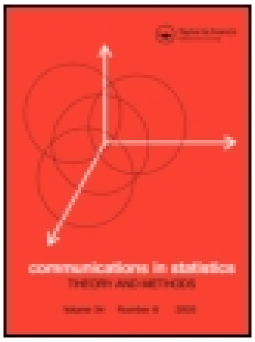

Communications in Statistics—Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

A Class of Two-Sample Nonparametric Tests for Panel Count Data

Do-Hwan Park , Jianguo Sun & Xingqiu Zhao

Communications in Statistics—Theory and Methods, 36: 1611–1625, 2007

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920601125912

Taylor & Francis
Taylor & Francis Group

Survival Analysis

A Class of Two-Sample Nonparametric Tests for Panel Count Data

DO-HWAN PARK[1], JIANGUO SUN[2],
AND XINGQIU ZHAO[3]

[1]Department of Mathematics and Statistics, University of Nevada, Reno, Nevada, USA
[2]Department of Statistics, University of Missouri, Columbia, Missouri, USA
[3]Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

Panel count data frequently occur in many situations including medical follow-up studies and reliability experiments. For two-sample comparison based on panel count data, several procedures have been proposed including Thall and Lachin (1988) and Sun and Fang (2003). In this article, a new class of nonparametric test procedures are presented. The test is a generalization of that for the same problem for failure time data and overcomes some shortcomings of the existing methods. Monte Carlo simulation studies are conducted to evaluate the presented approach and suggest that it works well. An illustrative example is discussed.

Keywords Cancer study; Follow-up study; Panel count data; Point processes; Two-sample comparison.

Mathematics Subject Classification Primary 62G10; Secondary 62N05.

1. Introduction and Notation

Consider a recurrent event study and suppose that only panel count data are available (Sun and Fang, 2003; Zhang, 2002). By panel count data, we mean that each study subject is observed only at discrete observation time points instead of continuously over an interval. Suppose that the study involves n independent subjects and define $N_i(t)$ as a point process recording the number of occurrences of the event up to time t, $i = 1, \ldots, n$. Then by panel count data, only the values of $N_i(t)$ at the observation time points are available. This article considers the two-sample analysis of panel count data. Specifically, suppose that the study

Received November 4, 2005; Accepted September 15, 2006
Address correspondence to Do-Hwan Park, Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557, USA; E-mail: dopark@unr.edu

involves two groups: control (group 1) and treatment (group 2) groups. Let $\Lambda_1(t)$ and $\Lambda_2(t)$ denote the mean functions of $N_i(t)$ corresponding to the control and treatment groups, respectively. Our main goal is to test the hypothesis $H_0 : \Lambda_1(t) = \Lambda_2(t)$.

Several authors have considered the analysis of panel count data. For example, Kalbfleisch and Lawless (1985) studied panel count data arising from Markov chains and Sun and Kalbfleisch (1995) and Wellner and Zhang (2000) investigated estimation of the mean function of the $N_i(t)$'s when all subjects are from the same population. Sun and Wei (2000) and Zhang (2002) discussed regression analysis of panel count data. For testing the hypothesis H_0 based on panel count data, two model-free approaches are available in the literature. One is given by Thall and Lachin (1988) who suggested to transform the two-sample comparison problem to a multivariate comparison problem and then to apply a multivariate Wilcoxon-like rank test. For the transformation, one needs to partition the whole study period into several fixed, consecutive, and non overlapping intervals. It is apparent that the test result may depend on these grouping intervals.

Sun and Fang (2003) gave the other model-free approach and suggested to base the test on the comparison between the estimators of $\Lambda_2(t)$ and the common mean function $\Lambda_0(t)$ under H_0. In particular, their test statistic has the form

$$U_{SF} = \int w(t)\left\{\widehat{\Lambda}_2(t) - \widehat{\Lambda}_0(t)\right\} d\widetilde{N}(t),$$

where $\widehat{\Lambda}_2(t)$ and $\widehat{\Lambda}_0(t)$ are the isotonic regression estimators of $\Lambda_2(t)$ and $\Lambda_0(t)$ defined in the next section and $w(t)$ and $\widetilde{N}$ are some function and process defined by the observed data. One drawback of the above procedure is that the validity of the statistic requires the assumption that the group or treatment indicators can be regarded as independent and identically distributed random variables, which may not be true in practice.

Panel count data arise in many medical follow-up studies and reliability experiments. For example, Schoenfield and Lachin (1981) and Thall and Lachin (1988) described a follow-up study on patients with floating gallstones. In the study, 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, the observed information includes 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 patients in different treatment groups. The example is discussed in more details below. In addition, panel count data also occur in AIDS clinical trials, animal tumorgenicity experiments, and sociological studies.

In panel count data, each subject or underlying continuous point process is observed only at the finite number of time points. In contrast, if every subject is observed continuously over an interval $(0, \tau]$, the resulting data are usually referred to as recurrent event data (Andersen et al., 1993; Cook and Lawless, 1997; Lawless and Nadeau, 1995; Lin et al., 2000; Wang and Chang, 1999). A number of authors have studied the analysis of recurrent event data. For example, Andersen et al. (1993) gave an excellent book that includes most of commonly used statistical methods for the analysis of recurrent event data.

The remainder of the article is organized as follows. Section 2 discusses nonparametric test of the hypothesis H_0 when only panel count data are available and presents a class of nonparametric test statistics. The statistics, motivated by similar statistics in survival analysis, are formulated as the integrated weighted differences between estimated mean functions corresponding to the two treatments. To estimate the mean function, the isotonic regression estimate is used (Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000). In Sec. 3, the asymptotic normal distribution is established for the test statistics. Section 4 investigates finite sample properties of the proposed test statistics through simulation studies and Sec. 5 gives the application of the proposed methods to the aforementioned gallstone follow-up study. In Sec. 6, some concluding remarks follow.

2. A Class of Nonparametric Tests

Let the $N_i(t)$'s, $\Lambda_1(t)$, and $\Lambda_2(t)$ be defined as before. For subject i, let $0 < t_{i,1} < \cdots < t_{i,k_i}$ denote the time points at which $N_i(t)$ is observed, $i = 1, \ldots, n$. Also, let n_1 and n_2 $(n_1 + n_2 = n)$ denote the numbers of subjects in control and treatment groups, respectively, and $n_{i,j} = N_i(t_{i,j})$, the observed value of $N_i(t)$ at $t_{i,j}$, $j = 1, \ldots, k_i$, $i = 1, \ldots, n$.

To test the hypothesis H_0, let $\widehat{\Lambda}_{n_1}$ and $\widehat{\Lambda}_{n_2}$ be the estimators of Λ_1 and Λ_2 based on samples from subjects in control and treatment groups, respectively. Motivated by the idea commonly used in survival studies (Pepe and Fleming, 1989), we propose to use the statistic

$$U_n = \sqrt{\frac{n_1 n_2}{n}} \int_0^\tau W_n(t)\{\widehat{\Lambda}_{n_1}(t) - \widehat{\Lambda}_{n_2}(t)\} dG_n(t)$$

where τ is the largest observation time, $W_n(t)$ is a bounded weight process, and

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

The statistic U_n is the integrated weighted differences between $\widehat{\Lambda}_{n_1}$ and $\widehat{\Lambda}_{n_2}$ and is sensitive especially to stochastically ordered mean functions. Statistics similar to U_n are commonly used in survival analysis. For two sample survival comparison with right-censored data, for example, Pepe and Fleming (1989) proposed some test statistics that have the same format as U_n with replacing $\widehat{\Lambda}_{n_1}$ and $\widehat{\Lambda}_{n_2}$ by estimated survival functions. Petroni and Wolfe (1994) and Zhang et al. (2001) used similar statistics for the same comparison problem based on interval-censored data.

Note that for testing H_0, the statistic U_n compares two individual mean functions directly. In contrast, the statistic U_{SF} used in Sun and Fang (2003) compares one individual mean function with the whole mean function. Thus it seems natural to expect that the test given here has better power than that given in Sun and Fang (2003) as supported by the simulation results given in Sec. 4.

In the above definition, $\widehat{\Lambda}_{n_1}$ and $\widehat{\Lambda}_{n_2}$ could be any estimators. In the following, we will focus on U_n with them being the isotonic regression estimators of Λ_1 and Λ_2 given in Sun and Kalbfleisch (1995) and Wellner and Zhang (2000). To introduce the isotonic regression estimator, for simplicity, assume that H_0 is true and let $\Lambda(t)$ denote the common mean function of the $N_i(t)$'s. Also, 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 $\bar{n}_l$ the number and mean value, respectively, of the observations made at s_l, $l = 1, \ldots, m$. Then the isotonic regression estimator $\widehat{\Lambda}_n(t)$ is defined as a non decreasing step function with possible jumps at the s_l's that minimizes

$$\sum_{l=1}^{m} w_l \{\bar{n}_l - \Lambda(s_l)\}^2$$

subject to the nondecreasing restriction. It can be shown that $\widehat{\Lambda}_n(t)$ has a closed form given by

$$\widehat{\Lambda}_n(s_l) = \max_{r \le l} \min_{s \ge l} \frac{\sum_{v=r}^{s} w_v \bar{n}_v}{\sum_{v=r}^{s} w_v} = \min_{s \ge l} \max_{r \le l} \frac{\sum_{v=r}^{s} w_v \bar{n}_v}{\sum_{v=r}^{s} w_v}, \quad l = 1, \ldots, m$$

(Robertson et al., 1988).

We can rewrite the test statistic U_n as

$$U_n = \sqrt{\frac{n_1 n_2}{n^3}} \sum_{i=1}^{n} \sum_{j=1}^{k_i} W_n(t_{i,j}) \{\widehat{\Lambda}_{n_1}(t_{i,j}) - \widehat{\Lambda}_{n_2}(t_{i,j})\}.$$

That is, U_n is a Wilcoxon-type statistic. Similar approaches are often used in the analysis of repeated measurement data (Davis and Wei, 1988). For the selection of the weight process $W_n(t)$, a simple and natural choice is $W_n^{(1)}(t) = 1$. Another natural choice is $W_n^{(2)}(t) = Y_n(t) = \sum_{i=1}^{n} I(t \le t_{i,k_i})/n$. In this case, weights are proportional to the number of subjects under observation. One could also use

$$W_n^{(3)}(t) = \frac{Y_{n_1}(t)\, Y_{n_2}(t)}{Y_n(t)},$$

where $Y_{n_1}(t)$ and $Y_{n_2}(t)$ are defined as $Y_n(t)$ with the summation being over subjects only in the control and treatment groups, respectively. The weight processes similar to $W_n^{(3)}$ are commonly used when recurrent event data are observed (Andersen et al., 1993). In the next section, we will establish the asymptotic distribution of U_n.

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-valued random variable and $T = \{T_{k,j}, j = 1, \ldots, k, k = 1, 2, \ldots\}$ is a random triangular array and that the 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 $X = (K, T_K, N_K)$, 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 $X_i = (K_i, T_{K_i}, N_{i,K_i})$, $i = 1, \ldots, n$ are n i.i.d. copies of X. Some comments on this are given below. To establish the asymptotic results about $\widehat{\Lambda}_n(t)$ and U_n, we need the following regularity conditions.

A. The mean function Λ_0 satisfies that $\Lambda_0(\tau) \le M$ for some constant $M \in (0, \infty)$ and it is strictly increasing.
B. There exists a constant K_0 such that $Pr\{K \le K_0\} = 1$ and the random variables $T_{k,j}$'s take values in a bounded set $[0, \tau]$, where $\tau \in (0, \infty)$.

C. $Pr\{\limsup_{n\to\infty} \max_i N_i(\tau) < \infty\} = 1$ and $E(N_i(t))^2 \leq M_1$ for all $t \leq \tau$ where M_1 is a constant.

Let $\widehat{\Lambda}_n(t)$ denote the isotonic regression estimate of $\Lambda_0(t)$ under H_0 given in the previous section. First we give the asymptotic distribution of $\widehat{\Lambda}_n(t)$.

Theorem 3.1 (Asymptotic Normality of Functional of $\widehat{\Lambda}_n$). *Suppose that conditions A, B, and C hold. Also suppose that $W(t)$ is a bounded weight process such that $W \circ \Lambda_0^{-1}$ is a bounded Lipschitz function. Let $G(t) = E\big(\sum_{j=1}^{K} 1_{\{T_{K,j}\leq t\}}\big)$. Then as $n \to \infty$,*

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

in distribution, where U_w has a normal distribution with mean zero and variance that can be consistently estimated by

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

Now we are ready to establish the asymptotic distribution of U_n. Let S_l denote the set of indices for subjects in group l, $l = 1, 2$.

Theorem 3.2. *Suppose that conditions A, B, and C hold. Also suppose that $W_n(t)$ is a bounded weight process and that there exists a bounded function $W(t)$ such that $W \circ \Lambda_0^{-1}$ is a bounded Lipschitz function and*

$$\sup_n E\int_0^\tau |\sqrt{n}\{W_n(t) - W(t)\}|^2 dG_n(t) < \infty.$$

Also suppose that $n_1/n \to p_1$ and $n_2/n \to p_2$ as $n \to \infty$, where $0 < p_1,\ p_2 < 1$ and $p_1 + p_2 = 1$. Then under $H_0 : \Lambda_1 = \Lambda_2 = \Lambda_0$, U_n has an asymptotic normal distribution with mean zero and variance that can be consistently estimated by

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

where

$$\hat{\sigma}_l^2 = \frac{1}{n_l}\sum_{i\in S_l}\left[\sum_{j=1}^{k_i} W_n(t_{k_i,j})\big\{N_i(t_{k_i,j}) - \widehat{\Lambda}_{n_l}(t_{k_i,j})\big\}\right]^2,$$

$l = 1, 2$.

The proof of the above theorems is sketched in the Appendix. It can be easily shown that the three weight processes discussed in Sec. 2 satisfy the conditions required by the theorem. By applying Theorem 3.2, the test of the hypothesis H_0 can be carried out using the statistic $U_n^* = U_n/\hat{\sigma}_U$ based on the standard normal distribution.

Table 1
Estimated sizes and powers for Poisson processes

	β				
Test statistic	−0.2	−0.1	0.0	0.1	0.2
U_{SF}	0.919	0.395	0.048	0.431	0.955
U_n with $W_n^{(1)}$	0.923	0.410	0.051	0.441	0.958
U_n with $W_n^{(2)}$	0.908	0.393	0.053	0.430	0.948
U_n with $W_n^{(3)}$	0.908	0.393	0.053	0.428	0.948

4. Numerical Studies

In this section we report some results obtained from simulation studies conducted for investigating the finite sample properties of the proposed test statistic U_n. To generate panel count data, we mimic medical follow-up studies such as the example discussed in the next section. In these situations, subjects are usually prescheduled to be examined at prespecified time points for a prespecified number of times, but the actual numbers of examinations and examination times may vary. In particular, we first generated K_i, the number of observation times, from the uniform distribution $U\{1, \ldots, c\}$ with $c = 10$ or 40 and given K_i, observation times t_{ij}'s were then also generated from $U\{1, \ldots, c\}$ for simplicity. Note that one could generate the t_{ij}'s from more general uniform distributions and the results should be similar.

For the N_i's, we assume that they are non-homogeneous Poisson or mixed Poisson processes. In particular, for given t_{ij}'s and v_i defined below, we suppose that $N_i(t_{ij})$ follows a Poisson distribution with case I: $\Lambda_1(t_{ij}) = v_i t_{ij}$ and $\Lambda_2(t_{ij}) = v_i t_{ij} \exp(\beta)$ or case II: $\Lambda_1(t_{ij}) = v_i 5.5 t_{i,j}^{1/2}$ and $\Lambda_2(t_{ij}) = v_i t_{i,j}$, where β is a parameter representing the difference between the two groups. For the Poisson process, all v_i's were set to be equal to 1 and for the mixed Poisson process, the v_i's were generated from a Gamma distribution with mean one and variance 0.25. The results reported below are based on 1,000 replications.

Table 1 presents the estimated sizes and powers of the proposed test statistic U_n at significance level $\alpha = 0.05$ for case I with $\beta = -0.2, -0.1, 0, 0.1$, and 0.2, $c = 10$ and $n_1 = n_2 = 100$. Here we considered all three weight processes suggested in Sec. 2 and assumed that the N_i's are Poisson processes. For comparison, we also calculated the estimated sizes and powers of the test procedure given in Sun and

Table 2
Estimated sizes and powers for mixed Poisson processes

	β				
Test statistic	−0.2	−0.1	0.0	0.1	0.2
U_{SF}	0.481	0.185	0.047	0.175	0.502
U_n with $W_n^{(1)}$	0.502	0.196	0.051	0.185	0.512
U_n with $W_n^{(2)}$	0.494	0.182	0.049	0.187	0.510
U_n with $W_n^{(3)}$	0.494	0.181	0.049	0.187	0.509

Table 3
Estimated powers with different sample sizes

Test statistic	Sample size			
	$n_1 = n_2 = 30$	$n_1 = n_2 = 50$	$n_1 = n_2 = 100$	$n_1 = n_2 = 200$
U_{SF}	0.457	0.650	0.928	0.999
U_n with $W_n^{(1)}$	0.611	0.676	0.933	1.000
U_n with $W_n^{(2)}$	0.714	0.812	0.983	1.000
U_n with $W_n^{(3)}$	0.715	0.812	0.984	1.000

Fang (2003) and included them in the table with $W_n^{(1)}(t)$. It can be seen that for the situation considered here, all four tests gave reasonable sizes and their powers are close to each other and the best performance was given by U_n with the weight process $W_n^{(1)}(t)$.

Table 2 studied the same situation as in Table 1 except that the $N_i(t)$'s are mixed Poisson processes. The results are similar to those given in Table 1. As expected, compared to Table 1, the power decreases when more variability exists. To see the effect of sample sizes on the power, Table 3 presented the estimated powers of the above four test procedures at significance level $\alpha = 0.05$ for case II with $c = 40$, $v_i = 1$ and $n_1 = n_2 = 30$, 50, 100, or 200. As expected, the power increases when the sample size increases and as seen in Tables 1 and 2, the proposed test with $W_n^{(1)}(t)$ has better performance than that based on U_{SF}. Two things that are different from the results given in Tables 1 and 2 are that the power of the proposed method could

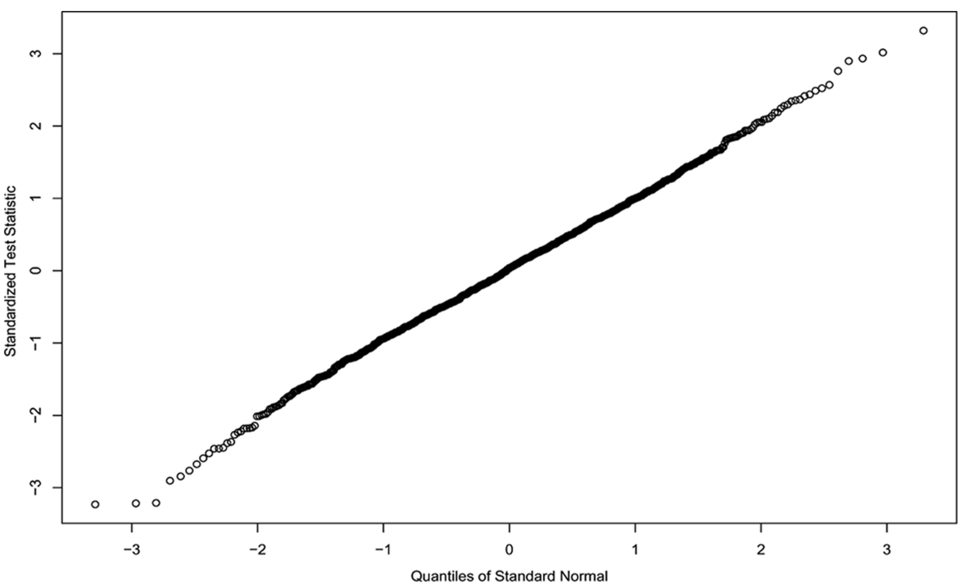

Figure 1. Poisson processes with $\beta = 0.0$.

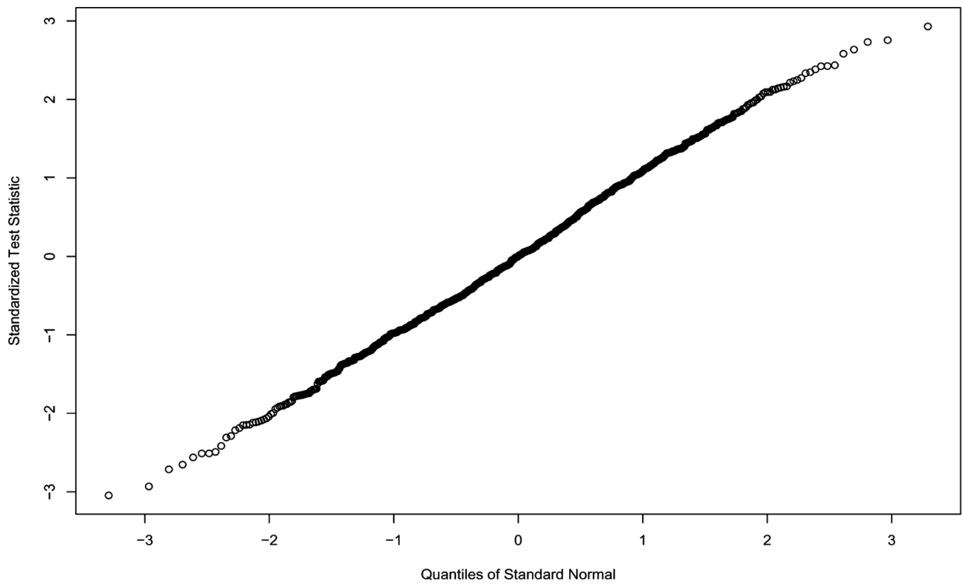

Figure 2. Mixed Poisson processes with $\beta = 0.0$.

be significantly higher than that of the method given in Sun and Fang (2003) and the weight processes $W_n^{(2)}(t)$ and $W_n^{(3)}(t)$ are competitive with the weight processes $W_n^{(1)}(t)$ depending on situations.

To evaluate the normal distribution approximation to the distribution of U_n, we studied the quantile plots of the standardized test statistic U_n against the standard normal distribution. Figures 1 and 2 display the plots for the situations considered in Tables 1 and 2 with $W_n(t) = W_n^{(1)}(t)$ and $\beta = 0$, respectively, and suggest that the approximation seems good. Similar plots were obtained for other situations.

5. An Analysis of the National Cooperative Gallstone Study

This section applies the proposed test in the previous sections to the floating gallstones study discussed above. The observed data are given in Table 1 of Thall and Lachin (1988) and 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 whole study consists of 916 patients who were randomized to placebo, low dose, or high dose group and followed for up to two years. As mentioned earlier, the data include the successive visit-times in study weeks and the associated counts of episodes of nausea. As many longitudinal follow-up studies, at the beginning of the study, all patients were scheduled to have clinical visits or be examined at 1, 2, 3, 6, 9, and 12 months. Again as these studies, the actual visit times differ from patient to patient. Most of the patients visited about six times within the first year, but there exist some patients who had only one visit and others who had nine visits. The goal here is to compare the effects of the two treatments on the incidence rate of nausea.

For the comparison, we treat the placebo group as group 1 ($\Lambda_1(t)$) and the high-dose chenodiol group as group 2 ($\Lambda_2(t)$). The application of the proposed method to the data gave $U_n^* = -0.748$ with $W_n(t) = W_n^{(1)}(t)$. This gives a p-value of 0.454 for testing the equality of the two mean functions. Letting $W_n(t) = W_n^{(2)}(t)$ or $W_n(t) = W_n^{(3)}(t)$, we obtained $U_n^* = -0.814$ or -0.818, respectively. They gave similar p-values. The above results indicate that the overall incidence rates of nausea were not significantly different for the patients in the two treatment groups. In comparison, the use of the approach in Sun and Fang (2003) gave a p-value of 0.1428. Note that the method given in Sun and Fang (2003) requires that the group indicators can be regarded as independent and identically distributed variables, which may not be true given the sample size difference between the two groups. For the problem, Thall and Lachin (1988) partitioned the observation period (one year) into six intervals and gave a quite smaller p-value. This is expected since as shown in Thall and Lachin (1988) and Sun and Kalbfleisch (1995), the two underlying mean functions seem to be overlapping and the method given by Thall and Lachin (1988) basically adds the differences over different intervals together.

6. Concluding Remarks

This article discusses the two sample comparison problem of point processes when only panel count data are available and a class of nonparametric tests for the problem is proposed. Simulation studies are conducted and suggest that the proposed method works well for practical situations. The presented approach applies to more general situations than the existing methods (Sun and Fang, 2003; Thall and Lachin, 1988). Another difference between the method given here and the test statistic presented in Sun and Fang (2003) is that the former can be generalized to general K-sample comparison problem, while the latter cannot. In addition to the presented test procedure, some asymptotic results are also established for the isotonic regression estimate of the mean function of the underlying point process, the estimate commonly used in this situation.

Throughout the article we have assumed that observation times follow the same distribution for subjects in the two treatment groups for two reasons. One is that this seems to be the case for the example discussed in Sec. 5 and actually holds for most of medical studies with periodic follow-up such as clinical trials; the other reason is the fact that the distributions of observation times cannot be allowed to be completely different between the two treatment groups. This can be seen through a simple example. Suppose that all observation times for subjects in one group are smaller than these for subjects in the other group. Then it is likely that no comparison can be made for the two groups.

One possible direction for future research is to replace in the statistic U_n isotonic regression estimates with maximum likelihood estimates for the mean function. A possible advantage could be the gain of efficiency assuming that the maximum likelihood estimate is more efficient than the isotonic regression estimate. Also it should be noted that unlike the isotonic regression estimate, the maximum likelihood estimate has no closed form and its determination needs a great deal of computational effort. Furthermore, there is very limited research available discussing its asymptotic properties.

Appendix: Proofs

A.1. Proof of Theorem 3.1

First note that

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

where

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

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

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, $P_nf = \frac{1}{n}\sum_{i=1}^n f_i$ 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 $\hat{\sigma}_w^2$. Thus 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 the condition C implies

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

almost surely. So, for every $\varepsilon > 0$, there exists a constant $M_\varepsilon > \Lambda_0(\tau)$ such that

$$\sup_n Pr\{\widehat{\Lambda}_n(\tau) > M_\varepsilon\} < \varepsilon.$$

Let

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

and

$$\mathscr{F}_\varepsilon = \{\Lambda : \Lambda \in \mathscr{F}, \Lambda(\tau) \le M_\varepsilon\}.$$

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

$$\widehat{\Lambda}_{n,\varepsilon} = \arg\max_{\Lambda\in\Omega\cap\mathscr{F}_\varepsilon}\left\{\sum_{i=1}^{n}\sum_{j=1}^{K_i}[N_i(T_{K_i,j})\log\Lambda(T_{K_i,j})-\Lambda(T_{K_i,j})]\right\}$$

where Ω is the class of non decreasing step functions with possible jumps only at the observation time points $\{T_{K_i,j}, j=1,\ldots,K_i, i=1,\ldots,n\}$. It is equivalent to $\widehat{\Lambda}_{n,\varepsilon}$

in Sun and Fang (2003) which is proved in Wellner and Zhang (2000). Let $I_{1n,\varepsilon}$ denote the version of I_{1n} obtained by replacing $\widehat{\Lambda}_n$ with $\widehat{\Lambda}_{n,\varepsilon}$. Then, to prove that I_{1n} converges to zero in probability, it is sufficient to show that $I_{1n,\varepsilon} = o_p(1)$ since

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

By using arguments similar to those in Sun and Fang (2003), it can be shown that $I_{1n,\varepsilon} = o_p(1)$. Now we show the convergence of I_{2n}. Let the s_l's, w_l's, and $\overline{N}_l = \bar{n}_l$ be defined as in Sec. 2. Using the same block argument as these of Proposition 1.2 in Part II of Groeneboom and Wellner (1992), we have that for any real function f,

$$\sum_{l=1}^{m} f(\widehat{\Lambda}_n(s_l))w_l[\widehat{\Lambda}_n(s_l) - \overline{N}_l] = 0.$$

Hence we can rewrite I_{2n} as

$$\begin{aligned} I_{2n} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \{W(T_{K_i,j}) - W \circ \Lambda_0^{-1}(\widehat{\Lambda}_n(T_{K_i,j}))\}\{\widehat{\Lambda}_n(T_{K_i,j}) - N_i(T_{K_i,j})\} \\ &= \sqrt{n}P_n\left\{\sum_{j=1}^{K}[W(T_{K,j}) - W \circ \Lambda_0^{-1}(\widehat{\Lambda}_n(T_{K,j}))][\widehat{\Lambda}_n(T_{K,j}) - N(T_{K,j})]\right\} \\ &= \Delta_{1n} + \Delta_{2n} + \Delta_{3n}, \end{aligned}$$

where

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

$$\Delta_{2n} = \sqrt{n}P\left\{\sum_{j=1}^{K}[W(T_{K,j}) - W \circ \Lambda_0^{-1}(\widehat{\Lambda}_n(T_{K,j}))][\widehat{\Lambda}_n(T_{K,j}) - \Lambda_0(T_{K,j})]\right\},$$

and

$$\Delta_{3n} = \sqrt{n}P\left\{\sum_{j=1}^{K}[W(T_{K,j}) - W \circ \Lambda_0^{-1}(\widehat{\Lambda}_n(T_{K,j}))][\Lambda_0(T_{K,j}) - N(T_{K,j})]\right\} = 0.$$

For Δ_{1n}, let $W_0 = W \circ \Lambda_0^{-1}$ and $\Delta_{1n,\varepsilon}$ denote the version of Δ_{1n} obtained by replacing $\widehat{\Lambda}_n$ with $\widehat{\Lambda}_{n,\varepsilon}$. Since W_0 is a bounded Lipschitz function, then it can be shown that

$$\mathscr{H}_\varepsilon = \left\{\sum_{j=1}^{K}[W_0(\Lambda_0(T_{K,j})) - W_0(\Lambda(T_{K,j}))][\Lambda(T_{K,j}) - N(T_{K,j})] : \Lambda \in \mathscr{F}_\varepsilon\right\}$$

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) gives that

$$d(\widehat{\Lambda}_{n,\varepsilon}, \lambda_0) \leq d(\widehat{\Lambda}_n, \lambda_0) \longrightarrow 0$$

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 uniformly 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 that $\Delta_{1n} = o_p(1)$ since

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

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

$$\begin{aligned} |\Delta_{2n}| &= \left| \sqrt{n} P\left\{ \sum_{j=1}^K \left[W_0(\Lambda_0(T_{K,j})) - W_0(\widehat{\Lambda}_n(T_{K,j}))\right]\left[\widehat{\Lambda}_n(T_{K,j}) - \Lambda_0(T_{K,j})\right] \right\} \right| \\ &= \left| \sqrt{n} \int_0^\tau \left[W_0(\Lambda_0(t) - W_0(\widehat{\Lambda}_n(t))\right]\left[\widehat{\Lambda}_n(t) - \Lambda_0(t)\right] dG(t) \right| \leq c_1 \sqrt{n} d^2(\widehat{\Lambda}_n, \Lambda_0) \end{aligned}$$

where c_1 is a constant. To prove that $\sqrt{n} d^2(\widehat{\Lambda}_n, \Lambda_0) = o_p(1)$, we only need to show that $\sqrt{n} d^2(\widehat{\Lambda}_{n,\varepsilon}, \Lambda_0) = o_p(1)$. As shown below, $d(\widehat{\Lambda}_{n,\varepsilon}, \Lambda_0) = O_p(n^{-\frac{1}{3}})$. This shows that $\Delta_{2n} = o_p(1)$ and thus completes the proof.

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

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

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

$$\begin{aligned} \mathbb{M}(\Lambda_0) - \mathbb{M}(\Lambda) &= P\left(\sum_{j=1}^K \Lambda_0(T_{K,j}) \log \frac{\Lambda_0(T_{K,j})}{\Lambda(T_{K,j})} - \left(\frac{\Lambda_0(T_{K,j})}{\Lambda(T_{K,j})} - 1 \right) \Lambda(T_{K,j}) \right) \\ &= P\left(\sum_{j=1}^K \Lambda(T_{K,j}) h\left(\frac{\Lambda_0(T_{K,j})}{\Lambda(T_{K,j})} \right) \right) = \int \Lambda(t) h\left(\frac{\Lambda_0(t)}{\Lambda(t)} \right) dG(t) \\ &\geq \frac{1}{5} \int \frac{(\Lambda_0(t) - \Lambda(t))^2}{\Lambda(t)} dG(t) \geq \frac{1}{5M_\varepsilon} d^2(\Lambda, \Lambda_0), \end{aligned}$$

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

$$\mathscr{F}_{\delta,\varepsilon} = \{\Lambda : d(\Lambda, \Lambda_0) \leq \delta, \Lambda \in \mathscr{F}_\varepsilon\} \ (\delta > 0), \quad \mathscr{M}_{\delta,\varepsilon} = \{m_\Lambda(X) - m_{\Lambda_0}(X) : \Lambda \in \mathscr{F}_{\delta,\varepsilon}\}.$$

Since we have that

$$\log N_{[]}(\eta, \mathscr{M}_{\delta,\varepsilon}, L_2(P)) \leq c_2 \eta^{-1}$$

where c_2 is a constant which depends only on M_ε, then

$$\int_0^\delta \sqrt{1+\log N_{[]}\big(\eta, \mathscr{M}_{\delta,\varepsilon}, L_2(P)\big)}d\eta \le \int_0^\delta \sqrt{1+c_2\eta^{-1}}\, d\eta$$
$$= 2c_2 \int_{\sqrt{1+c_2\delta^{-1}}}^\infty \frac{u^2}{(u^2-1)^2}\, du$$
$$\le c_3 \int_{\sqrt{1+c_2\delta^{-1}}}^\infty \frac{1}{u^2}\, du \le \frac{c_3}{\sqrt{c_2}}\delta^{\frac{1}{2}}$$

Hence, applying Lemma 3.4.2 of Van der Vaart and Wellner (1996), we have that

$$E^*\|\sqrt{n}(P_n - P)\|_{\mathscr{M}_{\delta,\varepsilon}} \le c_4\phi_n(\delta)$$

where E^* denotes the outer expectation, and $\phi_n(\delta) = \delta^{\frac{1}{2}} + \delta^{-1}n^{-\frac{1}{2}}$. Using Theorem 3.2.5 of Van der Vaart and Wellner (1996), $d\big(\widehat{\Lambda}_{n,\varepsilon}, \Lambda_0\big)$ converges in probability to zero of order at least $n^{-\frac{1}{3}}$.

A.2. Proof of Theorem 3.2

Let S_1 and S_2 be defined as before and define

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

$l = 1, 2$. To see the asymptotic distribution of U_n, note that we can rewrite U_n as

$$U_n = \sqrt{\frac{n_2}{n}}U_n^{(1)} - \sqrt{\frac{n_1}{n}}U_n^{(2)},$$

where

$$U_n^{(l)} = \sqrt{n_l}\int_0^\tau W_n(t)\big\{\widehat{\Lambda}_{n_l}(t) - \Lambda_0(t)\big\}dG_n(t),$$

$l = 1, 2$. Also note that $U_n^{(l)} = I_{1n}^{(l)} + I_{2n}^{(l)} + I_{3n}^{(l)}$ where

$$I_{1n}^{(l)} = \sqrt{n_l}\int_0^\tau \{W_n(t) - W(t)\}\big\{\widehat{\Lambda}_{n_l}(t) - \Lambda_0(t)\big\}dG_n(t),$$

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

$$I_{3n}^{(l)} = \sqrt{n_l}\int_0^\tau W(t)\big\{\widehat{\Lambda}_{n_l}(t) - \Lambda_0(t)\big\}dG(t).$$

First we show that $I_{1n}^{(l)} = o_p(1)$, $l = 1, 2$. Using Cauchy–Schwarz inequality and the proof of Theorem 3.1, we have that

$$
\begin{aligned}
&E\{|I_{1n}^{(l)}|1_{\{\widehat{\Lambda}_{n_l}(\tau)\le M_\varepsilon\}}\} \\
&\le E\left\{\sqrt{n_l}\left[\int_0^\tau \left(W_n(t) - W(t)\right)^2 dG_n(t)\right]^{1/2}\left[\int_0^\tau \left(\widehat{\Lambda}_{n_l,\varepsilon}(t) - \Lambda_0(t)\right)^2 dG_n(t)\right]^{1/2}\right\} \\
&\le \left\{E\int_0^\tau \left[\sqrt{n}(W_n(t) - W(t))\right]^2 dG_n(t)\right\}^{1/2}\left\{E\left[\int_0^\tau \left(\widehat{\Lambda}_{n_l,\varepsilon}(t) - \Lambda_0(t)\right)^2 dG_n(t)\right]\right\}^{1/2} \\
&\longrightarrow 0
\end{aligned}
$$

since

$$\int_0^\tau \left(\widehat{\Lambda}_{n_l,\varepsilon}(t) - \Lambda_0(t)\right)^2 dG_n(t) = o_p(1)$$

and it is bounded. So, $I_{1n}^{(l)} = o_p(1)$, $l = 1, 2$. As the proof of Theorem 3.1, it can be shown that $I_{2n}^{(l)} = o_p(1)$, $l = 1, 2$. Also, it follows from Theorem 3.1 that

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

$l = 1, 2$. Hence, we have that $U_n^{(l)}$ converges in distribution to random variable $U_w^{(l)}$ that has a normal distribution with mean zero and variance that can be estimated by $\hat{\sigma}_l^2$ given in Theorem 3.2, $l = 1, 2$. This proves the theorem since $U_w^{(1)}$ and $U_w^{(2)}$ are independent.

References

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

Cook, R. J., Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statist. Med.* 16:911–924.

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

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

Huang, J., Wellner, J. A. (1995). Asymptotic normality of the NPMLE of linear functionals for interval 5censored data, case I. *Statistica Neerlandica* 49:153–163.

Kalbfleisch, J. D., Lawless, J. F. (1985). The analysis of panel data under a markov assumption. *J. Amer. Statist. Assoc.* 80:863–871.

Lawless, J. F., Nadeau, J. C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics* 37:158–168.

Lin, D. Y., Wei, J. L., Yang, I., Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *J. Roy. Statist. Soc. B* 62:711–730.

Pepe, M. S., Fleming, T. R. (1989). Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. *Biometrics* 45:497–507.

Petroni, G. R., Wolfe, R. A. (1994). A two sample test for stochastic ordering with interval-censored data. *Biometrics* 50:77–87.

Robertson, T., Wright, F. T., Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. New York: Wiley.

Schoenfield, L. J., Lachin, J. M. (1981). National cooperative gallstone study: a controlled trial of the efficiency and safety of chenodeoxycholic acid for dissolution of gallstones. *Ann. Internal Med.* 95(3):257–282.

Sun, J., Kalbfleisch, J. D. (1995). Estimation of the mean function of point processes based on panel data. *Statistica Sinica* 5:279–289.

Sun, J., Wei, L. J. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *J. Roy. Statist. Soc. B* 62:293–302.

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

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

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

Wang, M. C., Chang, S. H. (1999). Nonparametric estimation of a recurrent survival function. *J. Amer. Statist. Assoc.* 94:146–153.

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

Zhang, Y. (2002). A semiparametric pseudolikelihood estimation method for panel count data. *Biometrika* 89:39–48.

Zhang, Y., Liu, W., Zhan, Y. (2001). A nonparametric two-sample test of the failure function with interval censoring case 2. *Biometrika* 88:677–686.