Statistica Sinica **9**(1999), 211-227

A CLASS OF NONPARAMETRIC K-SAMPLE TESTS FOR SEMI-MARKOV COUNTING PROCESSES

I-Shou Chang, Yuan-Chuan Chuang and Chao A. Hsiung*

National Central University, Ming-Chuan University and
**National Health Research Institute and Academia Sinica*

Abstract: A K-sample testing problem is studied for multivariate semi-Markov counting processes. Asymptotic distributions and efficiency of a class of nonparametric test statistics are established for certain local alternatives. The concept of the asymptotic efficiency states that for every nonparametric test in this class, there is a parametric submodel for which the optimal test has the same asymptotic power as the nonparametric test. The theory is illustrated by a simulation study and by analyzing the multivariate failure time data of Thompson et al. (1978).

Key words and phrases: Asymptotic distribution, asymptotic efficiency, K-sample test, parametric submodel, semi-Markov counting process.

1. Introduction

Multivariate failure time data arises from studies involving the recording of times to two or more distinct events or "failures" on an individual subject. The failure may be repetitions of essentially the same event or may be events of entirely different types. Interesting multivariate failure time data was reported by Thompson et al. (1978) from an experimental animal carcinogenesis study. In order to analyze these data, Gail, Santner and Brown (1980) proposed several models and tests for the comparison of multivariate failure time data arising from two treatment groups.

Two of the important models studied by Gail et al. (1980) are the m-site model and the semi-Markov model. A careful analysis shows that the K-sample problem for the m-site model provides an excellent example of the K-sample problem for the multiplicative intensity model of counting processes. This was later studied by, among others, Gill (1980), Hjort (1985), and Andersen, Borgan, Gill and Keiding (1982), (1993), (henceforth ABGK (1982) and ABGK (1993), respectively). The purpose of this paper is to present an analogous theory for the K-sample problem for semi-Markov model of counting processes. In particular, we will propose a class of nonparametric test statistics, derive their asymptotic distributions, and establish their asymptotic efficiency. It is expected that this theory will provide a useful alternative for analyzing multivariate failure time data.

In order to facilitate the discussion, we need to introduce the following notation. For $k = 1, \ldots, K\,(K \geq 2)$, $j = 1, \ldots, J_k$, let $N_{jk}(t)$ denote the number of events experienced up to time t by the jth individual in the kth experimental group. Assume that the intensity $\lambda_{jk}(\cdot)$ of $N_{jk}(\cdot)$ is of the semi-Markov form

$$\lambda_{jk}(t) = \Big(\sum_{i=0}^{\infty} h_{ki}(t - T_{jki}) 1_{(T_{jki}, T_{jk(i+1)}]}(t)\Big) Y_{jk}(t), \tag{1.1}$$

relative to a filtration $\mathcal{F}_{jk,t}$, for $k = 1, \ldots, K$, $j = 1, \ldots, J_k$. Here $Y_{jk}(\cdot)$ is a bounded predictable process, $h_{ki}(\cdot)$ is a nonnegative deterministic function, and $T_{jki} = \inf\{t > 0 | N_{jk}(t) = i\}$. We also assume throughout the paper that for every fixed t, the σ-fields $\mathcal{F}_{jk,t}$, for $k = 1, \ldots, K$, $j = 1, \ldots, J_k$, are independent.

Let H_0 denote the null hypothesis that $h_{1i}(\cdot) = \cdots = h_{Ki}(\cdot)$ for every $i = 0, 1, 2, \ldots$. Let H_1 denote the alternative hypothesis that $h_{li}(\cdot) \neq h_{ki}(\cdot)$ for some $l \neq k$, and some $i = 0, 1, 2, \ldots$.

We point out that (1.1) formalizes and generalizes the semi-Markov model discussed in Gail et al. (1980), where h_{ki} was assumed to be a constant. We note also that, assuming $K = 2$ and $h_{2i}(t) = h_{1i}(t)e^{\alpha_i}$ for some constant α_i, Gail et al. (1980) proposed tests for the hypothesis H_0. There are several ways to incorporate covariates in (1.1) for study of the treatment effect. See for example Cox (1986), Commenges (1986), Chang and Hsiung (1994) and Chang (1995). In particular, Chang and Hsiung (1994) established the asymptotic normality and efficiency of a Cox-type estimator proposed by Prentice, Williams and Peterson (1981).

The statistical problem of interest is to test the hypothesis H_0 based on the data $\{N_{jk}(t), Y_{jk}(t) |\, k = 1, \ldots, K,\ j = 1, \ldots, J_k,\ 0 \leq t \leq t_0\}$.

The plan of this paper is as follows. Section 2 presents a class of nonparametric test statistics for the null hypothesis H_0. Section 3 establishes the asymptotic distributions of the test statistics under both null hypothesis and local alternatives by making use of the Martingale Central Limit Theorem and Le Cam's third lemma. Although this is a standard approach, additional work is needed to introduce a discrete time filtration and exhibit certain martingale structures in it so as to present an approximation useful in the derivation of asymptotic distributions.

Section 4 establishes the asymptotic efficiency of the nonparametric test statistics. We will introduce a class of parametric submodels and show that each of the nonparametric test statistics is asymptotically equivalent to the asymptotically optimal test statistics in one of the parametric submodels. It seems that this approach to the asymptotic efficiency of a nonparametric test statistic for counting process appeared only in ABGK (1993), which deals with the multiplicative intensity model.

Finally, in Section 5, we illustrate our theory by a simulation study and by analyzing the multivariate failure time data of Thompson et al. (1978).

Readers interested in identifying data for which this semi-Markov model is needed are encouraged to consult Chang et al. (1997), which provides goodness-of-fit tests for certain semi-Markov counting process.

2. A Class of Nonparametric Tests

Let $(\Omega^{(\mathbf{J})}, \mathcal{F}^{(\mathbf{J})}, P^{(\mathbf{J})})$ be a probability space such that for every $k = 1, \ldots, K$, $N_{1k}(\cdot), \ldots, N_{J_k k}(\cdot)$ is an i.i.d. sequence of counting processes defined on it, and (1.1) is satisfied for every $N_{jk}(\cdot)$. Here $\mathbf{J} = (J_1, \ldots, J_K)$. We note that the assumption made in the lines following (1.1) implies that $\{N_{jk}(\cdot)|j = 1, \ldots, J_k, k = 1, \ldots, K\}$ is a family of independent processes.

When H_0 is true, we let $h_i(\cdot) = h_{ki}(\cdot)$ and denote by $P_0^{(\mathbf{J})}$ the corresponding probability measure.

Let $M_{jk}(t) = N_{jk}(t) - \int_0^t \sum_{i=0}^{\infty} h_i(s - T_{jki}) 1_{(T_{jki}, T_{jk(i+1)}]}(s) Y_{jk}(s) ds$, which is a martingale under H_0, for every $k = 1, \ldots, K$, $j = 1, \ldots, J_k$.

Let $t > 0$. Let $\mathcal{F}_{jk, T_{jki}+t}$ denote the history of $N_{jk}(\cdot)$ up to the time $T_{jki} + t$. Let S be a stopping time relative to the filtration $\{\mathcal{F}_{jk, T_{jki}+t} | t \geq 0\}$. Since $T_{jk(i+1)} - T_{jki}$ is an $\mathcal{F}_{jk, T_{jki}+t}$-stopping time, so is $S \wedge (T_{jk(i+1)} - T_{jki})$. Here $a \wedge b$ is the minimum of a and b. This shows that $T_{jki} + S \wedge (T_{jk(i+1)} - T_{jki})$ is an $\mathcal{F}_{jk,t}$-stopping time. With this in mind, we know that $\mathrm{E}\, M_{jk}((T_{jki} + S \wedge (T_{jk(i+1)} - T_{jki})) \wedge t_0) = 0$. This implies that

$$M_{jki}(t) \equiv M_{jk}((T_{jki} + t \wedge (T_{jk(i+1)} - T_{jki})) \wedge t_0) - M_{jk}(T_{jki} \wedge t_0) \qquad (2.1)$$

is an $\mathcal{F}_{jk, T_{jki}+t}$-martingale. (cf. Chang and Hsiung (1994)).

The relation (2.1) suggests a multiplicative intensity model as follows. Let $N_{jki}(t) = N_{jk}((T_{jki} + t \wedge (T_{jk(i+1)} - T_{jki})) \wedge t_0) - N_{jk}(T_{jki} \wedge t_0)$, and let $Y_{jki}(t) = Y_{jk}(T_{jki} + t) 1_{(T_{jki}, T_{jk(i+1)}]}(T_{jki} + t) 1_{(0, t_0]}(T_{jki} + t)$. A straightforward calculation shows that

$$M_{jki}(t) = N_{jki}(t) - \int_0^t h_i(s) Y_{jki}(s) ds. \qquad (2.2)$$

It follows from (2.1) and (2.2) that $\{N_{jki}(\cdot)|\, k = 1, \ldots, K,\, j = 1, 2, \ldots, J_k\}$ form a multiplicative intensity model for every given i, relative to the filtration $\mathcal{G}_t^{(\mathbf{J},i)}$, the σ-field generated by $\mathcal{F}_{jk, T_{jki}+t}$ for $k = 1, \ldots, K$, $j = 1, \ldots, J_k$.

The above observation based on a random time change technique for semi-Markov counting process was noted by Prentice, Williams and Peterson (1981) and Voelkel and Crowley (1984).

Following ABGK (1982), we define

$$Z_l^{(\mathbf{J},i)}(t) = J_{\cdot}^{-1/2} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \int_0^t G_i^{(\mathbf{J})}(s) \Big(\delta_{lk} - \frac{Y_{\cdot li}^{(J_l)}(s)}{Y_{\cdot\cdot i}^{(\mathbf{J})}(s)} \Big) dN_{jki}(s). \qquad (2.3)$$

Since

$$Z_l^{(\mathbf{J},i)}(t) = J_\cdot^{-1/2} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \int_0^t G_i^{(\mathbf{J})}(s) \Big(\delta_{lk} - \frac{Y_{\cdot li}^{(J_l)}(s)}{Y_{\cdot\cdot i}^{(\mathbf{J})}(s)} \Big) dM_{jki}(s),$$

we know that (2.3) is a martingale under H_0. Here $J_\cdot = J_1 + \cdots + J_K$, $G_i^{(\mathbf{J})}(\cdot)$ is a $\mathcal{G}_t^{(\mathbf{J},i)}$-predictable process, δ_{lk} is 1 if $l = k$, and is 0 otherwise, $Y_{\cdot li}^{(J_l)}(\cdot) = \sum_{j=1}^{J_l} Y_{jli}(\cdot)$, $Y_{\cdot\cdot i}^{(\mathbf{J})}(\cdot) = \sum_{k=1}^{K} \sum_{j=1}^{J_k} Y_{jki}(\cdot)$ and $\frac{0}{0} \equiv 0$.

Let $\mathbf{Z}^{(\mathbf{J},i)}(\cdot) = (Z_1^{(\mathbf{J},i)}(\cdot), \ldots, Z_K^{(\mathbf{J},i)}(\cdot))'$. Let i be given. When $J_1 = \cdots = J_K = J$, ABGK (1982, 1993) showed that, in the problem of testing the hypothesis $h_{1i}(\cdot) = \cdots = h_{Ki}(\cdot)$ based on $\{(N_{jki}(t), Y_{jki}(t)) | 0 \le t \le t_0, k = 1, \ldots, K, j = 1, \ldots, J\}$, a particular test statistic based on $\mathbf{Z}^{(\mathbf{J},i)}(t_0)$ is asymptotically chi-squared under both null and alternative hypotheses. Moreover it is optimal in the sense that it is asymptotically equivalent to the best test statistic in a parametric submodel (cf. ABGK (1993), p. 615).

Let $\hat{\Sigma}^{(\mathbf{J})}(t)$ be the matrix whose (l, m)-entry is

$$\hat{\Sigma}_{lm}^{(\mathbf{J})}(t) = J_\cdot^{-1} \sum_{i=0}^{\infty} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \int_0^t \Big(G_i^{(J)}(s) \Big)^2 \frac{Y_{\cdot mi}(s)}{Y_{\cdot\cdot i}(s)} \Big(\delta_{lm} - \frac{Y_{\cdot li}(s)}{Y_{\cdot\cdot i}(s)} \Big) dN_{jki}(s),$$

which is symmetric and asymptotically equivalent to

$$J_\cdot^{-1} \sum_{i=0}^{\infty} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \int_0^t \Big(G_i^{(J)}(s) \Big)^2 \Big(\delta_{lk} - \frac{Y_{\cdot li}(s)}{Y_{\cdot\cdot i}(s)} \Big) \Big(\delta_{mk} - \frac{Y_{\cdot mi}(s)}{Y_{\cdot\cdot i}(s)} \Big) dN_{jki}(s).$$

Let $\hat{\Sigma}^{(\mathbf{J})^-}(t)$ be a generalized inverse of $\hat{\Sigma}^{(\mathbf{J})}(t)$. The statistic we propose for testing H_0 is

$$X^{(\mathbf{J})} \equiv \mathbf{Z}^{(\mathbf{J})}(t_0)' \hat{\sum}^{(\mathbf{J})^-}(t_0) \mathbf{Z}^{(\mathbf{J})}(t_0), \tag{2.4}$$

where $\mathbf{Z}^{(\mathbf{J})}(t_0) = \sum_{i=0}^{\infty} \mathbf{Z}^{(\mathbf{J},i)}(t_0)$.

We will reject H_0 if $X^{(\mathbf{J})} \ge C_\alpha$, where C_α is the $(1-\alpha)$th quantile of the central chi-squared distribution with degree of freedom $K - 1$. The local asymptotic power can be obtained by computing $P(X \ge C_\alpha)$, where X has a noncentral chi-squared distribution with degree of freedom $K - 1$ and noncentrality $\eta(t_0)'\Sigma(t_0)^-\eta(t_0)$, as given in Theorem 3.2. In practice, both η and Σ are to be estimated.

3. Asymptotic Distributions of the Test Statistics

In this section we first establish the asymptotic distributions of $\mathbf{Z}^{(\mathbf{J})}(\cdot)$ under both null hypothesis and local alternatives, from which we can derive the asymptotic distributions of $X^{(\mathbf{J})}$. However, since $\mathbf{Z}^{(\mathbf{J})}(\cdot)$ is not a martingale, we will introduce a discrete time filtration and exhibit certain martingale structures relative to this filtration and show that $\mathbf{Z}^{(\mathbf{J})}(\cdot)$ can be approximated by a martingale $\bar{\mathbf{Z}}^{(\mathbf{J})}(\cdot)$. With this, we can apply the Martingale Central Limit Theorem to obtain the asymptotic distribution of $\mathbf{Z}^{(\mathbf{J})}(\cdot)$ under null hypothesis. Its asymptotic distribution under local alternatives is established by introducing a suitable local model, specifying its likelihood and then applying Le Cam's third lemma.

In order to increase the readability of this section, we will present the results in Subsection 3.1 and give the technical proofs in Subsection 3.2.

3.1. Local models, likelihoods and asymptotic distributions

The asymptotics of this paper are developed for the local model defined by the bounded deterministic functions $\gamma_{ki}(\cdot)$, $k = 1, \ldots, K$, $i = 0, 1, 2, \ldots$, as follows.

Let $Q^{(\mathbf{J})}$ denote a probability measure on $(\Omega^{(\mathbf{J})}, \mathcal{F}^{(\mathbf{J})})$ so that $N_{1k}(\cdot), \ldots, N_{J_k k}(\cdot)$ are i.i.d. counting processes with the intensity $\lambda_{jk}(\cdot)$ of $N_{jk}(\cdot)$ equal to

$$\left\{\sum_{i=0}^{\infty} h_i(t - T_{jki})\Big(1 + J.^{-1/2}\gamma_{ki}(t - T_{jki})\Big)1_{(T_{jki}, T_{jk(i+1)}]}(t)\right\}Y_{jk}(t), \tag{3.1}$$

for $k = 1, \ldots, K$, $j = 1, \ldots, J_k$ and the $N_{jk}(\cdot)$'s independent. Here we require $1 + J.^{-1/2}\gamma_{jk}(\cdot) \geq 0$.

We note that the existence of the probability measure $Q^{(\mathbf{J})}$ follows from the existence of $P_0^{(\mathbf{J})}$, Doléans-Dade's exponential martingale theorem, and the direct Radon-Nikodym derivative theorem (cf. Brémaud (1981), p. 165–167). In fact, we have the following log-likelihood process $\mathcal{L}^{(\mathbf{J})}(\cdot)$ on $\mathcal{F}_t^{(\mathbf{J})}$, the σ-field generated by $\mathcal{F}_{jk,t}$, $k = 1, \ldots, K$, $j = 1, \ldots, J_k$:

$$\begin{aligned}\mathcal{L}^{(\mathbf{J})}(t) &\equiv \log \frac{dQ^{(\mathbf{J})}}{dP_0^{(\mathbf{J})}} \\ &= \sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t \sum_{i=0}^{\infty} \log(1 + J.^{-1/2}\gamma_{ki}(s - T_{jki}))1_{(T_{jki}, T_{jk(i+1)}]}(s)dN_{jk}(s) \\ &\quad - \sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t \sum_{i=0}^{\infty} h_i(s - T_{jki})J.^{-1/2}\gamma_{ki}(s - T_{jki})1_{(T_{jki}, T_{jk(i+1)}]}(s)Y_{jk}(s)ds.\end{aligned} \tag{3.2}$$

The following assumptions are needed for the main theorems in this paper.

(A1) $G_i^{(\mathbf{J})}(t)$ converges in probability to a deterministic function $g_i(t)$, as $J.$ goes to infinity, for every $i = 0, 1, 2, \ldots$ and every $t \in [0, t_0]$.

(A2) There exists a constant $M > 0$ such that $h_i(t)$, $|\gamma_{ki}(t)|$, $|G_i^{(\mathbf{J})}(t)|$, $|g_i(t)|$, and $Y_{jk}(t)$ are all bounded by M for every $i = 0, 1, 2 \ldots$, $k = 1, \ldots, K$, $J_k = 1, 2, \ldots$, $j = 1, \ldots, J_k$ and $t \in [0, t_0]$.

(A3) For every $k = 1, \ldots, K$, $j = 1, \ldots, J_k$, $N_{jk}(\cdot)$ is non-explosive, i.e., $\mathrm{E}\,(t_0 - T_{jki})^+$ goes to 0 as i goes to infinity.

(A4) For every $k = 1, \ldots, K$, there exists a constant $\alpha_k > 0$ such that $J_k/J.$ converges to α_k as $J.$ goes to infinity.

Let $y_{ki}(t) = \mathrm{E}\,Y_{1ki}(t)$ and $y_{\cdot i}(t) = \sum_{k=1}^K y_{ki}(t)$. Let

$$\sigma^2(t_0) = \sum_{i=0}^{\infty} \sum_{k=1}^{K} \int_0^{\infty} \alpha_k \gamma_{ki}^2(s) h_i(s) y_{ki}(s) ds,$$

let $\Sigma(t_0)$ be a matrix with its (l, k)-entry

$$\Sigma_{lk}(t_0) = \sum_{i=0}^{\infty} \int_0^{\infty} \alpha_k g_i^2(s) h_i(s) y_{li}(s) \Big(\delta_{lk} - \frac{y_{ki}(s)}{y_{\cdot i}(s)}\Big) ds$$

and let $\eta(t_0)$ be a vector with l-component

$$\eta_l(t_0) = \sum_{i=0}^{\infty} \sum_{k=1}^{K} \int_0^{\infty} \alpha_k g_i(s) \gamma_{ki}(s) h_i(s) y_{ki}(s) \Big(\delta_{lk} - \frac{y_{li}(s)}{y_{\cdot i}(s)}\Big) ds.$$

Then we have the following weak convergences.

Theorem 3.1. *Assume* (A1)–(A4) *hold. Then,*

(i) *Under* $P_0^{(\mathbf{J})}$, $(\mathcal{L}^{(\mathbf{J})}(t_0), \mathbf{Z}^{(\mathbf{J})}(t_0))'$ *converges weakly, as* $J.$ *goes to infinity, to a normal random vector with mean* $(-\frac{1}{2}\sigma^2(t_0), \mathbf{0})'$ *and covariance matrix* $\begin{pmatrix} \sigma^2(t_0) & \eta(t_0)' \\ \eta(t_0) & \Sigma(t_0) \end{pmatrix}$;

(ii) *Under* $Q^{(\mathbf{J})}$, $(\mathcal{L}^{(\mathbf{J})}(t_0), \mathbf{Z}^{(\mathbf{J})}(t_0))'$ *converges weakly, as* $J.$ *goes to infinity, to a normal random vector with mean* $(\frac{1}{2}\sigma^2(t_0), \eta(t_0))'$ *and the same covariance matrix as in* (i).

Proposition 3.1. *Assume* (A1)–(A4) *hold.* $\hat{\Sigma}^{(\mathbf{J})}(t_0)$ *is an asymptotically consistent estimator for* $\Sigma(t_0)$ *under both* $P_0^{(\mathbf{J})}$ *and* $Q^{(\mathbf{J})}$.

The following theorem establishes the weak convergences of the test statistics $X^{(\mathbf{J})}$. Let $H_1^{(\mathbf{J})}$ denote the hypothesis that (3.1) is valid.

Theorem 3.2. *Assume* (A1)–(A4) *hold. Then,*

(i) *Under the null hypothesis* H_0, $X^{(\mathbf{J})}$ *converges weakly to* $\chi^2(K-1)$, *the central chi-squared distribution with degree of freedom* $K-1$, *as* $J.$ *goes to infinity;*

(ii) *Under the local alternatives* $H_1^{(\mathbf{J})}$, $X^{(\mathbf{J})}$ *converges weakly to* $\chi^2(K-1, \eta(t_0)' \Sigma(t_0)^- \eta(t_0))$, *the noncentral chi-squared distribution with degree of freedom* $K-1$ *and noncentrality* $\eta(t_0)'\Sigma(t_0)^-\eta(t_0)$, *as* $J.$ *goes to infinity.*

3.2. Proofs

We omit the proof for Theorem 3.2, since it follows from Theorem 3.1 by applying the continuous mapping theorem and Proposition 3.1. Proposition 3.1 is easy and we do not prove it either. We now prove Theorem 3.1, which is preceded by several lemmas.

We first introduce a martingale $\bar{\mathbf{Z}}^{(\mathbf{J})}(\cdot)$ to approximate $\mathbf{Z}^{(\mathbf{J})}(\cdot)$.

Let

$$\begin{aligned}\tilde{Z}_l^{(\mathbf{J},i)}(t) &= J_\cdot^{-1/2}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t g_i(s)\Big(\delta_{lk}-\frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)dM_{jki}(s)\\ &= J_\cdot^{-1/2}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t g_i(s)\Big(\delta_{lk}-\frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)1_{(T_{jki},T_{jk(i+1)}]}(T_{jki}+s)\\ &\quad 1_{(0,t_0]}(T_{jki}+s)dM_{jk}(T_{jki}+s)\\ &= J_\cdot^{-1/2}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_{T_{jki}\wedge t_0}^{\left(T_{jki}+t\wedge(T_{jk(i+1)}-T_{jki})\right)\wedge t_0} g_i(s-T_{jki})\\ &\quad \Big(\delta_{lk}-\frac{y_{li}(s-T_{jki})}{y_{\cdot i}(s-T_{jki})}\Big)dM_{jk}(s). \end{aligned} \tag{3.3}$$

Since $(T_{jki}+t\wedge(T_{jk(i+1)}-T_{jki}))\wedge t_0 = T_{jk(i+1)}\wedge t_0$ when $t\ge t_0$, we know, if $t\ge t_0$, then

$$\tilde{Z}_l^{(\mathbf{J},i)}(t)=J_\cdot^{-1/2}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^{t_0} g_i(s-T_{jki})\Big(\delta_{lk}-\frac{y_{li}(s-T_{jki})}{y_{\cdot i}(s-T_{jki})}\Big)1_{(T_{jki},T_{jk(i+1)}]}(s)dM_{jk}(s).$$

Let $\tilde{Z}_l^{(\mathbf{J})}(\cdot)=\sum_{i=0}^{\infty}\tilde{Z}_l^{(\mathbf{J},i)}(\cdot)$ and let

$$\bar{Z}_l^{(\mathbf{J})}(t)=J_\cdot^{-1/2}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t\sum_{i=0}^{\infty} g_i(s-T_{jki})\Big(\delta_{lk}-\frac{y_{li}(s-T_{jki})}{y_{\cdot i}(s-T_{jki})}\Big)1_{(T_{jki},T_{jk(i+1)}]}(s)dM_{jk}(s),$$

which is a martingale relative to the calender time filtration $\mathcal{F}_t^{(\mathbf{J})}$. Let $\tilde{\mathbf{Z}}^{(\mathbf{J})}(\cdot)=(\tilde{Z}_1^{(\mathbf{J})}(\cdot),\ldots,\tilde{Z}_K^{(\mathbf{J})}(\cdot))'$, and $\bar{\mathbf{Z}}^{(\mathbf{J})}(\cdot)=(\bar{Z}_1^{(\mathbf{J})}(\cdot),\ldots,\bar{Z}_K^{(\mathbf{J})}(\cdot))'$. Note that

$$\tilde{Z}_l^{(\mathbf{J})}(t)=\bar{Z}_l^{(\mathbf{J})}(t_0) \tag{3.4}$$

if $t\ge t_0$, which is needed in the proof of Theorem 3.1.

Lemma 3.1. *If for each i, $M_i(t)$ is a mean zero $\mathcal{G}_t^{(\mathbf{J},i)}$-martingale and is $\mathcal{G}_0^{(\mathbf{J},i+1)}$-measurable, then for every given $t\ge 0$, $\sum_{i=0}^{I}M_i(t)$ is a martingale with respect to the filtration $\{\mathcal{G}_0^{(\mathbf{J},I+1)}, I=0,1,2,\ldots\}$.*

Lemma 3.2. *Assume* (A1)–(A3) *hold. Then, under* $P_0^{(\mathbf{J})}$, $\lim_{J.\to\infty} \mathrm{E}\,(Z_l^{(\mathbf{J})}(t_0) - \tilde{Z}_l^{(\mathbf{J})}(t_0))^2 = 0$.

Let $\tilde{Y}_{jki}(t,s) = Y_{jk}(T_{jki}+s)1_{(T_{jki},T_{jk(i+1)}]}(T_{jki}+s)1_{(0,t]}(T_{jki}+s)$, $\tilde{y}_{ki}(t,s) = \mathrm{E}\,\tilde{Y}_{1ki}(t,s)$ and $\tilde{y}_{\cdot i}(t,s) = \sum_{k=1}^{K}\tilde{y}_{ki}(t,s)$. Note that $Y_{jki}(\cdot) = \tilde{Y}_{jki}(t_0,\cdot)$ and $y_{ki}(\cdot) = \tilde{y}_{ki}(t_0,\cdot)$.

Lemma 3.3. *Assume* (A1)–(A4) *hold. Then under* $P_0^{(\mathbf{J})}$ *as J. goes to infinity,* $(\mathcal{L}_1^{(\mathbf{J})}(t), \bar{\mathbf{Z}}^{(\mathbf{J})}(t))'$ *converges weakly to a normal random vector with mean* $\mathbf{0}$ *and covariance matrix* $\begin{pmatrix}\sigma^2(t) & \eta(t)'\\ \eta(t) & \Sigma(t)\end{pmatrix}$ *for* $t \geq 0$, *where*

$$\mathcal{L}_1^{(\mathbf{J})}(t) = J.^{-1/2}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t\sum_{i=0}^{\infty}\gamma_{ki}(s-T_{jki})1_{(T_{jki},T_{jk(i+1)}]}(s)dM_{jk}(s), \quad (3.5)$$

$$\sigma^2(t) = \sum_{i=0}^{\infty}\sum_{k=1}^{K}\int_0^{\infty}\alpha_k\gamma_{ki}^2(s)h_i(s)\tilde{y}_{ki}(t,s)ds,$$

$\Sigma(t)$ *is a matrix whose* (l,k)*-entry is*

$$\Sigma_{lk}(t) = \sum_{i=0}^{\infty}\int_0^{\infty}\alpha_k g_i^2(s)h_i(s)\Big(\delta_{lk}\tilde{y}_{li}(t,s) - \frac{\tilde{y}_{li}(t,s)y_{ki}(s) + y_{li}(s)\ tildey_{ki}(t,s)}{y_{\cdot i}(s)} + \frac{y_{li}(s)y_{ki}(s)\tilde{y}_{\cdot i}(t,s)}{y_{\cdot i}^2(s)}\Big)ds,$$

and $\eta(t)$ *is a vector whose* l*-component is*

$$\eta_l(t) = \sum_{i=0}^{\infty}\sum_{k=1}^{K}\int_0^{\infty}\alpha_k g_i(s)\gamma_{ki}(s)h_i(s)\Big(\delta_{lk} - \frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)\tilde{y}_{ki}(t,s)ds.$$

Lemma 3.4. *Assume* (A2)–(A4) *hold. Then under* $P_0^{(\mathbf{J})}$ *as J. goes to infinity,* $\mathcal{L}_1^{(\mathbf{J})}(t) - \mathcal{L}^{(\mathbf{J})}(t)$ *converges in probability to* $\frac{1}{2}\sigma^2(t)$.

Proof of Theorem 3.1. It follows from Lemma 3.2, Lemma 3.4, and (3.4) that $(\mathcal{L}^{(\mathbf{J})}(t_0) - \frac{1}{2}\sigma^2(t_0), \mathbf{Z}^{(J)}(t_0))'$ and $(\mathcal{L}_1^{(\mathbf{J})}(t_0), \bar{\mathbf{Z}}^{(\mathbf{J})}(t_0))'$ have the same limiting distribution. This together with Lemma 3.3 shows that $(\mathcal{L}^{(\mathbf{J})}(t_0), \mathbf{Z}^{(\mathbf{J})}(t_0))'$ has the desired limiting distribution under $P_0^{(\mathbf{J})}$. Finally, the weak convergence of $(\mathcal{L}^{(\mathbf{J})}(t_0), \mathbf{Z}^{(\mathbf{J})}(t_0))'$ under $Q^{(\mathbf{J})}$ follows from Le Cam's third lemma in a standard way. This completes the proof.

Proof of Lemma 3.1. Since $\mathcal{G}_0^{(\mathbf{J},i)} \subset \mathcal{G}_0^{(\mathbf{J},I+1)}$ for $i \leq I$ and $\sum_{i=0}^{I} M_i(t)$ is

$\mathcal{G}_0^{(\mathbf{J},I+1)}$-measurable, this lemma is proved by observing the following equations.

$$\mathrm{E}\Big(\sum_{i=0}^{I} M_i(t)|\mathcal{G}_0^{(\mathbf{J},I)}\Big) = \sum_{i=0}^{I-1} M_i(t) + \mathrm{E}\,(M_I(t)|\mathcal{G}_0^{(\mathbf{J},I)}) = \sum_{i=0}^{I-1} M_i(t).$$

Proof of Lemma 3.2. It follows from (2.3) and (3.3) that both $Z_l^{(\mathbf{J},i)}(t)$ and $\tilde{Z}_l^{(\mathbf{J},i)}(t)$ satisfy the conditions in Lemma 3.1. With this, we observe that

$$\begin{aligned}
&\mathrm{E}\Big(\sum_{i=0}^{I} Z_l^{(\mathbf{J},i)}(t_0) - \sum_{i=0}^{I} \tilde{Z}_l^{(\mathbf{J},i)}(t_0)\Big)^2 \\
=&\sum_{i=0}^{I} \mathrm{E}\,(Z_l^{(\mathbf{J},i)}(t_0) - \tilde{Z}_l^{(\mathbf{J},i)}(t_0))^2 \qquad (3.6)\\
=&\sum_{i=0}^{I} \mathrm{E}\Big\{J_{\cdot}^{-1/2}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^{t_0}\Big[G_i^{(\mathbf{J})}(s)\Big(\delta_{lk}-\frac{Y_{\cdot li}^{(J_l)}(s)}{Y_{\cdot\cdot i}^{(\mathbf{J})}(s)}\Big)-g_i(s)\Big(\delta_{lk}-\frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)\Big]dM_{jki}(s)\Big\}^2\\
=&\sum_{i=0}^{I} \mathrm{E}\,J_{\cdot}^{-1}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^{t_0}\Big[G_i^{(\mathbf{J})}(s)\Big(\delta_{lk}-\frac{Y_{\cdot li}^{(J_l)}(s)}{Y_{\cdot\cdot i}^{(\mathbf{J})}(s)}\Big)-g_i(s)\Big(\delta_{lk}-\frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)\Big]^2\\
&\quad h_i(s)Y_{jki}(s)ds\\
=&\sum_{i=0}^{I}\sum_{k=1}^{K}\int_0^{t_0}\mathrm{E}\Big[G_i^{(\mathbf{J})}(s)\Big(\delta_{lk}-\frac{Y_{\cdot li}^{(J_k)}(s)}{Y_{\cdot\cdot i}^{(\mathbf{J})}(s)}\Big)-g_i(s)\Big(\delta_{lk}-\frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)\Big]^2\\
&\quad h_i(s)J_{\cdot}^{-1}\sum_{j=1}^{J_k}Y_{jki}(s)ds. \qquad (3.7)
\end{aligned}$$

We note that the first equality uses Lemma 3.1 and the martingale property, and the third equality uses the quadratic variation formula of a stochastic integral.

Let

$$F_{J.ki}(t) = \mathrm{E}\Big\{\Big[G_i^{(\mathbf{J})}(t)\Big(\delta_{lk}-\frac{Y_{\cdot li}^{(J_k)}(t)}{Y_{\cdot\cdot i}^{(\mathbf{J})}(t)}\Big)-g_i(t)\Big(\delta_{lk}-\frac{y_{li}(t)}{y_{\cdot i}(t)}\Big)\Big]^2 h_i(t)J_{\cdot}^{-1}\sum_{j=1}^{J_k}Y_{jki}(t)\Big\}.$$

From (A2) there is a constant $C > 0$ such that

$$|F_{J.ki}(t)| \le C, \qquad (3.8)$$

for every $i = 0, 1, \ldots,\ k = 1, \ldots, K,\ J. = 1, 2, \ldots$, and $t \in [0, t_0]$. Thus, it follows from (A1), (3.8), the Law of Large Numbers, and the Dominated Convergence Theorem, that for every $t \in [0, t_0]$, $F_{J.ki}(t)$ converges to zero as $J.$ goes to infinity. This implies that (3.6) converges to zero as $J.$ goes to infinity.

Calculations similar to (3.7) and (3.8), together with Fatou's Lemma, give the following inequality:

$$\begin{aligned}\mathrm{E}\Big\{\sum_{i=I+1}^{\infty}\Big(Z_l^{(\mathbf{J},i)}(t_0)-\tilde{Z}_l^{(\mathbf{J},i)}(t_0)\Big)\Big\}^2 &\le \frac{C}{J.}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=I+1}^{\infty}\mathrm{E}\,1_{(T_{jki}\wedge t_0,T_{jk(i+1)}\wedge t_0]}(s)\\ &\le C\cdot\sum_{k=1}^{K}\mathrm{E}\,(t_0-T_{1kI})^+, \qquad (3.9)\end{aligned}$$

for a suitable constant C. Because of the non-explosiveness of $N_{jk}(\cdot)$, we know (3.9) is small if I is large.

We can now choose a large I to make (3.9) small and then a large $J.$ to make (3.6) small. This shows that $\mathrm{E}\,(Z_l^{(J)}(t_0)-\tilde{Z}_l^{(J)}(t_0))^2$ is small if $J.$ is large and completes the proof.

Proof of Lemma 3.3. Since both $\mathcal{L}_1^{(\mathbf{J})}(\cdot)$ and $\bar{\mathbf{Z}}^{(\mathbf{J})}(\cdot)$ are martingales, the proof is a standard application of the Martingale Central Limit Theorem (cf. ABGK (1982), (1993), or Fleming and Harrington (1991)). We omit the details of the proof.

Proof of Lemma 3.4. Let

$$\mathcal{L}_2^{(\mathbf{J})}(t)=-\frac{1}{2J.}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t\sum_{i=0}^{\infty}\gamma_{ki}^2(s-T_{jki})1_{(T_{jki},T_{jk(i+1)}]}(s)dN_{jk}(s).$$

Define $\mathcal{L}_3^{(\mathbf{J})}(t)$ by requiring $\mathcal{L}^{(\mathbf{J})}(t)=\mathcal{L}_1^{(\mathbf{J})}(t)+\mathcal{L}_2^{(\mathbf{J})}(t)+\mathcal{L}_3^{(\mathbf{J})}(t)$. It is clear that $\mathcal{L}_2^{(\mathbf{J})}(t)$ converges in probability to

$$\begin{aligned}&-\frac{1}{2}\sum_{k=1}^{K}\alpha_k\mathrm{E}\int_0^t\sum_{i=0}^{\infty}\gamma_{ki}^2(s-T_{1ki})1_{(T_{1ki},T_{1k(i+1)}]}(s)h_i(s-T_{1ki})Y_{1k}(s)ds\\ &=-\frac{1}{2}\sum_{i=0}^{\infty}\sum_{k=1}^{K}\int_0^{\infty}\alpha_k\gamma_{ki}^2(s)h_i(s)\mathrm{E}\,\tilde{Y}_{1ki}(t,s)ds\\ &=-\frac{1}{2}\sum_{i=0}^{\infty}\sum_{k=1}^{K}\int_0^{\infty}\alpha_k\gamma_{ki}^2(s)h_i(s)\tilde{y}_{ki}(t,s)ds,\end{aligned}$$

as $J.$ goes infinity. Since $|\log(1+x)-x+\frac{1}{2}x^2|\le cx^3$ for some constant C on a neighborhood of 0, we have

$$\begin{aligned}|\mathcal{L}_3^{(\mathbf{J})}(t)|&\le\frac{C}{J.^{\frac{3}{2}}}\sum_{k=1}^{K}\sum_{j=1}^{J_k}\int_0^t\sum_{i=0}^{\infty}\gamma_{ki}^3(s-T_{jki})1_{(T_{jki},T_{jk(i+1)}]}(s)dN_{jk}(s)\\ &\le\frac{CM^3}{J.^{\frac{3}{2}}}\sum_{k=1}^{K}\sum_{j=1}^{J_k}(N_{jk}(t)+1),\end{aligned}$$

which converges to 0 as $J_\cdot$ goes to infinity. This completes the proof.

4. Asymptotic Efficiency of the Nonparametric K-sample Tests

In this section we introduce a parametric submodel and study the hypothesis testing problem of $H_0^{(\mathbf{J})}$ versus $H_1^{(\mathbf{J})}$ in the submodel. In particular, by considering the likelihood process in this submodel, we are able to present its optimal test statistic in Theorem 4.1. Subsection 4.2 shows that the nonparametric test statistic of Section 3 is asymptotically equivalent to the statistic introduced in Subsection 4.1 for the parametric submodel, which establishes the asymptotic efficiency of the test statistic.

4.1. The parametric submodel

Assume that the function $\gamma_{ki}(\cdot)$ in (3.1) is of the form

$$\gamma_{ki}(t) = \gamma_i(t)\Big(\phi_k + \frac{\sum_{l=1}^K \psi_l y_{li}(t)}{y_{\cdot i}(t)}\Big), \tag{4.1}$$

where $\phi_1, \ldots, \phi_K$, $\psi_1, \ldots, \psi_K$ are unknown parameters and $\gamma_i(\cdot)$ is a known bounded deterministic function. With this assumption, we have a parametric submodel and we will treat $\psi_1, \ldots, \psi_K$ as nuisance parameters. The null hypothesis now becomes $H_0^{(\mathbf{J})} : \phi_1 = \cdots = \phi_K$ and the alternative becomes $H_1^{(\mathbf{J})} : \phi_l \neq \phi_k$ for some $l \neq k$.

We set the following notation to ease the discussion. Let

$$\begin{aligned}
\phi &= (\phi_1, \ldots, \phi_K)', \quad \psi = (\psi_1, \ldots, \psi_K)', \\
U_l^{(\mathbf{J})} &= J_\cdot^{-1/2} \sum_{i=0}^{\infty} \sum_{j=1}^{J_l} \int_0^{t_0} \gamma_i(s) dM_{jli}(s), \quad \text{for} \quad l = 1, \ldots, K, \\
V_l^{(\mathbf{J})} &= J_\cdot^{-1/2} \sum_{i=0}^{\infty} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \int_0^{t_0} \gamma_i(s) \frac{y_{li}(s)}{y_{\cdot i}(s)} dM_{jki}(s), \quad \text{for} \quad l = 1, \ldots, K, \\
\mathbf{U}^{(\mathbf{J})} &= (U_1^{(\mathbf{J})}, \ldots, U_K^{(\mathbf{J})})', \quad \mathbf{V}^{(\mathbf{J})} = (V_1^{(\mathbf{J})}, \ldots, V_K^{(\mathbf{J})})', \\
\sigma^2(\phi, \psi, t_0) &= \sum_{i=0}^{\infty} \sum_{k=1}^{K} \int_0^{\infty} \alpha_k \gamma_i^2(s) \Big(\phi_k + \frac{\sum_{l=1}^K \psi_l y_{li}(s)}{y_{\cdot i}(s)}\Big)^2 h_i(s) y_{ki}(s) ds, \\
\mathcal{I}_{lk}^{11} &= \delta_{lk} \sum_{i=0}^{\infty} \int_0^{\infty} \alpha_k \gamma_i^2(s) y_{li}(s) h_i(s) ds, \text{ for } l = 1, \ldots, K,\ k = 1, \ldots, K, \\
\mathcal{I}_{lk}^{22} &= \sum_{i=0}^{\infty} \int_0^{\infty} \gamma_i^2(s) \frac{y_{li}(s) y_{ki}(s)}{y_{\cdot i}(s)} \sum_{m=1}^{K} \frac{\alpha_m y_{mi}(s)}{y_{\cdot i}(s)} h_i(s) ds, \text{ for } l = 1, \ldots, K, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad k = 1, \ldots, K, \\
\mathcal{I}_{lk}^{12} &= \mathcal{I}_{lk}^{21} = \mathcal{I}_{lk}^{22}, \quad \mathcal{I}^{11} = \Big(\mathcal{I}_{lk}^{11}\Big)_{K \times K}, \quad \mathcal{I}^{12} = \mathcal{I}^{21} = \mathcal{I}^{22} = (\mathcal{I}_{lk}^{22})_{K \times K},
\end{aligned}$$

and

$$\mathcal{I} = \begin{pmatrix} \mathcal{I}^{11} \, \mathcal{I}^{12} \\ \mathcal{I}^{21} \, \mathcal{I}^{22} \end{pmatrix}_{2K \times 2K}.$$

Let $Q^{(\mathbf{J},\phi,\psi)}$ denote the probability measure specified by (4.1), and let $\mathcal{L}^{(\mathbf{J})}(\phi,\psi,t_0) = \log \frac{dQ^{(\mathbf{J},\phi,\psi)}}{dP_0^{(\mathbf{J})}}$. From Theorem 3.1 we have the following lemma.

Lemma 4.1. (i) *Under $P_0^{(\mathbf{J})}$ as $J_\cdot$ goes to infinity, $\mathcal{L}^{(\mathbf{J})}(\phi,\psi,t_0)$ converges weakly to a normal random variable with mean $-\frac{1}{2}\sigma^2(\phi,\psi,t_0)$ and variance $\sigma^2(\phi,\psi,t_0)$.*

(ii) *Under $Q^{(\mathbf{J},\phi,\psi)}$ as $J_\cdot$ goes to infinity, $\mathcal{L}^{(J)}(\phi,\psi,t_0)$ converges weakly to a normal random variable with mean $\frac{1}{2}\sigma^2(\phi,\psi,t_0)$ and variance $\sigma^2(\phi,\psi,t_0)$.*

Lemma 4.2. $\binom{\phi}{\psi}' \mathcal{I} \binom{\phi}{\psi} = \sigma^2(\phi,\psi,t_0)$.

Proof of Lemma 4.2. Let $e_l = (\overbrace{0,\ldots,0,1}^{l},0,\ldots,0)' \in R^{2K}$ and

$$e_{lk} = (\overbrace{0,\ldots,0,1}^{l},0,\ldots,0,1,\overbrace{0,\cdots,0}^{K-k})' \in R^{2K}.$$

The lemma follows from the following straightforward calculations: $\sigma^2(e_l,\mathbf{0},t_0) = \mathcal{I}_{ll}^{11}$, $\sigma^2(e_{lk},\mathbf{0},t_0) = \mathcal{I}_{ll}^{11} + \mathcal{I}_{kk}^{11}$, $\sigma^2(\mathbf{0},e_l,t_0) = \mathcal{I}_{ll}^{22}$, $\sigma^2(\mathbf{0},e_{lk},t_0) = \mathcal{I}_{ll}^{22} + 2\mathcal{I}_{lk}^{22} + \mathcal{I}_{kk}^{22}$, and $\sigma^2(e_l,e_k,t_0) = \mathcal{I}_{ll}^{11} + 2\mathcal{I}_{lk}^{12} + \mathcal{I}_{kk}^{22}$.

Lemma 4.3. *Under both $P_0^{(\mathbf{J})}$ and $Q^{(\mathbf{J},\phi,\psi)}$,*

$$\mathcal{L}^{(\mathbf{J})}(\phi,\psi,t_0) = \begin{pmatrix} \phi \\ \psi \end{pmatrix}' \begin{pmatrix} \mathbf{U}^{(\mathbf{J})} \\ \mathbf{V}^{(\mathbf{J})} \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \phi \\ \psi \end{pmatrix}' \mathcal{I} \begin{pmatrix} \phi \\ \psi \end{pmatrix} + o_p(1). \tag{4.2}$$

Proof of Lemma 4.3. It follows from (3.2), (3.5), Lemma 3.4 and Lemma 4.2 that, as $J_\cdot$ goes to infinity,

$$\begin{aligned}
&\mathcal{L}^{(\mathbf{J})}(\phi,\psi,t_0) \\
&= J_\cdot^{-1/2} \sum_{i=0}^{\infty} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \int_0^{t_0} \gamma_i(s) \Big(\phi_k + \frac{\sum_{l=1}^{K} \psi_l y_{li}(s)}{y_{\cdot i}(s)}\Big) dM_{jki}(s) - \frac{1}{2}\sigma^2(\phi,\psi,t_0) + o_p(1) \\
&= \sum_{k=1}^{K} \frac{\phi_k}{J_\cdot^{1/2}} \sum_{i=0}^{\infty} \sum_{j=1}^{J_k} \int_0^{t_0} \gamma_i(s) dM_{jki}(s) + \sum_{l=1}^{K} \frac{\psi_l}{J_\cdot^{1/2}} \sum_{i=0}^{\infty} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \int_0^{t_0} \gamma_i(s) \frac{y_{li}(s)}{y_{\cdot i}(s)} dM_{jki}(s) \\
&\quad - \frac{1}{2}\sigma^2(\phi,\psi,t_0) + o_p(1) \\
&= \begin{pmatrix} \phi \\ \psi \end{pmatrix}' \begin{pmatrix} \mathbf{U}^{(\mathbf{J})} \\ \mathbf{V}^{(\mathbf{J})} \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \phi \\ \psi \end{pmatrix}' \mathcal{I} \begin{pmatrix} \phi \\ \psi \end{pmatrix} + o_p(1).
\end{aligned}$$

This completes the proof.

With Lemma 4.1, Lemma 4.3 and the Cramér-Wold device, we can obtain the following weak convergences.

Lemma 4.4. (i) *Under $P_0^{(\mathbf{J})}$ as J. goes to infinity, $(\mathbf{U}^{(\mathbf{J})}, \mathbf{V}^{(\mathbf{J})})'$ converges weakly to a normal random vector with mean $\mathbf{0}$ and covariance matrix $\mathcal{I}$.*

(ii) *Under $Q^{(\mathbf{J},\phi,\psi)}$ as J. goes to infinity, $(\mathbf{U}^{(\mathbf{J})}, \mathbf{V}^{(\mathbf{J})})'$ converges weakly to a normal random vector with mean $\mathcal{I} \cdot (\phi, \psi)'$ and covariance matrix $\mathcal{I}$.*

It follows from Lemma 4.4 that the optimal test for the parametric submodel testing problem is based on $\tilde{\mathbf{U}}^{(\mathbf{J})}$, defined by

$$\tilde{\mathbf{U}}^{(\mathbf{J})} \equiv \mathbf{U}^{(\mathbf{J})} - \mathcal{I}^{12}(\mathcal{I}^{22})^{-1}\mathbf{V}^{(\mathbf{J})}, \tag{4.3}$$

whose asymptotic distributions are given in the following lemma.

Lemma 4.5. (i) *Under $H_0^{(\mathbf{J})}$ as J. goes to infinity, $\tilde{\mathbf{U}}^{(\mathbf{J})}$ converges weakly to a normal random vector with mean $\mathbf{0}$ and covariance matrix $\mathcal{I}^{11|22}$, where $\mathcal{I}^{11|22} \equiv \mathcal{I}^{11} - \mathcal{I}^{22}$.*

(ii) *Under $H_1^{(\mathbf{J})}$ as J. goes to infinity, $\tilde{\mathbf{U}}^{(\mathbf{J})}$ converges weakly to a normal random vector with mean $\mathcal{I}^{11|22} \cdot \phi$ and covariance matrix $\mathcal{I}^{11|22}$.*

Proof of Lemma 4.5. Let $(\tilde{\mathcal{I}}_1, \ldots, \tilde{\mathcal{I}}_{2K})' = \mathcal{I} \cdot (\phi, \psi)'$. It follows from Lemma 4.4, the Continuous Mapping Theorem and multivariate normal theory that, under $Q^{(\mathbf{J},\phi,\psi)}$ and as J. goes to infinity, $\tilde{\mathbf{U}}^{(\mathbf{J})}$ converges weakly to a normal distribution with mean $(\tilde{\mathcal{I}}_1, \ldots, \tilde{\mathcal{I}}_K)' - \mathcal{I}^{12}(\mathcal{I}^{22})^{-1}(\tilde{\mathcal{I}}_{K+1}, \ldots, \tilde{\mathcal{I}}_{2K})'$ and variance $\mathcal{I}^{11} - \mathcal{I}^{12}(\mathcal{I}^{22})^{-1}\mathcal{I}^{21}$. Then the second weak convergence of the lemma follows by observing that $(\tilde{\mathcal{I}}_1, \ldots, \tilde{\mathcal{I}}_K)' - \mathcal{I}^{12}(\mathcal{I}^{22})^{-1}(\tilde{\mathcal{I}}_{K+1}, \ldots, \tilde{\mathcal{I}}_{2K})' = (\mathcal{I}^{11} - \mathcal{I}^{12}) \cdot \phi = \mathcal{I}^{11|12} \cdot \phi$ and that $\mathcal{I}^{11} - \mathcal{I}^{12}(\mathcal{I}^{22})^{-1}\mathcal{I}^{21} = \mathcal{I}^{11} - \mathcal{I}^{22} = \mathcal{I}^{11|12}$. The fact that $(\mathcal{I}^{11} - \mathcal{I}^{12}) \cdot \phi = 0$ if and only if $\phi_1 = \cdots = \phi_K$ establishes the first weak convergence and completes the proof.

From Lemma 4.5, the test statistic $W^{(\mathbf{J})} \equiv (\tilde{\mathbf{U}}^{(\mathbf{J})})'(\mathcal{I}^{11|12})^{-}\tilde{\mathbf{U}}^{(\mathbf{J})}$ is an asymptotically optimal test for the parametric submodel testing problem in the sense that it is asymptotically most powerful in an invariant class of tests, with power invariant under transformations of the parameters of interest $(\phi_2 - \phi_1, \ldots, \phi_K - \phi_{K-1})$ which preserve the effective information for these parameters, i.e., tests whose power is constant on the ellipsoids $\phi'\mathcal{I}^{11|12}\phi = \text{constant}$. See Choi (1989) for a related concept of optimality.

The asymptotic distribution of $W^{(\mathbf{J})}$ is now derived as a consequence of Lemma 4.5.

Theorem 4.1. *Under $H_0^{(\mathbf{J})}$, $W^{(\mathbf{J})}$ converges weakly to $\chi^2(K-1)$; under $H_1^{(\mathbf{J})}$, $W^{(J)}$ converges weakly to $\chi^2(K-1, \phi'\mathcal{I}^{11|12}\phi)$, where $K-1 = $ rank of $\mathcal{I}^{11|12}$.*

4.2. Asymptotic efficiency of the nonparametric tests

From Theorem 3.2 and Theorem 4.1, the nonparametric test statistic $X^{(\mathbf{J})}$ and the asymptotically optimal test statistic $W^{(\mathbf{J})}$ for the parametric submodel (4.1) have the same asymptotic distribution under the null hypothesis. By considering a specific parametric submodel of the form (4.1), we can show further that $X^{(\mathbf{J})}$ and $W^{(\mathbf{J})}$ are asymptotically equivalent.

Theorem 4.2. *If $\gamma_i = g_i$ for $i = 0, 1, 2, \ldots$ in* (4.1)*, then the asymptotic distribution of $W^{(\mathbf{J})}$ is the same as that of $X^{(\mathbf{J})}$ under the local alternative $H_1^{(\mathbf{J})}$ for* (4.1)*.*

Proof of Theorem 4.2. From Theorem 3.2 and Theorem 4.1, we need only show that

$$\eta(t_0)'\Sigma(t_0)^-\eta(t_0) = \phi'\mathcal{I}^{11|12}\phi. \tag{4.4}$$

Observe that

$$\begin{aligned}
\eta_l(t_0) &= \sum_{i=0}^{\infty}\sum_{k=1}^{K}\int_0^\infty \alpha_k g_i(s)\gamma_{ki}(s)h_i(s)y_{ki}(s)\Big(\delta_{lk} - \frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)ds \\
&= \sum_{i=0}^{\infty}\sum_{k=1}^{K}\int_0^\infty \alpha_k g_i^2(s)\Big(\phi_k + \frac{\sum_{m=1}^K \psi_m y_{mi}(s)}{y_{\cdot i}(s)}\Big)h_i(s)y_{ki}(s)\Big(\delta_{lk} - \frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)ds \\
&= \sum_{i=0}^{\infty}\sum_{k=1}^{K}\int_0^\infty \alpha_k g_i^2(s)h_i(s)y_{ki}(s)\Big(\phi_k\delta_{lk} + \frac{\sum_{m=1}^K \psi_m y_{mi}(s)}{y_{\cdot i}(s)}\delta_{lk} \\
&\qquad -\phi_k\frac{y_{li}(s)}{y_{\cdot i}(s)} - \frac{\sum_{m=1}^K \psi_m y_{mi}(s)}{y_{\cdot i}(s)}\frac{y_{li}(s)}{y_{\cdot i}(s)}\Big)ds \\
&= \sum_{i=0}^{\infty}\int_0^\infty g_i^2(s)h_i(s)\Big(\alpha_l y_{li}(s)\phi_l + \sum_{m=1}^{K}\frac{\psi_m y_{mi}(s)}{y_{\cdot i}(s)}\alpha_l y_{li}(s) \\
&\qquad -\sum_{k=1}^{K}\alpha_k y_{ki}(s)\phi_k\frac{y_{li}(s)}{y_{\cdot i}(s)} - \sum_{m=1}^{K}\frac{\psi_m y_{mi}(s)}{y_{\cdot i}(s)}y_{li}(s)\sum_{k=1}^{K}\frac{\alpha_k y_{ki}(s)}{y_{\cdot i}(s)}\Big)ds \\
&= \sum_{i=0}^{\infty}\int_0^\infty g_i^2(s)h_i(s)y_{li}(s)\Big(\alpha_l\phi_l - \sum_{k=1}^{K}\frac{\alpha_k\phi_k y_{ki}(s)}{y_{\cdot i}(s)} \\
&\qquad\qquad + \sum_{m=1}^{K}\frac{\psi_m y_{mi}(s)}{y_{\cdot i}(s)}\Big(\alpha_l - \sum_{k=1}^{K}\frac{\alpha_k y_{ki}(s)}{y_{\cdot i}}\Big)\Big)ds \\
&= ((\mathcal{I}^{11} - \mathcal{I}^{22})\phi)_l = (\mathcal{I}^{11|12}\phi)_l,
\end{aligned} \tag{4.5}$$

and

$$\Sigma_{lk}(t_0) = \sum_{i=0}^{\infty}\int_0^\infty \alpha_k g_i^2(s)h_i(s)y_{li}(s)\Big(\delta_{lk} - \frac{y_{ki}(s)}{y_{\cdot i}(s)}\Big)ds$$

$$= ((\mathcal{I}^{11} - \mathcal{I}^{22}))_{lk} = (\mathcal{I}^{11|12})_{lk}. \tag{4.6}$$

With (4.5) and (4.6), we get (4.4) immediately. This completes the proof.

5. Simulation and Example

The first part of this section describes a computer simulation of a two-sample semi-Markov model under the null hypothesis. We generate counting processes $N_{jk}(t), 0 \le t \le 1$, satisfying the following properties. $\{N_{jk}(\cdot)\}$ are independent for $k = 1, 2$ and $j = 1, 2, \ldots, J$. The intensity of $N_{jk}(\cdot)$ is of the form $\lambda_{jk}(\cdot)$ given in (1.1), with $h_{1i}(t) = (9 - i)^+ \cdot 1.2$, for $i = 0, 1, 2, \ldots$, and $Y_{jk}(\cdot) = 1$, and we set $h_{2i}(\cdot) = h_{1i}(\cdot)$, for $i = 0, 1, 2, \ldots$

Let $X^{(\mathbf{J})}$ be the statistic in (2.4) with weight process $G_i^{(\mathbf{J})}(\cdot) \equiv 1$, for $i = 0, 1, 2, \ldots$, i.e.,

$$X^{(\mathbf{J})} = \Big(\sum_{i=0}^{\infty} \sum_{j=1}^{J} \int_0^1 \frac{Y_{\cdot 2i}(s)}{Y_{\cdot\cdot i}(s)} dN_{j1i}(s) - \frac{Y_{\cdot 1i}(s)}{Y_{\cdot\cdot i}(s)} dN_{j2i}(s) \Big)^2 \Big(\sum_{i=0}^{\infty} \sum_{j=1}^{J} \int_0^1 \frac{Y_{\cdot 2i}(s)}{Y_{\cdot\cdot i}(s)} dN_{j1i}(s) \Big)^{-1},$$

where $Y_{jki}(s) \equiv 1_{(T_{jki}, T_{jk(i+1)}]}(T_{jki} + s) 1_{(0,1]}(T_{jki} + s)$.

A total of $n = 5000$ data sets are generated. From these data sets the empirical distribution $\hat{F}_n^{(\mathbf{J})}$ of $X^{(\mathbf{J})}$ is computed at quantile points $(\chi^2(1))^{-1}(t)$ for $t = 0.05, 0.1, \ldots, 0.95$, with $J = 50$, where $\chi^2(1)$ is the central chi-squared distribution with one degree of freedom. These numbers are presented in Table 5.1. It indicates clearly that $\hat{F}_n^{(\mathbf{J})}((\chi^2(1))^{-1}(t))$ is quite close to t.

Table 5.1. The empirical distribution of $X^{(\mathbf{J})}$ at quantile points $(\chi^2(1))^{-1}(t)$ for different values of t, under the null hypothesis $h_{2i}(\cdot) = h_{1i}(\cdot)$, for $i = 0, 1, 2, \ldots$

t	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60
$(\chi^2(1))^{-1}(t)$	3.841	2.705	2.072	1.642	1.323	1.074	0.873	0.708
$\hat{F}_n^{(\mathbf{J})}((\chi^2(1))^{-1}(t))$	0.949	0.898	0.846	0.794	0.745	0.696	0.645	0.596

0.55	0.50	0.45	0.40	0.35	0.30	0.25	0.20	0.15	0.10	0.05
0.571	0.455	0.357	0.275	0.206	0.148	0.102	0.064	0.036	0.016	0.004
0.552	0.504	0.447	0.398	0.349	0.299	0.251	0.197	0.147	0.098	0.051

We now illustrate the theory of this paper by analyzing the multivariate failure time data of Thompson et al. (1978) in an experimental animal carcinogenesis study. The data summarized here are taken from Gail et al. (1980).

Seventy-six animals were injected with a carcinogen for mammary cancer at day zero, then all animals were given retinyl acetate to prevent cancer for sixty days. After 60 days, the 48 animals which remained tumor-free were randomly assigned to continued retinoid prophylaxis (Treatment Group 1) or control (Treatment Group 2). Rats were palpated for tumors twice weekly, and observation ended 182 days after the initial carcinogen injection. The times to development of mammary cancer for 23 rats in treatment group 1 and 25 rats in the control group are given.

Because rats were palpated only twice a week, two or more tumors may be found at the same time on a given rat. In order to analyze these data with our theory, in which it is required that no two events occur at the same time, we transform the data by the following random mapping. If x is one of the times to tumor of a certain rat reported in Gail et al. (1980), we replace it by $m(x)$, where $m(x)$ is a number chosen randomly from $(\frac{x-60-3.5}{182-60}, \frac{x-60}{182-60})$ according to the uniform distribution.

Accepting the above randomization device, the result is as follows. The test statistic $X^{(\mathbf{J})}$ with $G_i^{(\mathbf{J})} = 1$ in (2.4) has value 12.9 for the transformed data, a p-value of 3.3×10^{-4}. This suggests that H_0 is to be rejected, very nuch in line with the results reported in Gail et al. (1980). We will investigate the appropriateness of the randomization device for semi-Morkov counting process in a future study.

Acknowledgements

This work was supported by grants from National Science Council of ROC. Part of the work was included in the 1995 doctoral dissertation of Yuan-Chuan Chuang, Department of Mathematics, National Central University . The authors would like to thank Dr. David Harrington for some helpful comments and two anonymous referees, whose suggestions led to a significant improvement of the paper.

References

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1982). Linear nonparametric tests for comparison of counting processes, with application to censored survival data. *Internat. Statist. Rev.* **50**, 219-258.

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

Brémaud, P. (1981). *Point Processes and Queues, Martingale Dynamics.* Springer-Verlag, New York.

Chang, I.-S. and Hsiung, C. A. (1994). Information and asymptotic efficiency in some generalized proportional hazards models for counting processes. *Ann. Statist.* **22**, 1275-1298.

Chang, I.-S., Hsiung, C. A. and Chuang, Y.-C. (1997). Goodness-of-fit tests for survival models with intermediate states. Submitted.

Chang, S.-H. (1995). Regression analysis for recurrent event data. Ph.D. Thesis, John Hopkins University.

Choi, S. (1989). On asymptotically optimal tests. Ph.D. Thesis, University of Rochester.

Commenges, D. (1986). Semi-Markov and non-homogeneous Markov models in medicine. In *Semi-Markov models* (Edited by J. Janssen), 423-436. Plenum Press, New York.

Cox, D. R. (1986). Some remarks on semi-Markov processes in medical studies. In *Semi-Markov models* (Edited by J. Janssen), 411-422. Plenum Press, New York.

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

Gail, M. H., Santer, T. J. and Brown, C. C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics* **36**, 255-266.

Gill, R. D. (1980). *Censoring and Stochastic Integrals.* Mathematical Centre Tracts 124, Mathematisch Centrum, Amsterdam.

Hjort, N. L. (1985). Contribution to the discussion of Andersen and Borgan's "Counting process models for life history data: A review". *Scand. J. Statist.* **12**, 141-150. *Biometrika* **79**, 495-512.

Prentice, R. L., Williams, B. J. and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68**, 373-379.

Thompson, H. F., Grubbs, C. J., Moon, R. C. and Sporn, M. B. (1978). Continual requirement of retinoid for maintenance of mammary cancer inhibition. *Proceedings of the Annual Meeting of the American Association for Cancer Research* **19**, 74.

Voelkel, J. G. and Crowley, J. (1984). Nonparametric inference for a class of semi–Markov processes with censored observations. *Ann. Statist.* **12**, 142-160.

Department of Mathematics, National Central University, Chung-Li, Taiwan.

E-mail: chang@math.ncu.edu.tw

Department of Statistics, Ming-Chuan University, Taipei, Taiwan.

E-mail: ycchuang@mcu.edu.tw

Division of Biostatistics, National Health Research Institute and Institute of Statistical Science, Academia Sinica, Taipei, Taiwan.

E-mail: hsiung@stat.sinica.edu.tw

(Received October 1996; accepted February 1998)