Statistica Sinica **10**(2000), 1257-1266

ESTIMATION IN A PROPORTIONAL HAZARD MODEL FOR SEMI-MARKOV COUNTING PROCESS

I-Shou Chang, Chao A. Hsiung and Su-Mei Wu

National Central University, National Health Research Institutes and National Central University

Abstract: Estimation is studied in a regression model for counting processes whose baseline intensity processes are of semi-Markov form. Asymptotic normality is established for a Breslow-type estimator of the cumulative baseline hazard for each gap time of the counting process.

Key words and phrases: Breslow-type estimator, counting process, proportional hazard model, semi-Markov model.

1. Introduction

1.1. Let $N_1(t), N_2(t), \ldots$ be a sequence of independent and identically distributed counting processes. Let $T_{ji} = inf\{t > 0 \mid N_j(t) = i\}$. Assume that relative to a filtration $\mathcal{F}_{j,t}$, $N_j(t)$ has an intensity process of the form

$$\lambda_j(t) = \lambda_{j0}(t)Y_j(t)\exp(\theta_0 Z_j(t)), \tag{1.1}$$

where $Y_j(\cdot) \geq 0$ and $Z_j(\cdot)$ are bounded $\mathcal{F}_{j,t}$-predictable processes, $\theta_0 \in \Theta$ (a bounded subset of $\Re$) is the relative risk coefficient, and

$$\lambda_{j0}(t) = \sum_{i=1}^{I} h_i^{(0)}(t - T_{ji-1})\mathbf{1}_{(T_{ji-1},T_{ji}]}(t) \tag{1.2}$$

for some functions $h_i^{(0)} \geq 0$. Here $\mathbf{1}_A$ is the indicator function of a set A.

Let $\Lambda_i(t) = \int_0^t h_i^{(0)}(u)\,du$. The statistical problem is to estimate $\Lambda_1(t), \Lambda_2(t), \ldots, \Lambda_I(t)$ and θ_0 based on the data

$$\{N_j(t), Y_j(t), Z_j(t) | 0 \leq t \leq t_0, j = 1, \ldots, J\}, \tag{1.3}$$

assuming that $\mathcal{F}_{1,\infty}, \ldots, \mathcal{F}_{J,\infty}$ are independent. Here $0 < t_0 \leq \infty$ is a termination time.

We note that (1.1) provides a model for certain recurrent event data. In a medical context, $N_j(t)$ is the number of events experienced by the jth individual at time t, while $Y_j(t)$ and $Z_j(t)$ indicate, respectively, the censoring status and the covariate of this individual at time t.

When $I = 1$, (1.1) is the classical Cox model for survival time studied by Andersen and Gill (1982), among others. When $h_1 = h_2 = \cdots = h_I$, (1.1) includes as a special case the modulated renewal processes studied by Oakes and Cui (1994) (cf. Cox (1972) and Cox (1997)). The model (1.1) in this generality was studied by Prentice, Williams and Petersen (1981), Keiding (1986), Cox (1986), Commenges (1986) and Chang and Hsiung (1994), among others.

Except for the case $I = 1$ and the modulated renewal processes, estimation for model (1.1) was studied mainly for θ_0. In this paper we estimate $\Lambda_1, \ldots, \Lambda_I$ and θ_0 simultaneously, which is useful even in the estimation of Λ_1 for a modulated renewal process.

1.2. The estimators to be studied in this paper are described as follows. Let

$$M_j(t) = N_j(t) - \int_0^t \lambda_j(u)du,$$
$$M_{ji}(t) = M_j((T_{ji-1} + t \wedge (T_{ji} - T_{ji-1})) \wedge t_0) - M_j(T_{ji-1} \wedge t_0).$$

Since $M_j(t)$ is an $\mathcal{F}_{j,t}$-martingale and T_{ji-1} is an $\mathcal{F}_{j,t}$-stopping time, we know $M_j(T_{ji-1}+t)$ is an $\mathcal{F}_{j,T_{ji-1}+t}$-martingale. This, together with the fact that $T_{ji} - T_{ji-1}$ is an $\mathcal{F}_{j,T_{ji-1}+t}$-stopping time, shows that $M_{ji}(t)$ is an $\mathcal{F}_{j,T_{ji-1}+t}$-martingale. (cf. Chang and Hsiung (1994)).

Let

$$Y_{ji}(t) = \mathbf{1}_{(0,T_{ji}-T_{ji-1}]}(t)\mathbf{1}_{(0,t_0]}(T_{ji-1} + t)Y_j(T_{ji-1} + t),$$
$$Z_{ji}(t) = Z_j(T_{ji-1} + t),$$
$$N_{ji}(t) = N_j((T_{ji-1} + t \wedge (T_{ji} - T_{ji-1})) \wedge t_0) - N_j(T_{ji-1} \wedge t_0).$$

Then

$$M_{ji}(t) = N_{ji}(t) - \int_0^t h_i^{(0)}(u)Y_{ji}(u)e^{\theta_0 Z_{ji}(u)}\,du, \tag{1.4}$$

which says that, relative to $\mathcal{F}_{j,T_{ji-1}+t}$, the counting process $N_{ji}(t)$ is a Cox regression model with intensity $h_i^{(0)}(t)Y_{ji}(t)e^{\theta_0 Z_{ji}(t)}$, for each fixed i.

For $k = 0, 1, 2$, we set $S_{J,i}^{(k)}(\theta, t) = \frac{1}{J}\sum_{j=1}^J Y_{ji}(t)Z_{ji}^k(t)e^{\theta Z_{ji}(t)}$. Let

$$G_J^{(i)}(\theta, t) = \sum_{j=1}^J \int_0^t \Big(Z_{ji}(u) - \frac{S_{J,i}^{(1)}(\theta, u)}{S_{J,i}^{(0)}(\theta, u)}\Big)\,dN_{ji}(u),$$
$$G_J(\theta) = \sum_{i=1}^I G_J^{(i)}(\theta, \alpha_i) \tag{1.5}$$

for some $\alpha_1, \ldots, \alpha_I$ to be defined in Section 2.

The estimator $\widehat{\theta}_J$ proposed by Prentice et al. (1981) is a root of $G_J(\cdot) = 0$. For the estimation of $\Lambda_i(t)$, we propose the Breslow-type estimator

$$\widehat{\Lambda}_J^{(i)}(t) = \sum_{j=1}^{J} \int_0^t \frac{1}{\sum_{k=1}^{J} Y_{ki}(u) e^{\widehat{\theta}_J Z_{ki}(u)}} \, dN_{ji}(u). \tag{1.6}$$

We note that, when $\theta_0 = 0$, $\widehat{\Lambda}_J^{(i)}(t)$ reduces to the Nelson-Aalen estimator and $\exp(-\widehat{\Lambda}_J^{(i)}(t))$ approximates the Kaplan-Meier estimator of the gap distribution $T_{1i} - T_{1i-1}$. Readers are referred to Fleming and Harrington (1991) for the definitions of these terms and the related properties.

The asymptotic properties of $\widehat{\theta}_J$ are studied by Chang and Hsiung (1994). In this paper we establish the asymptotic distribution of $J^{1/2}(\widehat{\Lambda}_J^{(1)}(\cdot) - \Lambda_1(\cdot), \ldots, \widehat{\Lambda}_J^{(I)}(\cdot) - \Lambda_I(\cdot))$. This suggests that we can estimate $\Lambda_1(\cdot)$ by a linear combination of $\widehat{\Lambda}_J^{(i)}(\cdot)'$s when $\Lambda_1 = \cdots = \Lambda_I$. In fact, care is taken so that the data is fully utilized in the sense that the α_i in (1.5) is chosen as large as possible.

This paper is organized as follows. Section 2 contains the main results concerning the asymptotic normality of the Breslow-type estimator, and Section 3 consists of the lemmas needed in establishing the main results.

2. Asymptotic Normality of the Breslow Estimator

Let $H = \{h | h = (h_1, \ldots, h_I), h_i : [0, \alpha_i) \to [0, \infty)\}$. Here $\alpha_i = \inf\{u | \int_0^u h_i^{(0)}(t) dt = \infty\}$, which implies that $P(T_{ji} - T_{ji-1} \geq \alpha_i) = 0$. The parameter space is $\Theta \times H$. We assume $\alpha_i \leq t_0$, for $i = 1, \ldots, I$.

Let $h_{(0)} = (h_1^{(0)}, \ldots, h_I^{(0)})$, $s_i^{(k)}(\theta, t) = E(S_{J,i}^{(k)}(\theta, t))$ for $k = 0, 1, 2$. Then $s_i^{(0)}, s_i^{(1)}, s_i^{(2)}$ are bounded on $\Theta \times [0, \alpha_i)$. Assume $s_i^{(0)}$ is positive, and, for $i = 1, \ldots, I$, $k = 0, 1, 2$, that $s_i^{(k)}$ is continuous on $\Theta \times [0, \alpha_i)$, and

$$\int_0^{\alpha_i} h_i^{(0)}(u) s_i^{(0)}(\theta_0, u) \, du < \infty, \tag{2.1}$$

$$\lim_{t \to \alpha_i} \sup_{J=1,2,\ldots} E\Big(\int_t^{\alpha_i} S_{J,i}^{(0)}(\theta_0, u) h_i^{(0)}(u) \, du \Big)^2 = 0. \tag{2.2}$$

Under these conditions, we show that $\sup_{J=1,2,\ldots} E(\sup_{0 \leq t < \alpha_i} |G_J^{(i)}(\theta_0, t)|^4) < \infty$, an important step to establish uniform integrability in the proof of Lemma 3.2.

To simplify notation, we extend the domain of definition of $G_J^{(i)}(\theta, t)$ in (1.5) by $G_J^{(i)}(\theta, t) = G_J^{(i)}(\theta, \alpha_i)$ if $t > \alpha_i$. Similar extensions are made also for $h_i^{(0)}$, $S_{J,i}^{(k)}$ and $s_i^{(k)}$. Let $\widehat{\theta}_J$ denote the solution to $G_J(\theta, t_0) = 0$.

Let

$$H_J^{(i)}(\theta,t) = -\sum_{j=1}^{J}\int_0^t \frac{S_{J,i}^{(1)}(\theta,u)}{(S_{J,i}^{(0)}(\theta,u))^2} dN_{ji}(u), \tag{2.3}$$

$$W_J^{(i)}(t) = \sum_{j=1}^{J}\int_0^t \frac{1}{S_{J,i}^{(0)}(\theta_0,u)}\, dM_{ji}(u). \tag{2.4}$$

Then

$$\begin{aligned}
&J^{1/2}(\widehat{\Lambda}_J^{(i)}(t) - \Lambda_i(t)) \\
=&J^{-1/2}\sum_{j=1}^{J}\int_0^t \Big(\frac{1}{S_{J,i}^{(0)}(\widehat{\theta}_J,u)} - \frac{1}{S_{J,i}^{(0)}(\theta_0,u)}\Big)\, dN_{ji}(u) \\
&+J^{1/2}\Big(\sum_{j=1}^{J}\int_0^t \frac{1}{JS_{J,i}^{(0)}(\theta_0,u)} dN_{ji}(u) - \int_0^t h_i^{(0)}(u)\mathbf{1}_{\{\sum_{k=1}^{J} Y_{ki}(u)>0\}} du\Big) + o_p(1) \\
=&J^{-1/2}H_J^{(i)}(\theta_J^*,t)(\widehat{\theta}_J - \theta_0) + J^{-1/2}W_J^{(i)}(t) + o_p(1) \\
=&\frac{1}{J}H_J^{(i)}(\theta_J^*,t)\Big(\frac{-1}{J}\frac{\partial}{\partial\theta}G_J(\widetilde{\theta}_J,t_0)\Big)^{-1}\Big(J^{-1/2}G_J(\theta_0,t_0)\Big) + J^{-1/2}W_J^{(i)}(t) + o_p(1),
\end{aligned} \tag{2.5}$$

where both θ_J^* and $\widetilde{\theta}_J$ are between $\widehat{\theta}_J$ and θ_0.

Let

$$V(t_0) = \sum_{i=1}^{I}\int_0^{t_0} \Big(s_i^{(2)}(\theta_0,u) - \frac{(s_i^{(1)}(\theta_0,u))^2}{s_i^{(0)}(\theta_0,u)}\Big)h_i^{(0)}(u)\, du. \tag{2.6}$$

The main result of this paper is the following theorem concerning the asymptotic distribution of $J^{1/2}(\widehat{\Lambda}_J^{(i)}(t) - \Lambda_i(t))$.

Theorem 2.1. *Assume* (1.1), (1.2), (2.1), (2.2) *and* $V(t_0) > 0$. *Then for* $0 \le t_i \le \alpha_i' < \alpha_i$, $i = 1,2,\ldots,I$, $J^{1/2}(\widehat{\Lambda}_J^{(1)}(t_1) - \Lambda_1(t_1),\ldots,\widehat{\Lambda}_J^{(I)}(t_I) - \Lambda_I(t_I))$ *converges weakly to a mean 0 Gaussian process* $(U_1(t_1),\ldots,U_I(t_I))$ *satisfying*

$$\begin{aligned}
E(U_l(t_l)U_k(t_k)) =& \mathbf{1}_{\{l=k\}}\int_0^{t_l\wedge t_k} \frac{h_l^{(0)}(u)}{s_l^{(0)}(\theta_0,u)}\, du \\
&+V^{-1}(t_0)\Big(\int_0^{t_l} \frac{s_l^{(1)}(\theta_0,u)}{s_l^{(0)}(\theta_0,u)}h_l^{(0)}(u)\, du\Big)\Big(\int_0^{t_k} \frac{s_k^{(1)}(\theta_0,u)}{s_k^{(0)}(\theta_0,u)}h_k^{(0)}(u)\, du\Big).
\end{aligned}$$

Since the theorem follows directly from Lemmas 3.2-3.5, (2.5) and the continuous mapping theorem for weak convergence of stochastic processes, we remark only that the asymptotic normality of $J^{1/2}(\widehat{\theta}_J - \theta_0)$ is a by-product of this argument.

Corollary 2.2. *Under the conditions of Theorem 2.1, and assuming* $h_1^{(0)} = h_2^{(0)} = \cdots = h_I^{(0)} = h_0$, *let* $a_i(t) \geq 0$, *and* $\sum_{i=1}^I a_i(t) = 1$. *Then, for* $0 \leq t \leq \alpha' < \min_i \alpha_i$, $J^{1/2}(\sum_{i=1}^I a_i(t)\widehat{\Lambda}_J^{(i)}(t) - \Lambda_1(t))$ *converges weakly to a mean 0 Gaussian process* $U(t)$ *satisfying* $EU^2(t) = \sum_{1\leq i,j\leq I} a_i(t)a_j(t)V_{ij}(t)$, *where* $V_{ij}(t) = EU_i(t)U_j(t)$.

Remark 1. When we have modulated renewal processes, we can first estimate $V_{ij}(t)$ and then find $\widehat{a}_J^{(i)}(t)$ so that $\sum_{i=1}^I \widehat{a}_J^{(i)}(t) = 1$ and $\sum_{i=1}^I \widehat{a}_J^{(i)}(t)\widehat{\Lambda}_J^{(i)}(t)$ is asymptotically normal with the smallest variance among all estimators $\sum_{i=1}^I a_i(t)$ $\widehat{\Lambda}_J^{(i)}(t)$ studied in Corollary 2.2.

Remark 2. In order to elaborate on the conditions (2.1) and (2.2), we consider the important situation that $Y_j(t) = \mathbf{1}_{(0,C_j]}(t)$ for some independent censoring variable C_j.

In this case, for some constant $C > 0$,

$$\begin{aligned}
&\int_0^{\alpha_i} h_i^{(0)}(u)s_i^{(0)}(\theta_0, u)\,du = \int_0^{\alpha_i} h_i^{(0)}(u)EY_{1i}(u)e^{\theta_0 Z_{1i}(u)}\,du \\
&\leq C\int_0^{\alpha_i} h_i^{(0)}(u)E\mathbf{1}_{(0,T_{1i}-T_{1i-1}]}(u)\mathbf{1}_{(0,t_0]}(T_{1i-1}+u)\mathbf{1}_{(0,C_1]}(T_{1i-1}+u)\,du \\
&\leq C\int_0^{\alpha_i} h_i^{(0)}(u)E\mathbf{1}_{(0,T_{1i}-T_{1i-1}]}(u)E\mathbf{1}_{(0,C_1]}(u)\,du \\
&= C\int_0^{\alpha_i} h_i^{(0)}(u)P(T_{1i}-T_{1i-1}\geq u)P(C_1\geq u)du, \qquad (2.7)
\end{aligned}$$

and

$$\begin{aligned}
&E\Big(\int_0^{\alpha_i} S_{J,i}^{(0)}(\theta_0,u)h_i^{(0)}(u)\,du\Big)^2 \\
&= E\Big(\int_0^{\alpha_i}\Big(\frac{1}{J}\sum_{j=1}^J Y_{ji}(u)e^{\theta_0 Z_{ji}(u)}\Big)h_i^{(0)}(u)\,du\Big)^2 \\
&\leq CE\Big(\int_0^{\alpha_i}\Big(\frac{1}{J}\sum_{j=1}^J Y_{ji}(u)\Big)h_i^{(0)}(u)\,du\Big)^2 \\
&\leq CE\Big(\int_0^{\alpha_i} Y_{1i}(u)h_i^{(0)}(u)\,du\Big)^2 \\
&\leq CE\Big(\int_0^{\alpha_i} h_i^{(0)}(u)\mathbf{1}_{(0,T_{1i}-T_{1i-1}]}(u)\mathbf{1}_{(0,C_1]}(u)\,du\Big)^2 \\
&\leq CE\Big(\int_0^{\alpha_i} h_i^{(0)}(u)\mathbf{1}_{(0,T_{1i}-T_{1i-1}]}(u)\,du\Big)\Big(\int_0^{\alpha_i} h_i^{(0)}(u)\mathbf{1}_{(0,C_1]}(u)\,du\Big) \\
&= C\Big(\int_0^{\alpha_i} h_i^{(0)}(u)P(T_{1i}-T_{1i-1}\geq u)\,du\Big)\Big(\int_0^{\alpha_i} h_i^{(0)}(u)P(C_1\geq u)\,du\Big). \qquad (2.8)
\end{aligned}$$

Thus, a sufficient condition for (2.1) and (2.2) is $\int_0^{\alpha_i} h_i^{(0)}(u)P(T_{1i} - T_{1i-1} \geq u)\,du < \infty$ and $\int_0^{\alpha_i} h_i^{(0)}(u)P(C_1 \geq u)\,du < \infty$. Other sufficient conditions can be derived by examining the calculations in (2.7) and (2.8).

3. Lemmas and Proofs

Lemma 3.1. *Assume that* (2.1) *and* (2.2) *hold. Then*

$$\sup_{J=1,2,\ldots} E(\sup_{0\leq t\leq\alpha_i} (J^{-1/2}G_J^{(i)}(t))^4) < \infty, \tag{3.1}$$

$$\sup_{J=1,2,\ldots} E(\sup_{0\leq t\leq\alpha_i'} (J^{-1/2}W_J^{(i)}(t))^4) < \infty, \tag{3.2}$$

where $0 < \alpha_i' < \alpha_i$ *(*θ_0 *is suppressed in* $G_J^{(i)}$*).*

Proof. We prove (3.1), (3.2) can be done similarly. It follows from the Burkholder-Davis-Gundy inequality (cf. Burkholder, Davis and Gundy (1972) and Lenglart, Lepingle and Pratelli (1980)) that, for some constant $C_1 > 0$,

$$E(\sup_{0\leq t\leq\alpha_i} (G_J^{(i)}(t))^4) \leq C_1 E[G_J^{(i)}]_{\alpha_i}^2, \tag{3.3}$$

where $[G_J^{(i)}]_{\alpha_i}$ is the optional quadratic variation process of $G_J^{(i)}$.

Since $G_J^{(i)}$ is a process of finite variation, we know from Elliott (1982, p.97), for example, that

$$\begin{aligned} E[G_J^{(i)}]_{\alpha_i}^2 &= E\Big(\sum_{t\leq\alpha_i} (\Delta G_J^{(i)}(t))^2\Big)^2 \\ &= E\Big(\sum_{j=1}^J \int_0^{\alpha_i} \Big(Z_{ji}(u) - \frac{S_{J,i}^{(1)}(\theta_0,u)}{S_{J,i}^{(0)}(\theta_0,u)}\Big)^2 dN_{ji}(u)\Big)^2 \\ &\leq 2E < G_J^{(i)} >_{\alpha_i}^2 + 2E(\eta_J(\alpha_i))^2, \end{aligned} \tag{3.4}$$

where $< \cdot >_{\alpha_i}$ is the predictable variation process, and

$$\eta_J(\alpha_i) = \sum_{j=1}^J \int_0^{\alpha_i} \Big(Z_{ji}(u) - \frac{S_{J,i}^{(1)}(\theta_0,u)}{S_{J,i}^{(0)}(\theta_0,u)}\Big)^2 dM_{ji}(u).$$

Using the predictable variation formula, we have

$$\begin{aligned} &\sup_{J=1,2,\ldots} \frac{1}{J^2} E(\eta_J(\alpha_i))^2 \\ = &\sup_{J=1,2,\ldots} \frac{1}{J} E\Big(\frac{1}{J}\sum_{j=1}^J \int_0^{\alpha_i} \Big(Z_{ji}(u) - \frac{S_{J,i}^{(1)}(\theta_0,u)}{S_{J,i}^{(0)}(\theta_0,u)}\Big)^4 h_i^{(0)}(u)Y_{ji}(u)e^{\theta_0 Z_{ji}(u)}du\Big) \end{aligned}$$

$$\leq \sup_{J=1,2,\ldots} \frac{C_2}{J} E\Big(\int_0^{\alpha_i} S_{J,i}^{(0)}(\theta_0, u) h_i^{(0)}(u)\, du \Big) < \infty, \tag{3.5}$$

by (2.1), and

$$\begin{aligned} &\sup_{J=1,2,\ldots} \frac{1}{J^2} E < G_J^{(i)} >_{\alpha_i}^2 \\ &= \sup_{J=1,2,\ldots} E\Big(\frac{1}{J} \sum_{j=1}^{J} \int_0^{\alpha_i} \Big(Z_{ji}(u) - \frac{S_{J,i}^{(1)}(\theta_0, u)}{S_{J,i}^{(0)}(\theta_0, u)} \Big)^2 h_i^{(0)}(u) Y_{ji}(u) e^{\theta_0 Z_{ji}(u)}\, du \Big)^2 \\ &\leq \sup_{J=1,2,\ldots} C_3 E\Big(\int_0^{\alpha_i} S_{J,i}^{(0)}(\theta_0, u) h_i^{(0)}(u)\, du \Big)^2 < \infty, \end{aligned} \tag{3.6}$$

by (2.2). Here C_2 and C_3 are some constants. From (3.3), (3.4), (3.5) and (3.6), we get (3.1). This completes the proof.

Lemma 3.2. *Assume that* (2.1) *and* (2.2) *hold. Then*

(i) $J^{-1/2}(G_J^{(1)}(\theta_0, \cdot), \ldots, G_J^{(I)}(\theta_0, \cdot), W_J^{(1)}(\cdot), \ldots, W_J^{(I)}(\cdot))$ *converges weakly to a multivariate mean* 0 *independent Gaussian martingale* $(G_1(\cdot), \ldots, G_I(\cdot), W_1(\cdot), \ldots, W_I(\cdot))$ *on* $[0, \alpha']$ *satisfying*

$$EG_i^2(t) = \int_0^t \Big(s_i^{(2)}(\theta_0, u) - \frac{(s_i^{(1)}(\theta_0, u))^2}{s_i^{(0)}(\theta_0, u)} \Big) h_i^{(0)}(u)\, du, \tag{3.7}$$

$$EW_i^2(t) = \int_0^t \frac{h_i^{(0)}(u)}{s_i^{(0)}(\theta_0, u)}\, du, \tag{3.8}$$

where $0 \leq t \leq \alpha' < \min_{1 \leq i \leq I} \alpha_i$.

(ii) *The weak convergence of* $J^{-1/2}(G_J^{(1)}(\theta_0, t), \ldots, G_J^{(I)}(\theta_0, t))$ *is also valid for* $0 \leq t \leq t_0$.

We note that $EG_i^2(t) = EG_i^2(t \wedge \alpha_i)$ *for every* $t > 0$.

Proof. It follows from the Martingale Central Limit Theorem that both $J^{-1/2} G_J^{(i)}(\theta_0, \cdot)$ and $J^{-1/2} W_J^{(i)}(\cdot)$ converge weakly as random elements in the Skorohod space $D[0, \alpha']$. Let $G_i(\cdot)$ and $W_i(\cdot)$ denote their limiting processes respectively. It is straightforward to see both (3.7) and (3.8) are satisfied. This implies that $J^{-1/2}(G_J^{(1)}(\theta_0, \cdot), \ldots, G_J^{(I)}(\theta_0, \cdot), W_J^{(1)}(\cdot), \ldots, W_J^{(I)}(\cdot))$ is a tight sequence in the product space $(D[0, \alpha'])^{2I}$. With this understanding, it suffices to show that every weakly convergent subsequence has the same limiting distribution. For this, we will show that $EG_l(s)G_k(t) = EW_l(s)W_k(t) = 0$ for $1 \leq k \neq l \leq I$, and $0 \leq s \leq t \leq \alpha'$.

It follows from Lemma 3.1 and the Schwarz inequality that

$$\begin{aligned}&\sup_J E(\frac{1}{J}G_J^{(l)}(s)G_J^{(k)}(t))^2\\ \le&\sup_J E^{\frac{1}{2}}(J^{-1/2}G_J^{(l)}(s))^4E^{\frac{1}{2}}(J^{-1/2}G_J^{(k)}(t))^4\\ \le&(\sup_J E^{\frac{1}{2}}(J^{-1/2}G_J^{(l)}(s))^4)(\sup_J E^{\frac{1}{2}}(J^{-1/2}G_J^{(k)}(t))^4)<\infty.\end{aligned}$$

This shows that $\{J^{-1}G_J^{(l)}(s)G_J^{(k)}(t)|J=1,2,\ldots\}$ is uniformly integrable. This, together with weak convergence, shows that if $l<k$, $EG_l(s)G_k(t)=\lim_{J\to\infty}E\frac{1}{J}G_J^{(l)}(s)G_J^{(k)}(t)=\lim_{J\to\infty}\frac{1}{J}EG_J^{(l)}(s)E(G_J^{(k)}(t)|\mathcal{G}_{J,0}^{(k)})=0$, where $\mathcal{G}_{J,t}^{(k)}\equiv\sigma\{\mathcal{F}_{j,T_{jk}+t}|j=1,\ldots,J\}$. Similarly, we can show the orthogonality of other components. This proves (i). Since (ii) can be proved in the same manner, the proof is omitted.

Lemma 3.3.

(i) $\sup_{0\le t\le\alpha_i'}|J^{-1}H_J^{(i)}(\theta_0,t)+\int_0^t\frac{s_i^{(1)}(\theta_0,u)}{s_i^{(0)}(\theta_0,u)}h_i^{(0)}(u)\,du|$ *converges to 0 in probability, where* $0<\alpha_i'<\alpha_i$.

(ii) $\sup_{0\le t\le\alpha_i'}J^{-1}|H_J^{(i)}(\theta_J^*,t)-H_J^{(i)}(\theta_0,t)|$ *converges to 0 in probability, where* θ_J^* *is given in* (2.5), *and* $0<\alpha_i'<\alpha_i$.

The proof for Lemma 3.3 is omitted, see Andersen and Gill (1982).

Lemma 3.4. $\widehat{\theta}_J$ *converges to* θ_0 *in probability.*

Proof. Let

$$X_J^{(i)}(\theta,t)=J^{-1}\sum_{j=1}^J\int_0^t\Big((\theta-\theta_0)Z_{ji}(u)-\log\frac{S_{J,i}^{(0)}(\theta,u)}{S_{J,i}^{(0)}(\theta_0,u)}\Big)\,dN_{ji}(u),$$

$$A_J^{(i)}(\theta,t)=J^{-1}\sum_{j=1}^J\int_0^t\Big((\theta-\theta_0)Z_{ji}(u)-\log\frac{S_{J,i}^{(0)}(\theta,u)}{S_{J,i}^{(0)}(\theta_0,u)}\Big)h_i^{(0)}(u)Y_{ji}(u)e^{\theta_0Z_{ji}(u)}\,du,$$

which is equal to $\int_0^t((\theta-\theta_0)S_{J,i}^{(1)}(\theta_0,u)-\log\frac{S_{J,i}^{(0)}(\theta,u)}{S_{J,i}^{(0)}(\theta_0,u)}\cdot S_{J,i}^{(0)}(\theta_0,u))h_i^{(0)}(u)\,du$. Since $X_J^{(i)}(\theta,t)-A_J^{(i)}(\theta,t)$ is a martingale, we know

$$\begin{aligned}&E(X_J^{(i)}(\theta,\alpha_i)-A_J^{(i)}(\theta,\alpha_i))^2\\ &=J^{-2}\sum_{j=1}^J\int_0^{\alpha_i}\Big((\theta-\theta_0)Z_{ji}(u)-\log\frac{S_{J,i}^{(0)}(\theta,u)}{S_{J,i}^{(0)}(\theta_0,u)}\Big)^2h_i^{(0)}(u)Y_{ji}(u)e^{\theta_0Z_{ji}(u)}\,du\\ &\le C\cdot J^{-1}\int_0^{\alpha_i}S_{J,i}^{(0)}(\theta_0,u)h_i^{(0)}(u)\,du,\end{aligned}\tag{3.9}$$

which converges to 0 in probability as a consequence of (2.2).

Let

$$A^{(i)}(\theta,t)=\int_0^t\Big((\theta-\theta_0)s_i^{(1)}(\theta_0,u)-\log\frac{s_i^{(0)}(\theta,u)}{s_i^{(0)}(\theta_0,u)}\cdot s_i^{(0)}(\theta_0,u)\Big)h_i^{(0)}(u)\,du.$$

It follows from (2.1) and (2.2) that $A^{(i)}(\theta,\alpha_i)$ is well-defined, and for every $t\in(0,\alpha_i)$, $\lim_{J\to\infty}E(A_J^{(i)}(\theta,t)-A^{(i)}(\theta,t))^2=0$, and $\lim_{t\to\alpha_i}\sup_J E(A_J^{(i)}(\theta,\alpha_i)-A^{(i)}(\theta,t))^2=0$. These imply that $E(A_J^{(i)}(\theta,\alpha_i)-A^{(i)}(\theta,\alpha_i))^2$ converges to 0. This together with (3.9) shows that $X_J^{(i)}(\theta,\alpha_i)$ converges to $A^{(i)}(\theta,\alpha_i)$ in probability. Hence $\sum_{i=1}^I X_J^{(i)}(\theta,\alpha_i)$ also converges in probability to $\sum_{i=1}^I A^{(i)}(\theta,\alpha_i)$. It now follows from the convex analysis arguments used in Andersen and Gill (1982) that $\widehat{\theta}_J$ converges to θ_0 in probability. This completes the proof.

Lemma 3.5.
(i) $J^{-1}\frac{\partial}{\partial\theta}G_J(\theta_0,t_0)+V(t_0)$ *converges to 0 in probability,*
(ii) $J^{-1}\frac{\partial}{\partial\theta}G_J(\widetilde{\theta}_J,t_0)-J^{-1}\frac{\partial}{\partial\theta}G_J(\theta_0,t_0)$ *converges to 0 in probability, where* $\widetilde{\theta}_J$ *is given in* (2.5) *and* $V(t_0)$ *is given in* (2.6).

Since (i) can be proved using the arguments in the proof of Lemma 3.4, and (ii) is an easy consequence of Lemma 3.4, we omit the proof.

Acknowledgements

Thanks are due to an anonymous referee for valuable suggestions that led to an improvement of the presentation of the paper.

References

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10**, 1100-1120.

Burkholder, D. L., Davis, B. J. and Gundy, R. F. (1972). Inequalities for convex functions of operators on martingales. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2**, 223-240.

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.

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. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34**, 187-220.

Cox, D. R. (1986). Some remarks on semi-Markov processes in medical statistics. In *Semi-Markov Models* (Edited by J. Janssen), 411-421. Plenum Press, New York.

Cox, D. R. (1997). Some remarks on the analysis of survival data. In *Proceedings of the first Seattle Symposium in Biostatistics: Survival Analysis* (Edited by D. Y. Lin and T. R. Fleming), 1-10. Springer-Verlag, New York.

Elliott, R. J. (1982). *Stochastic Calculus and Applications.* Springer-Verlag, New York.

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

Keiding, N. (1986). Statistical analysis of semi-Markov processes. In *Semi-Markov Models* (Edited by J. Janssen), 301-315. Plenum Press, New York.

Lenglart, E., Lepingle, D. and Pratelli, M.(1980). Presentation unifiee de certaines inéqualities de la théorie des martingales. Séminaire de Probabilitiés. Lecture Notes in Math. **14,** 24-48. Springer-Verlag, New York.

Oakes, D. and Cui, L. (1994). On semiparametric inference for modulated renewal processes. *Biometrika* **81**, 83-90.

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

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

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

National Health Research Institutes, Academia Sinica, Nankang, Taipei, Taiwan.

E-mail: hsiung@nhri.org.tw

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

E-mail: shihwei@ms3.hinet.net

(Received December 1998; accepted March 2000)