

In this chapter, motivated by Troendle and Yu (2006), we develop nonparametric test procedures for case 1 interval-censored data. The test procedures are presented in Section 5.2. Also in Section 5.2, the asymptotic distribution of the proposed test statistic is derived. Section 5.3 reports some simulation results for evaluating the proposed test procedure which reveal that the approach works well for the practical situations considered. In Section 5.4, we apply the approach to a set of current status data from a tumorigenicity experiment while Section 5.5 contains some concluding

remarks.

5.2 A Nonparametric Test for the NBFH

Suppose that X and T have the cumulative distribution functions F_1 and G_1, respectively, and that they are independent. Here, X is the variable of interest and T is a censoring variable. Also suppose that Y and U have the cumulative distribution functions F_2 and G_2, respectively, and that they are independent. Here, Y is the variable of interest and U is a censoring variable. Suppose that the only observable variables are (X, δ) and (Y, η), where $\delta = I(X \leq T)$ and $\eta = I(Y \leq U)$. Let $(T_1, \delta_1), \ldots, (T_{n_1}, \delta_{n_2})$ be a random sample of size n_1 drawn from the distribution of (T, δ), and $(U_1, \eta_1), \ldots, (U_{n_2}, \eta_{n_2})$ be another random sample of size n_2 from the distribution of (U, η). The identity hypothesis is that $F_1$and F_2 are identical, i.e.,

$$F_1(x) = F_2(x), \quad -\infty < x < \infty. \tag{5.1}$$

Note that rejection of (5.1) does not really permit one to claim that the values in the first population are larger or smaller than those in the second population. To formally make such a claim, one might want to test a less specific null hypothesis whose rejection leads to a stronger claim. Define

$$p = P\{X > Y\} + \frac{1}{2}P\{X = Y\}.$$

Then the NBFH (nonparametric Behrens-Fisher hypothesis) is

$$p = \frac{1}{2}. \tag{5.2}$$

Note that a rejection of (5.2) with an estimate of p greater than $1/2$ would allow one to conclude that the first population is, in this sense, larger than the second population.

In comparison to the identity hypothesis, rejecting the NBFH allows a stronger claim because the null consists of a larger class of distributions.

Note that

$$p = \int (1 - F_1) dF_2,$$

but F_1 and F_2 are unknown. In order to estimate p, we first introduce the nonparametric maximum likelihood estimators of F_1 and F_2. The log-likelihood function for F_1 is

$$l_{n_1}(F) = \sum_{i=1}^{n_1} \{\delta_i \log F(T_i) + (1 - \delta_i) \log(1 - F(T_i))\}, \tag{5.3}$$

where F is a right-continuous distribution function. The nonparametric maximum likelihood estimator (NPMLE) of F_1, $\hat{F}_{1,n_1}$, is defined to be the nondecreasing, nonnegative step function with possible jumps only occurring at observation time points T_i, $i = 1, \ldots, n_1$, that maximizes $l_{n_1}(F)$. Similarly, the nonparametric maximum likelihood estimator of F_2, $\hat{F}_{2,n_2}$, can be defined.

Now, we need to determine $\hat{F}_{1,n_1}(t)$ and $\hat{F}_{2,n_2}(t)$. The nonparametric maximum likelihood estimators $\hat{F}_{1,n_1}(t)$ and $\hat{F}_{2,n_2}(t)$ of F_1 and F_2 can be characterized in terms of the self-consistency equations, and can be explicitly expressed via a max-min formula, respectively (Groeneboom and Wellner, 1992; Huang and Wellner, 1995). The self-consistency equations for $\hat{F}_{1,n_1}(t)$ and $\hat{F}_{2,n_2}(t)$ are given by

$$\hat{F}_{1,n_1}(t) = E_{\hat{F}_{1,n_1}} \{F_{1,n_1}(t) | T_1, \ldots, T_{n_1}, \delta_1, \ldots, \delta_{n_1}\} \tag{5.4}$$

and

$$\hat{F}_{2,n_2}(t) = E_{\hat{F}_{2,n_2}} \{F_{2,n_2}(t) | U_1, \ldots, U_{n_2}, \eta_1, \ldots, \eta_{n_2}\}, \tag{5.5}$$

where F_{1,n_1} and F_{2,n_2} are the (unobservable) empirical distribution functions of random variables $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$, respectively. Clearly, $\hat{F}_{1,n_1}$ is the condi-

tional expectation of the empirical distribution function F_{1,n_1} at t, given the available information $\{(T_1, \delta_1), \ldots, (T_{n_1}, \delta_{n_1})\}$; similarly, $\hat{F}_{2,n_2}$ is the conditional expectation of the empirical distribution function F_{2,n_2} at t, given the available information $\{(U_1, \eta_1), \ldots, (U_{n_2}, \eta_{n_2})\}$. As mentioned in Huang and Wellner (1995), the self-consistency equations do not uniquely determine the NPMLEs $\hat{F}_{1,n_1}$ and $\hat{F}_{2,n_2}$. Therefore, we prefer to use the max-min formula. In order to present this formula, we relabel the data $(T_1, \delta_1), \ldots, (T_{n_1}, \delta_{n_1})$ in terms of the ordered values of $T_1, \ldots, T_{n_1}$ as $(T_{(1)}, \delta_{(1)}), \ldots, (T_{(n_1)}, \delta_{(n_1)})$, and (U_1, η_1), $\ldots$, (U_{n_2}, η_{n_2}) in terms of the ordered values of $U_1, \ldots, U_{n_2}$ as $(U_{(1)}, \eta_{(1)}), \ldots, (U_{(n_2)}, \eta_{(n_2)})$, where $T_{(1)} \leq \cdots \leq T_{(n_1)}$ and $U_{(1)} \leq \cdots \leq U_{(n_2)}$. Then, $\hat{F}_{1,n_1}$ and $\hat{F}_{2,n_2}$ are given by

$$\hat{F}_{1,n_1}(T_{(i)}) = \max_{j \leq i} \min_{k \geq i} \frac{\sum_{j \leq r \leq k} \delta_{(r)}}{k - j + 1}, \qquad i = 1, \ldots, n_1 \tag{5.6}$$

and

$$\hat{F}_{2,n_2}(U_{(i)}) = \max_{j \leq i} \min_{k \geq i} \frac{\sum_{j \leq r \leq k} \eta_{(r)}}{k - j + 1}, \qquad i = 1, \ldots, n_2. \tag{5.7}$$

The max-min formula will be used below in simulation studies and in the example for computing the NPMLEs of $\hat{F}_{1,n_1}$ and $\hat{F}_{2,n_2}$.

The nonparametric maximum likelihood estimator $\hat{p}$ of p is then given by

$$\hat{p} = \int (1 - \hat{F}_{1,n_1}) d\hat{F}_{2,n_2}.$$

The asymptotic normality result for $\hat{p}$ is as follows.

Theorem 5.2.1 *Suppose that*

(i) The support of F_1 is a bounded interval $I_1 = [0, M_1]$, $G_1 \ll F_1$, $F_1 \ll G_1$, and F_1 and G_1 have densities f_1 and g_1 with respect to Lebesgue measure, respectively.

Also suppose that the support of F_2 is a bounded interval $I_2 = [0, M_2]$, $G_2 \ll F_2$, $F_2 \ll G_2$, and F_2 and G_2 have densities f_2 and g_2 with respect to Lebesgue measure, respectively.

(ii) F_1, F_2, f_1, f_2, g_1 *and* g_2 *satisfy*

$$\sigma_1^2 \equiv \int_{I_1} \frac{F_1(x)(1-F_1(x))}{g_1(x)} f_2^2(x)\, dx < \infty$$

and

$$\sigma_2^2 \equiv \int_{I_2} \frac{F_2(x)(1-F_2(x))}{g_2(x)} f_1^2(x)\, dx < \infty.$$

(iii) $(f_2/g_1) \circ F_1^{-1}$ *and* $(f_1/g_2) \circ F_2^{-1}$ *are bounded Lipschitz functions on* $[0,1]$.

(iv) Let $n = n_1 + n_2$, *and* $\frac{n_1}{n} \longrightarrow q \in (0,1)$.

Then

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} N(0, \sigma^2),$$

where

$$\sigma^2 = \frac{1}{q}\sigma_1^2 + \frac{1}{1-q}\sigma_2^2.$$

Proof. Note that

$$\sqrt{n}(\hat{p} - p) = \sqrt{n}\left\{\int (1 - \hat{F}_{1,n_1})\, d\hat{F}_{2,n_2} - \int (1 - F_1)\, dF_2\right\} = \sqrt{\frac{n}{n_1}} U_{1n} - \sqrt{\frac{n}{n_2}} U_{2n}$$

where

$$U_{1n} = \sqrt{n_1} \int \left\{(1 - \hat{F}_{1,n_1}) - (1 - F_1)\right\}\, d\hat{F}_{2,n_2}$$

and

$$U_{2n} = -\sqrt{n_2}\left\{\int (1 - F_1)\, d\hat{F}_{2,n_2} - \int (1 - F_1)\, dF_2\right\}.$$

Clearly, U_{2n} can be rewritten as

$$\begin{aligned} U_{2n} &= \sqrt{n_2}\left\{\int (1-\hat{F}_{2,n_2})\, dF_1 - \int (1-F_2)\, dF_1\right\} \\ &= \sqrt{n_2}\int \left\{(1-\hat{F}_{2,n_2}(x)) - (1-F_2(x))\right\} f_1(x)\, dx. \end{aligned}$$

From Theorem 5.1 of Huang and Wellner (1995), we have

$$U_{2n} \xrightarrow{d} N(0, \sigma_2^2),$$

where σ_2^2 is as given in the theorem.

Now we need to show that

$$U_{1n} \xrightarrow{d} N(0, \sigma_1^2).$$

Since $\hat{F}_{1,n_1}$ is independent of $(U_1, \eta_1) \ldots, (U_{n_2}, \eta_{n_2})$ and $\hat{F}_{2,n_2}$, then

$$\begin{aligned} & \int (\hat{F}_{1,n_1} - F_1) d\hat{F}_{2,n_2} \\ = & \int (\hat{F}_{1,n_1} - F_1)\, dE_{\hat{F}_{2,n_2}}(F_{2,n_2}|U_1, \ldots, U_{n_2}, \eta_1, \ldots, \eta_{n_2}) \\ = & E_{\hat{F}_{2,n_2}}\left\{\int (\hat{F}_{1,n_1} - F_1)\, dF_{2,n_2}|U_1, \ldots, U_{n_2}, \eta_1, \ldots, \eta_{n_2})\right\}. \end{aligned}$$

Let P_{n_2} be the empirical measure of the random variables $Y_1, \ldots, Y_{n_2}$. Then,

$$\sqrt{n_1}\int (\hat{F}_{1,n_1} - F_1)\, dF_{2,n_2} = V_n + \sqrt{\frac{n_1}{n_2}}\Delta_n$$

where

$$V_n = \sqrt{n_1}\int (\hat{F}_{1,n_1} - F_1)\, dF_2$$

and

$$\Delta_n = \sqrt{n_2}(P_{n_2} - P)(\hat{F}_{1,n_1} - F_1).$$

Define

$$\mathcal{F} = \{F : F \text{ is a distribution function defined on } [0, M_2]\},$$

and

$$\mathcal{G} = \{F(t) - F_1(t) \,:\, F \in \mathcal{F}\}.$$

Then, $\mathcal{F}$ is a $P-$*Donsker* from the proof of Corollary 5.1 of Huang and Wellner (1995) and thus $\mathcal{G}$ is $P-$ *Donsker*. Also note that $\hat{F}_{1,n_1} \in \mathcal{F}$ for all n sufficiently large and as $n \to \infty$, we have that

$$\int |\hat{F}_{1,n_1}(t) - F_1(t)|^2 \, dP \longrightarrow 0$$

in probability from the strong consistency of $\hat{F}_{1,n_1}$. It thus follows from this and the uniform asymptotic equicontinuity of the empirical process resulting from the Donsker property (van der Vaart and Wellner, 1996, pp. 168–171) that

$$\Delta_n \longrightarrow 0$$

in probability as $n \to \infty$. Thus, we have

$$U_{1n} = V_n + o_p(1),$$

where

$$V_n = \sqrt{n_1} \int (\hat{F}_{1,n_1}(x) - F_1(x)) \, f_2(x) dx.$$

Upon using Theorem 5.1 of Huang and Wellner (1995), we have

$$U_{1n} \xrightarrow{d} N(0, \sigma_1^2),$$

where σ_1^2 is as given in the theorem. This completes the proof. □

Note that the expression of the variance σ^2 involves unknown functions f_1, F_1, g_1, G_1, f_2, F_2, g_2 and G_2. Clearly, F_k can be consistently estimated by $\hat{F}_{k,n_i}$ and G_k can

be consistently estimated by the empirical estimator G_{k,n_k}, $k = 1, 2$. However, the estimators of f_k and g_k cannot be obtained by $\hat{F}_{k,n_i}$ and G_{k,n_k}, $k = 1, 2$ since such estimators are not smooth functions. For this reason, we apply the bootstrap resampling method to estimate the variance in our simulations and also in the application. Based on the theorem, we can then construct approximate confidence interval for p and also test the null hypothesis.

5.3 Simulation Study

To investigate the finite sample properties of the proposed procedure, simulation studies are conducted. In the simulation, we consider the two-sample comparison problem, and mimic the set-up commonly used in periodic follow-up studies to generate case 1 interval-censored data. All censoring times are generated from uniform distribution $U(0, \theta_1)$ for group 1 and $U(0, \theta_2)$ for group 2. The survival times are generated from the exponential distribution with rate λ_k and Weibull distribution with shape a_k and scale λ_k for group k, $k = 1, 2$, which are often assumed in survival studies. Here, we consider three distribution pairs:

Case 1: Population 1 = Exponential($\lambda_1 = 1$); population 2 = Exponential(λ_2).

Case 2: Population 1 = Weibull($a_1 = 3, \lambda_1 = 1$); population 2 = Weibull($a_2 = 3, \lambda_2$).

Case 3: Population 1 = Weibull($a_1 = 1, \lambda_1 = 1$); population 2 = Weibull($a_2 = 1/2, \lambda_2$).

For the distribution pairs in Cases 1 and 2, we can obtain null or alternative config-

Table 5.1: Percentage of null hypothesis rejection at significance level 5% based on 1000 replications for three cases

Dist. pair	λ_2	θ_1	θ_2	$n_1 = n_2 = 50$	$n_1 = n_2 = 100$
Case 1	1.00	2	5	0.045	0.051
	0.50	2	5	0.375	0.596
	0.30	2	5	0.815	1.000
	0.10	2	5	0.985	1.000
Case 2	1.00	2	2	0.045	0.049
	0.75	2	2	0.526	0.843
	0.50	2	2	0.961	1.000
	0.25	2	2	1.000	1.000
Case 3	1.33	2	10	0.055	0.052
	0.50	2	10	0.384	0.602
	0.30	2	10	0.618	0.915
	0.10	2	10	0.865	1.000

urations for either the identity hypothesis or the NBFH by adjusting λ. Distribution pair in Case 3 represents a case of the NBFH when $\lambda_2 = 1.33$. For each case, we consider two samples, $n_1 = n_2 = 50$ and 100, respectively. The NPMLEs $\hat{F}_{1,n_1}$ and $\hat{F}_{2,n_2}$ are computed by using the max-min formula given in (5.6) and (5.7). All the results reported here are based on 1000 Monte Carlo replications using R software.

Table 5.1 present the empirical sizes and powers of the proposed test based on simulated current status data for different values of θ_1 and θ_2 and three different distributions. It can be seen from the table that the proposed test procedure seems to have the right size (nominal level 0.05 is used). The results indicate that the proposed test performs well under all situations considered here. It can be seen that when the sample size increases, the power increases as expected.

To evaluate the normal distribution approximation given in the theorem to the finite distribution of the proposed test statistic, we studied the quantile plots of the test statistic against standard normal distribution under different set-ups. All of them suggest that the normal approximation works quite well.

5.4 Application

In this section, we apply the proposed test procedure to the data from a tumorigenicity experiment described in Hoel and Walburg (1972) and discussed by Hotel and Walburg (1972), Dinse and Lagakos (1983), Finkelstein and Wolfe (1985), Huang (1996), Sun (1999), and Shen (2000), among others. In this study, there were 144 male RFM mice. They were randomly assigned to one of two treatment groups: conventional and germfree environments, and then examined at death. The survival time of interest

is the lung tumor onset time, but it is not directly observable since lung tumors are usually regarded as relatively nonlethal in mice. The data set includes the death time of mice at study and the status of tumor onset at the death time. Among the 155 mice, there were 96 mice in a conventional environment and 48 in a germfree environment. Let X and Y be the tumor onset times of mice in the conventional and germfree environments, respectively, and let T and U be the death time of mice in the conventional and germfree environments, respectively. Let $\delta = 1(X \leq T)$ and $\eta = 1(Y \leq U)$. Then, the observed data consist of $\{(t_i, \delta_i), i = 1, \ldots, 96\}$ for the conventional environment, and $\{(u_i, \eta_i), i = 1, \ldots, 48\}$ for the germfree environment. In the conventional environment, 27 mice had lung tumor ($\delta = 1$) and 69 mice did not have lung tumor ($\delta = 0$); in the germfree environment, 35 mice had lung tumor ($\eta = 1$) and 13 mice did not have lung tumor ($\eta = 0$). The focus here is on the effect of treatment on the development of lung tumor.

Let p be the probability that tumor onset time of mice from the conventional environment exceeds that of mice from germfree environment. Using the observed data and the proposed method, we obtained $\hat{p} = 0.07939$ with the standard error 0.0694, and $Z = (\hat{p} - 0.5)/0.0694 = -6.0609$ which corresponds to a p-value $\ll 0.0001$ for the null hypothesis $p = 1/2$. From this, we conclude that there are strongly significant differences between the two treatments and that the mice in the germfree environment have lung tumors less than those in the conventional environment. The result agrees with those given in Hoel and Walburg (1972), Dinse and Lagakos (1883), Finkelstein and Wolfe (1985), Huang (1996), Sun (1999), and Shen (2000).

5.5 Concluding Remarks

This chapter discussed the nonparametric two-sample problem with case 1 interval-censored or current status data. For the considered problem, we obtained the asymptotic normality of the nonparametric maximum likelihood estimator of the probability that an observation from the first distribution exceeds an observation from the second distribution. A nonparametric test for the nonparametric Behrens-Fisher hypothesis was proposed and the asymptotic normality of the proposed test statistic was also established.

In comparison with the test procedures given in Sun (1999), our method is more nonparametric and robust since the distributions of censoring times are left unspecified, but the proportional hazards models for censoring times are assumed by Sun (1999).

In comparison with the test procedures given in Troendle and Yu (2006), the asymptotic distribution was derived here by using empirical process theory and it does not involve any permutation and simulation. In contrast, the test procedure for right-censored data given in Troendle and Yu (2006) is more complicated by using the imputed permutation and simulation distribution for the approximation of the null distribution. Therefore, we need to develop a similar test procedure for right-censored data to the proposed one.

To avoid resampling, we need to seek some consistent smoothed estimators $\hat{f}_k$ and $\hat{g}_k$ for f_k and g_k $(k = 1, 2)$ such that the variance σ^2 can be consistently estimated by $\hat{\sigma}^2$

$$\hat{\sigma}^2 = \frac{n}{n_1}\hat{\sigma}_1^2 + \frac{n}{n_2}\hat{\sigma}_2^2,$$

where

$$\hat{\sigma}_1^2 = \int_{I_1} \hat{F}_{1,n_1}(x)(1 - \hat{F}_{1,n_1}(x)) \left\{ \frac{\hat{f}_2(x)}{\hat{g}_1(x)} \right\}^2 \, dG_{1,n_1}(x)$$

and

$$\hat{\sigma}_2^2 = \int_{I_2} \hat{F}_{2,n_2}(x)(1 - \hat{F}_{2,n_2}(x)) \left\{ \frac{\hat{f}_1(x)}{\hat{g}_2(x)} \right\}^2 \, dG_{2,n_2}(x).$$

Chapter 6

Conclusions and Future Research

6.1 Conclusions

This dissertation discussed the problems of the multi-sample comparison of counting processes when only panel count data are available. For the problems considered, we have developed some new nonparametric tests based on the nonparametric maximum pseudo-likelihood approach and the nonparametric maximum likelihood approach, respectively.

Using the nonparametric maximum pseudo-likelihood approach, we have constructed two classes of nonparametric test statistics based on the integrated weighted differences between the estimated mean functions of the count processes, established their asymptotic distributions, and examined their finite-sample properties through Monte Carlo simulations. The simulation results indicated that the proposed methods are good for practical use. Compared to the method presented by Sun and Fang (2003), the proposed method is applicable to more general situations. Compared to the method

presented by Zhang (2006), the weight processes used in the test statistics here have more options.

Under the same setting for panel count data as Sun and Fang (2003), we have presented a new nonparametric test statistic based on the nonparametric maximum likelihood estimator of the mean function of the counting processes over observation times, derived its asymptotic distribution and also examined its finite-sample property through Monte Carlo simulations. The simulation results showed that the proposed method is good for practical use and also more powerful than the existing nonparametric tests based on the nonparametric maximum pseudo-likelihood estimator.

Under a more general setting for panel count data, we have presented two classes of nonparametric tests based on the accumulated weighted differences between the rates of increase of the estimated mean functions of the counting processes over observation times, wherein the nonparametric maximum likelihood approach is used to estimate the mean functions instead of the nonparametric maximum pseudo-likelihood. The asymptotic properties of the test statistics were established and their performance were evaluated through Monte Carlo simulations. Simulation studies revealed that the proposed method works well in practical situations, and are also more powerful than the tests constructed with the use of nonparametric maximum pseudo-likelihood estimators of the mean functions. Compared to the existing methods (Park, Sun and Zhao, 2007; Zhang, 2006; Sun and Fang, 2003; Thall and Lachin, 1988), the proposed methods apply to more general situations, more powerful, and more robust.

When each subject at study is observed only once, panel count data reduce to the case 1 interval-censored or current status data. In this case, the mean function of counting process is the distribution function of the event occurrence time, and

so the problem of interest is to compare the distribution functions between different samples. For nonparametric two-sample comparison based on censored or interval-censored data, most of the existing methods have focused on testing the hypothesis that specifies the two population distributions to be identical under the assumption that observation or censoring times have the same distribution. We have considered the nonparametric Behrens-Fisher hypothesis, and studied the asymptotic property of the nonparametric maximum likelihood estimator of the probability that an observation from the first distribution exceeds an observation from the second distribution. Based on the asymptotic result, we presented a nonparametric test for the NBFH, proved its asymptotic normality, and evaluated its performance by simulations in case of small samples. In comparison with the existing methods for testing equality of two distributions with current status data, it is not required in our case that observation or censoring times have the same distribution, and rejecting the nonparametric Behrens-Fisher hypothesis allows for a stronger claim. Compared to the method presented by Sun (1999), the proposed method is more nonparametric and more robust.

6.2 Future Research

6.2.1 Analysis of Panel Count Data with Unequal Observation Times

The existing nonparametric methods and the tests presented in Chapters 2-4 are based on the assumption that observation times have the same distribution for different treatment groups. This may not be true in practice. To remove this assumption, we

need to modify the test statistics. Zhang (2006) suggested a two-sample test statistic as

$$\nu_n(\tilde{\Lambda}_{n_1}, \tilde{\Lambda}_{n_2}) = \left(\frac{n_1 n_2}{n}\right)^{1/2} \left\{ \int \tilde{\Lambda}_{n_1}(t)\hat{\eta}_{n_1}(t)d\hat{\mu}_{n_1}(t) - \int \tilde{\Lambda}_{n_2}(t)\hat{\eta}_{n_2}(t)d\hat{\mu}_{n_2}(t) \right\},$$

where $\hat{\eta}_{n_l}(t)$ $(l = 1, 2)$ can be chosen as the reciprocal of the kernel-smoothed estimator of derivative of $\mu_l(t)$,

$$\mu_l(t) = E\left\{ \sum_{j=1}^{K^{(l)}} 1\left(T^{(l)}_{K^{(l)},j} \leq t\right) \right\}$$

for the lth sample, and

$$\hat{\mu}_l(t) = \frac{1}{n_l} \sum_{i=1}^{n_l} \left\{ \sum_{j=1}^{K_i^{(l)}} 1\left(T^{(l)}_{K_i^{(l)},j} \leq t\right) \right\}.$$

Here, computation of $\nu_n(\tilde{\Lambda}_{n_1}, \tilde{\Lambda}_{n_2})$ is not straightforward since one needs to find a consistent estimator for $\mu'(t)$. To overcome this drawback, we wish to present a new test statistic which only involves $\tilde{\Lambda}_{n_l}$ (or $\hat{\Lambda}_{n_l}$) and $\hat{\mu}_l(t)$ $(l = 1, 2)$ such that it is easily computable. Here, $\tilde{\Lambda}_{n_l}$ and $\hat{\Lambda}_{n_l}$ denote the NPMPLE and NPMLE of the true mean function Λ_l from sample l, respectively.

6.2.2 Analysis of Over/Under-dispersed Panel Count Data

It is assumed that the counting process arising from each subject is a non-homogeneous Poisson process. However, it is equi-dispersed meaning that the mean is equal to the variance. In practice, the count process is not always a Poisson process since under/over-dispersed count data often occur in practice, and it might be represented by extended Poisson process. The weighted Poisson distributions are modified Poisson distributions that provide a unified approach to handle both over-dispersion and under-dispersion; see, for example, Balakrishnan and Kozubowski (2008) and Kokonendji,

Mizère and Balakrishnan (2008), among others. It is therefore desirable to apply extended Poisson process models with the weighted Poisson distribution instead of Poisson distribution to the analysis of panel count data.

6.2.3 Nonparametric Behrens-Fisher Hypothesis Testing Based on Mixed Case Interval-Censored Data

We have discussed the problem of the nonparametric Brehrens-Fisher hypothesis testing for case 1 interval-censored data. Case 1 and Case 2 interval-censored data are special cases of mixed case interval-censored, with mixed case interval-censored data being a special type of panel count data. Let Y_i be the event occurrence time for subject i, define the counting process arising from subject i as a one-jump process

$$N_i(T_{K_i,j}) = I\left(Y_i \leq T_{K_i,j}\right).$$

This type of panel count data are referred to as mixed case interval-censored data by Schick and Yu (2000). In this case, the mean function of the counting process is the distribution function of the event time. So, we wish to discuss the problem of the nonparametric Brehrens-Fisher hypothesis testing for such data. The key is to study the asymptotic property of

$$\hat{p} = \int (1 - \hat{F}_{n_1}) d\hat{F}_{n_2},$$

where $\hat{F}_{n_l}$ is the nonparametric maximum likelihood estimator of the distribution function F_l of the event occurrence time for sample l, $l = 1, 2$.

6.2.4 Nonparametric Behrens-Fisher Hypothesis Testing Based on Partly Interval-Censored Data

We also wish to discuss the problem of the nonparametric Behrens-Fisher hypothesis testing based on partly interval-censored data. By partly interval-censored data, we mean that for some subjects, the exact event times are observed, but for the remaining subjects, the event time of interest is observed only as belonging to an interval instead of being exactly known or right-censored; see, for example, Peto and Peto (1972), Huang (1999), Kim (2003), and Zhao et al, the Framingham Heart Disease Study (Odell, Anderson and D'Agostino, 1992) and the Danish Diabetes Study (Ramlau-Hansen, Jespersen and Andersen, 1987). For the problem considered, we needs to study the asymptotic property of the nonparametric maximum likelihood estimator for

$$p = P\{X > Y\} + \frac{1}{2}P\{X = Y\}.$$

Bibliography

Aalen, O. O. (1975). *Statistical Inference for a Family of Counting Processes.* PhD Thesis, University of California, Berkeley.

Aalen, O. O. (1978). Nonparametric inference for a family of counting processses. *The annals of Statistics* **6**; 701–726.

Andersen, P. K. and Borgan, Ø. (1985). Counting process models for life history data: a review. *Scandinavian Journal of Statistics* **12**, 97–158.

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, Berlin.

Andersen, P. K. and Ronn, B. B. (1995). A nonparametric test for comparing two samples where all observations are either left- or right-censored. *Biometrics* **51**, 323–329.

Andrews, D. F. and Herzberg, A. M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker.* Springer-Verlag, New York.

Balakrishnan, N. and Kozubowski, T. J. (2008). A class of weighted Poisson processes. *Statistics & Probability Letters* (to appear).

Breslow, N. E. (1984). Extra-Poisson variation in log-linear models. *Applied Statistics* **33**, 38–44.

Brunner, E. and Munzel, U. (2000). The nonparametric Behrens-Fisher problem: asymptotic theory and a small-sample approximation. *Biometrical Journal* **42**, 17–25.

Byar, D. P. (1980). The veterans administration study of chemoprophylaxis for recurrent stage I bladder tumors: comparisons of placebo, pyridoxine, and topical thiotepa. In *Bladder Tumors and Other Topics in Urological Oncology*, pp. 363–370, edited by M. Pavone-Macaluso, P. H. Smith and F. Edsmyr. Plenum, New York.

Byar, D. P., Blackard, C. and the VACURG (1977). Comparisons of placebo, pyridoxine, and topical thiotepa in preventing recurrence of stage I bladder cancer. *Urology* **10**, 5556–5561.

Cameron, A. C. and Trivedi, P. K. (1998). *Regression Analysis of Count data.* London: Cambridge University Press.

Chen, B. E., Cook, R. J., Lawless, J. F., and Zhan, M. (2005). Statistical methods for multivariate interval-censored recurrent events. *Statistics in Medicine* **24**, 671–691.

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

Cook, R. J., Lawless, J. F. and Nadeau, C. (1996). Robust tests for treatment comparisons based on recurrent event response. *Biometrics* **52**, 557–571.

Davis, C. S. & Wei, L. J. (1988). Nonparametric methods for analyzing incomplete nondecreasing repeated measurements. *Biometrics* **44**, 1005–1018.

Dewanji, A. and Kalbfleisch, J. D. (1986). Nonparametric methods for survival/sacrifice experiments. *Biometrics* **42**, 325–341.

Dinse, G. E. (1994). A comparison of tumor incidence analyses applicable in single-sacrifice animal experiments. *Statistics in Medicine* **13**, 689–708.

Dinse, G. E. and Lagakos, S. W. (1983). Regression analysiss of tumor prevalence data. *Applied Statistics* **32**, 236–248.

Finkelstein, D. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics* **42**, 845–854.

Finkelstein, D. and Wolfe, R. (1985). A semiparametric model for regression analysis of interval-censored failure time data. *Biometrics* **41**, 933–945.

Fleming, T.R. and Harrington, D.P. (1991). *Counting Process and Survival Analysis.* New York: John Wiley.

Groeneboom, P. (1996). Inverse problems in statistics. *Proceedings of the St. Flour Summer in Probability. Lecture Notes in Math.* **1648**, 67–164. Springer, Berlin.

Groeneboom, P. and Wellner, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation.* Basel: Birkhäuser-Verlag.

Hoel, D. and Walburg, H. (1972). Statistical analysis of survival experiments. *Journal of National Cancer Institute* **49**, 361–372.

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

Huang, J. (1996). Efficient estimation for the proportional hazards model with interval censoring. *The Annals of Statistics* **24**, 540– 568.

Huang, J. (1999). Asymptotic properties of nonparametric estimation based on partly interval-censored data. *Statistica Sinica*, **9**, 501–519.

Huang, C.-Y., Wang, M.-C., and Zhang, Y. (2006). Analysing panel count data with informative observation times. *Biometrika*, **93**, 763–775.

Huang, J. and Wellner, J. A. (1995). Asymptotic normality of the NPMLE of linear functionals for interval censored data, case I. *Statistica Neerlandica* **49** 153–163.

Kalbfeisch, J. D. and Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association* **80**, 863–871.

Kalbfleisch, J.D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data.* Second edition, New York : Wiley.

Keiding, N. (1991). Age-specific incidence and prevalence; a statistical perspective (with discussion). *Journal of the Royal Statistical Society* A, **154**, 371–412.

Kim, J.S. (2003). Maximum likelihood estimation for the proportional hazards model with partly interval-censored data. *Journal of the Royal Statistical Society Series B*, **65**, 489–502.

Kokonendji, C. C., Mizère, D. and Balakrishnan, N. (2008). Connections of the Poisson weight function to overdispersion and underdispersion. *Journal of Statistical Planning and Inference* **138**, 1287–1296.

Li, G. (2003). Nonparametric likelihood ratio goodness-of-fit tests for survival data. *Journal of Multivariate Analysis* **86**, 166–182.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

Odell, P.M., Anderson, K.M. and D'Agostino, R.B. (1992). Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics*, **48**, 951–959.

Park, D-H., Sun, J. and Zhao, X. (2007). A class of two-sample nonparametric tests for panel count data. *Communications in Statistics: Theory and Methods* **36**, 1611–1625.

Pepe, M. S. and Fleming, T. R. (1989). Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. *Biometrics* **45**, 497–507.

Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society, Series A* **135**, 185–207.

Petroni, G. R. and Wolfe, R. A. (1994). A two sample test for stochastic ordering with interval-censored data. *Biometrics* **50**, 77–87.

Ramlau-Hansen, H., Jespersen, N.C. and Anderson, P.K. (1987). Life insurance for insulin-dependent diabetics. *Scandinavian Actuarial Journal* **39**, 19–36.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. New York: John Wiley & Sons.

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

Schoenfield, L. J., Lachin, J. M., et al. (1981). National cooperative gallstone study: A controlled trial of the efficiency and safety of chenodeoxycholic acid for dissolution of gallstones. *Annals of Internal Medicine* **95**, 257–292.

Shen, X. (2000). Linear regression with current status data. *Journal of the American Statistical Association* **451**, 842–852.

Sinha, D. and Maiti, T. (2004). A Bayesian approach for the analysis of panel-count data with dependent termination. *Biometrics* **60**, 34-40.

Sun, J. (1996). A nonparametric test for interval-censored failure time data with application to AIDS studies. *Statistics in Medicine* **15**, 1387–1395.

Sun, J. (1998). Interval censoring. In *Encyclopedia of Biostatistics*, pp. 2090–2095. New York: John Wiley.

Sun, J. (1999). A nonparametric test for current status data with unequal censoring. *Journal of the Royal Statistical Society, Series B* **61**, 243-250.

Sun, J. (2006). *Statistical Analysis of Interval-Censored Failure Time Data*. New York: Springer.

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 data. *Statistica Sinica* **5**, 279–289.

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* **62**, 293–302.

Sun, J., Zhao, Q. and Zhao, X. (2005). Generalized long-rank tests for interval-censored failure time data. *Scandinavian Journal of Statistics* **32**, 49–57.

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.

Thall, P. F. and Vail, S. C. (1990). Some covariance models for longitudinal count data with overdispersion. *Biometrics* **46**, 657-671.

Troendle, J. F. and Yu, K. F. (2006). Likelihood approaches to the non-parametric two-sample problem for right-censored data. *Statistics in Medicine* **25**, 2284–2298.

Turnbull, B. W. (1976). The empirical distribution with arbitrarily grouped censored and truncated data. *Journal of the Royal Statistical Society, Series B* **38**, 290–295.

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

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

Wellner, J. A., Zhang, Y. and Liu, H. (2004). A semiparametric regression model for panel count data: when do pseudo-likelihood estimators become badly inefficient?. *Proceedings of the Second Seattle Symposium in Biostatistics.* New York: Springer, pp. 143–174.

Yuen, K.-C., Shi, J. and Zhu, L. (2006). A k-sample test with interval censored data. *Biometrika* **93**, 315–328.

Zhang, Y. (2002). A semiparametric pseudolikelihood estimation method for panel count data. *Biometrika* **89**, 39–48.

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

Zhang, Y. and Jamshidian, M. (2003). The gamma-frailty Poisson model for the nonparametric estimation of panel count data. *Biometrics* **59**, 1099–1106.

Zhang, Y., Liu, W. and Zhan, Y. (2001). A nonparametric two-sample test of the failure function with interval censoring, case 2. *Biometrika* **88**, 677–686.

Zhao, X., Zhao, Q., Sun, J. and Kim, J. S. (2008). Generalized log rank tests for partly interval-censored failure time data. *Biometrical Journal* (to appear).

Zhu, J., Yuen, K. C., Sun, J. and Zhao, X. (2008). A nonparametric test for interval-censored failure time data with unequal censoring. *Communications in Statistics: Theory and Methods* **37**, No 12.