Journal of the Korean Statistical Society (2006), **35**: *3, pp 251–260*

ALMOST SURE LIMITS OF SAMPLE ALIGNMENTS IN PROPORTIONAL HAZARDS MODELS

JOHAN LIM[1] AND SEUNG-JEAN KIM[2]

ABSTRACT

The proportional hazards model (PHM) can be associated with a non-homogeneous Markov chain (NHMC) in the sense that sample alignments in the PHM correspond to trajectories of the NHMC. As a result the partial likelihood widely used for the PHM is a probabilistic function of the trajectories of the NHMC. In this paper, we show that, as the total number of subjects involved increases, the trajectories of the NHMC, *i.e.* sample alignments in the PHM, converges to the solution of an ordinary differential equation which, subsequently, characterizes the almost sure limit of the partial likelihood.

AMS 2000 subject classifications. Primary 62N99; Secondary 60K10.
Keywords. Non-homogeneous Markov chain, proportional hazards model.

1. INTRODUCTION

Let T be a continuous random time associated with an event, representing the survival or event time. The hazards rate of T is the instantaneous conditional probability defined by

$$\lambda(t) = \lim_{\Delta \to 0} \mathbb{P}(t \le T \le t + \Delta | T \ge t)/\Delta.$$

The probability that the survival time is larger than t is given by

$$\mathbb{P}(T > t) = \exp\Big(- \int_0^t \lambda(s)\, ds \Big).$$

Received May 2006; accepted July 2006.

[1]Corresponding author. Department of Applied Statistics, Yonsei University, Seoul 120-749, Korea (e-mail: johanlim@yonsei.ac.kr)

[2]Department of Electrical Engineering, Stanford University, Stanford, CA 94305, U.S.A. (e-mail: sjkim@stanford.edu)

Specifying the hazards rate is therefore equivalent to specifying the probability distribution of T. In practice, hazards rate model is more commonly used (Kalbfleish and Prentice, 1980; Therneau and Grambsch, 2000).

In practical applications, many assumptions are made on the shape of the hazards rate. In particular, the proportional hazards model (PHM) assumes that the hazards rate for a subject associated with a covariate x is

$$\lambda(t|x) = \lambda_0(t)\exp(x\beta), \tag{1.1}$$

where $\lambda_0(t)$ is called as a baseline hazards function. In other words, the proportional hazards assumes that the ratio of hazards rate for two subjects associated with covariates x_1 and x_2 is

$$\frac{\lambda(t|x_2)}{\lambda(t|x_1)} = \exp((x_2 - x_1)\beta),$$

which is constant over time. In particular, the semi-parametric model with an unspecified baseline function $\lambda_0(t)$ is called the Cox PHM. (For more on the model, see Cox (1972, 1975)).

In the Cox PHM, the main emphasis is placed on estimating the parameter β. A common method for finding β maximizes the partial likelihood (PL) introduced by Cox (1975). The PL for the observations $\{(x_i, t_i),\ i = 1, 2, \ldots, n\}$, where t_i is the observed survival time of the i^{th} subject with the covariate x_i, is

$$\mathrm{PL}(\beta) = \prod_{i=1}^{n} \left(\frac{\exp(x_i\beta)}{\sum_{j\in R_i} \exp(x_j\beta)} \right), \tag{1.2}$$

where R_i is the risk set at time t_i containing all indices of survivors at that time.

The PL is different from the true likelihood function in the sense that it uses only the information on the ranks of the data. For instance, let us consider a two sample problem with 3 subjects in each sample. Here, sample is the set of subjects. There are total $n = 6$ subjects. Three of them are associated with the covariate $x = 0$ and the other three are with $x = 1$, that is, a covariate indicates the sample the subject is from. Suppose that the hazards rate of the subjects in the sample 1 is 1 and that in the sample 2 is λ. To be specific, $\lambda_0(t) = 1$ and $\lambda = \exp(\beta)$ in (1.1). The following two scenarios (subjects in order) then have the same partial likelihood:

- Scenario 1: 2, 5, (7), (9), 5, (12),
- Scenario 2: 2, 6, (8), (10), 12, (17).

Here the numbers in parentheses represent the survival times of the subjects from the sample 1 and the numbers not in parenthesis are those for the subjects from sample 2. In both scenarios, the PL(λ) is given by

$$\mathrm{PL}(\lambda) = \Big(\frac{3}{3+3\lambda}\Big)\Big(\frac{2}{2+3\lambda}\Big)\Big(\frac{3\lambda}{1+3\lambda}\Big)\Big(\frac{2\lambda}{1+2\lambda}\Big)\Big(\frac{2\lambda}{1+2\lambda}\Big)\Big(\frac{1}{1+\lambda}\Big). \tag{1.3}$$

Note that it depends only on the observed alignment of subjects labeled by their sample indicators.

We first describe the NHMC model for the two-sample PHM, and then move toward the general case. Let the two-sample PHM have n subjects for each sample. There are total $2n$ subjects. We associate the PHM with a urn containing n white and n black balls where the color represent the sample. We assume that the weights of balls with the same color are equal but the weight of a white ball is different from that of a black one. We draw a ball randomly from the urn without replacement in accordance with the model: the probability of drawing a white (black) ball is proportional to the total weights of white (black) balls which are currently in the urn. Finally, we have an alignment of n white balls (weight $=1$) and n black balls (weight $=\lambda\geq 1$) in the order of the time they are chosen. Let X_k be the indicator function of the event that the ball on the k^{th} draw is white, and let $S_k = X_1 + X_2 + \cdots + X_k$ be the number of white balls we have until the k^{th} draw. Given $S_k = m$, the probability that a white ball appears in the $(k+1)^{th}$ draw is $(n-m)\lambda/\{(n-k+m)+(n-m)\lambda\}$, which only depends on the value of S_k. Thus, $\{S_k\}_{k=1}^{2n}$ can be identified with a non-homogeneous Markov chain (NHMC) whose transition probability is

$$S_{k+1} = \begin{cases} S_k + 1 \text{ with probability } \frac{(n-S_k)}{(n-S_k)+(n-k+S_k)\lambda}, \\ S_k \qquad \text{with probability } \frac{(n-k+S_k)\lambda}{(n-S_k)+(n-k+S_k)\lambda}. \end{cases} \tag{1.4}$$

This, in turn, implies that the PL becomes the probability of the trajectory $\{S_k\}_{k=1}^{2n}$ of the NHMC. For example, the PL of the aforementioned two scenarios is

$$\begin{aligned}\mathrm{PL}(\lambda) &= \mathbb{P}(S_1 = 1,\ S_2 = 2,\ S_3 = 2,\ S_4 = 2,\ S_5 = 3,\ S_6 = 3) \\ &= \mathbb{P}(S_1 = 1)\mathbb{P}(S_2 = 2|S_1 = 1)\cdots\mathbb{P}(S_6 = 3|S_5 = 3).\end{aligned} \tag{1.5}$$

Many interesting statistics can be associated with the above NHMC. In this paper, we characterize the almost sure limit of $S_{[nt]}/n$ as the solution to the ordinary dimensional differential equation:

$$d\mathbf{M}_\infty(t;\lambda) = \frac{(1-\mathbf{M}_\infty(t;\lambda))\lambda}{(1-\mathbf{M}_\infty(t;\lambda))\lambda + (1-t+\mathbf{M}_\infty(t;\lambda))}\,dt, \quad t\in[0,2] \tag{1.6}$$

with the initial condition $\mathbf{M}_\infty(0;\lambda)=0$. To be specific, in the two sample PHM, for every fixed $\lambda \geq 1$,

$$\lim_{n\to\infty}\sup_{t\in[0,2]}\left|\frac{1}{n}S_{[nt]} - \mathbf{M}_\infty(t;\lambda)\right| = 0 \quad \text{almost surely.} \tag{1.7}$$

The convergence in (1.7) shows that the number of white balls we have until the S_k^{th} trial in the urn example described above is approximately $n\mathbf{M}_\infty(k/n;\lambda)$ for n large enough.

It is easy to show that

$$\mathbf{M}_\infty(t;1) = \frac{t}{2} \equiv \underline{\mathbf{M}}_\infty(t), \qquad \forall t\in[0,2].$$

and

$$\lim_{\lambda\to\infty}\mathbf{M}_\infty(t;\lambda) = \min\ (t,1) \equiv \overline{\mathbf{M}}_\infty(t), \qquad 0\leq t\leq 2.$$

Moreover, the limit function $\mathbf{M}_\infty(t;\lambda)$ in (1.7) is monotone in λ, which is easy to see from

$$\mathbf{M}_\infty(t;\lambda) = \int_0^t \frac{(1-\mathbf{M}_\infty(s;\lambda))\lambda}{(1-\mathbf{M}_\infty(s;\lambda))\lambda + (1-s+\mathbf{M}_\infty(s;\lambda))}\,ds. \tag{1.8}$$

Along with the monotonicity of the integrand in (1.8):

$$1\leq\lambda_1\leq\lambda_2 \qquad \Longrightarrow \qquad \mathbf{M}_\infty(t;\lambda_1)\leq\mathbf{M}_\infty(t;\lambda_2), \quad 0\leq t\leq 2.$$

The lower and upper bound on $\mathbf{M}_\infty(t;\lambda)$ can now be computed as

$$\underline{\mathbf{M}}_\infty(t) = \frac{t}{2} \leq \mathbf{M}_\infty(t;\lambda) \leq \min\big(t,1\big) = \overline{\mathbf{M}}_\infty(t), \qquad 0\leq t\leq 2.$$

Figure 1.1 plots the almost sure limit of $\mathbf{M}_\infty(t)$ for two values of λ as well the lower bound $\underline{\mathbf{M}}_\infty(t)$ and the upper bound $\overline{\mathbf{M}}_\infty(t)$.

A general p-sample PHM with n subjects for each sample is analogous to a urn containing total np balls, where the balls are classified as their colors, denoted by $1,\ldots,p$, and there are n balls for each color. The weight of a ball for the k^{th} color is λ_k with $1=\lambda_1\leq\lambda_2\leq\cdots\leq\lambda_p$. Suppose that we draw a ball randomly from the urn. Then the order of balls drawn from the urn (*i.e.*, a

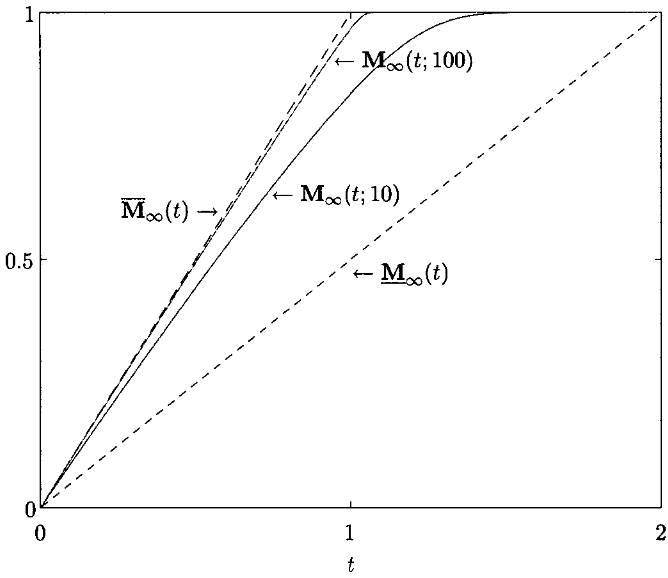

FIGURE 1.1 *Almost sure limit* $\mathbf{M}_\infty(t;\lambda)$ *for* $\lambda = 10$ *and* $\lambda = 100$ *and the upper bound* $\overline{\mathbf{M}}_\infty(t)$ *and lower bound* $\underline{\mathbf{M}}_\infty(t)$.

general p sample PHM) can be associated with a p-dimensional NHMC as follows. Let $S_k = (S_k^1, S_k^2, \ldots, S_k^p)$ where S_k^q is the number of balls of color q that have appeared until the k^{th} draw. Then, $\{S_k\}_{k=1}^{np}$ follows the p-dimensional NHMC whose transition probability, given S_k, is

$$S_{k+1} = S_k + e_q \quad \text{with probability } \frac{(n - S_k^q)\lambda_q}{\sum_l (n - S_k^l)\lambda_l}, \tag{1.9}$$

where e_q is the p-dimensional unit vector whose q^{th} element is 1 and other ones are zero.

The following theorem extends the convergence in (1.7) to the p-sample PHM and characterizes the almost sure limit of the trajectories of its NHMC.

THEOREM 1.1. *For a fixed* $\lambda = (\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p$ *with* $1 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_p$,

$$\lim_{n\to\infty} \sup_{t\in[0,p]} \max_{1\le k\le p} \left| \frac{1}{n} S_{[nt]}^k - \mathbf{M}_\infty^k(t;\lambda) \right| = 0 \quad \textit{almost surely}, \tag{1.10}$$

where $\mathbf{M}_\infty(t;\lambda) = (\mathbf{M}_\infty^1(t;\lambda), \ldots, \mathbf{M}_\infty^p(t;\lambda))$ *be the solution to the ordinary* p

dimensional differential equation on $[0, p]$

$$dM_\infty^k(t;\lambda) = \frac{(1 - M_\infty^k(t;\lambda))\lambda_k}{\sum_q (1 - M_\infty^q(t;\lambda))\lambda_q}\, dt, \quad \textit{for} \quad k = 1, 2, \ldots, p, \tag{1.11}$$

with $M_\infty(0, \lambda) = (0, \ldots, 0)$.

A possible application of Theorem 1.1 to statistics is the procedure to test $\lambda = \lambda_0$ against $\lambda \neq \lambda_0$ for the two sample PHM. The suggested testing statistics would be

$$\mathbf{T}_n = \frac{1}{n} \sup_{t \in [0,2]} \left| S_{[nt]} - n M_\infty(t, \lambda_0) \right|$$

whose almost sure limit for every λ is well understood.

2. Proof of Theorem 1.1

We first prove Theorem 1.1 for the two sample PHM ($p = 2$) and then extends it to the general p-sample PHM.

2.1. *Proof of the case* $p = 2$

We will prove the convergence in (1.7): for every fixed $\lambda > 0$,

$$\lim_{n \to \infty} \sup_{t \in [0,2]} \left| \frac{1}{n} S_{[nt]} - M_\infty(t) \right| = 0 \quad \text{almost surely.}$$

From the triangular inequality

$$\left| \frac{1}{n} S_{[nt]} - M_\infty(t) \right| \leq \left| \frac{1}{n} S_{[nt]} - \mathbb{E}\left(\frac{1}{n} S_{[nt]}\right) \right| + \left| \mathbb{E}\left(\frac{1}{n} S_{[nt]}\right) - M_\infty(t) \right|,$$

we decompose the proof into two parts:

LEMMA 2.1.

$$\lim_{n \to \infty} \sup_{t \in [0,2]} \left| \frac{1}{n} S_{[nt]} - \mathbb{E}\left(\frac{1}{n} S_{[nt]}\right) \right| = \max_{k=1}^{2n} \left| \frac{1}{n} S_k - \mathbb{E}\left(\frac{1}{n} S_k\right) \right| = 0 \quad \textit{almost surely,} \tag{2.1}$$

and

LEMMA 2.2.

$$\lim_{n \to \infty} \sup_{t \in [0,2]} \left| \mathbb{E}\left(\frac{1}{n} S_{[nt]}\right) - M_\infty(t) \right| = \max_{k=1}^{2n} \left| \mathbb{E}\left(\frac{1}{n} S_k\right) - M_\infty\left(\frac{k}{n}\right) \right| = 0. \tag{2.2}$$

The convergence in (1.7) is an immediate consequence of the two lemmas.

PROOF OF LEMMA 2.1. The proof is based on Markov's inequality (Billingsley, 1995, p. 80) along with the fact that the correlation between X_i and X_j is negative.

Let $Y_i = X_i - \mathbb{E}(X_i)$ and $S_k - \mathbb{E}(S_k) = \sum_{i=1}^k Y_i$. First, we will show the following results on the negative dependency between Y_is: for every distinct integers $1 \le i, j, k, l \le 2n$, (i) $\mathbb{E}(Y_i \cdot Y_i) < 0$, (ii) $\mathbb{E}(Y_i^3 Y_j) < 0$, (iii) $\mathbb{E}(Y_i^2 Y_j Y_k) < 0$ and (iv) $\mathbb{E}(Y_i Y_j Y_k Y_l) < 0$.

To show (i), it suffices to show that

$$\mathbb{P}(X_j = 1 | X_i = 1) < \mathbb{P}(X_j = 1), \tag{2.3}$$

which follows from

$$\begin{aligned} &\mathbb{P}(x_1, \dots, x_{i-1},\ X_i = 1,\ x_{i+1}, \dots, x_{j-1},\ X_j = 1) \\ &< \mathbb{P}(x_1, \dots, x_{i-1},\ X_i = 0,\ x_{i+1}, \dots, x_{j-1},\ X_j = 1) \end{aligned}$$

for any realization $x_1, \dots, x_{i-1},\ x_{i+1}, \dots, x_{j-1}$. The claim (ii) can be proved as follows:

$$\mathbb{E}(Y_i^3 Y_j) = \mathbb{E}[(X_i - \mathbb{E}(X_i)^3)(X_j - \mathbb{E}(X_j))] = \mathbb{E}(Y_i Y_j) \le 0.$$

We can also show (iii) and (iv) in a similar way.

Now we will show that

$$\sum_n \mathbb{P}\big(\max_{1 \le k \le 2n} |S_k - \mathbb{E}(S_k)| > n\epsilon\big) < \infty, \tag{2.4}$$

which results in (2.1) from the Borel Cantelli Lemma (Billingsley, 1995, Theorem 4.3). Note from the Markov inequality that

$$\begin{aligned} \mathbb{P}\big(\max_k |S_k - \mathbb{E}(S_k)| > n\epsilon\big) &\le \sum_{k=1}^n \mathbb{P}\Big(\Big|\sum_{i=1}^k Y_i\Big| > n\epsilon\Big) \\ = \sum_{k=1}^n \frac{1}{n^4 \epsilon^4} \Big\{ \sum_{i=1}^k \mathbb{E}(Y_i^4) &+ \sum_{i \ne j} \mathbb{E}(Y_i^3 Y_j) + \sum_{i \ne j} \mathbb{E}(Y_i^2 Y_j^2) \\ &+ \sum_{i \ne j \ne k} \mathbb{E}(Y_i^2 Y_j Y_k) + \sum_{i \ne j \ne k \ne l} \mathbb{E}(Y_i Y_j Y_k Y_l) \Big\} \qquad (2.5) \\ \le \sum_{k=1}^n \frac{1}{n^4 \epsilon^4} \Big\{ \sum_{i=1}^k \mathbb{E}(Y_i^4) &+ \sum_{i \ne j} \mathbb{E}(Y_i^2 Y_j^2) \Big\} \\ \approx O\Big(\frac{1}{n^2}\Big), & \end{aligned}$$

where $i \neq j \neq k$ (respectively, $i \neq j \neq k \neq l$) in (2.5) means that $i \neq j,\ i \neq k$ and $j \neq k$ (respectively, $i \neq j,\ i \neq k,\ i \neq l,\ j \neq k,\ j \neq l$ and $k \neq l$) □

PROOF OF LEMMA 2.2. Now we will show the convergence of $\mathbb{E}(S_{[nt]}/n)$ to $\mathbf{M}_\infty(t)$. To do it, we first note that, by conditioning S_k, $\mathbb{E}(S_{k+1}/n)$ can be written as

$$\mathbb{E}\Big(\frac{S_{k+1}}{n}\Big) = \mathbb{E}\Big(\frac{S_k}{n}\Big) + \frac{1}{n}\mathbb{E}\Big[f\Big(\frac{S_k}{n}\Big)\Big],$$

where

$$f(x) = \frac{(1-x)\lambda}{(1-x)\lambda + (1-\frac{k}{n}+x)}.$$

Then, $\mathbb{E}(S_{k+1}/n) - \mathbf{M}_\infty\{(k+1)/n\}$ is decomposed as

$$\begin{aligned}\mathbb{E}\Big(\frac{S_{k+1}}{n}\Big) - \mathbf{M}_\infty\Big(\frac{k+1}{n}\Big) &= \mathbb{E}\Big(\frac{S_k}{n}\Big) - \mathbf{M}_\infty\Big(\frac{k}{n}\Big) \\ &\quad + \mathbf{M}_\infty\Big(\frac{k}{n}\Big) - \mathbf{M}_n\Big(\frac{k}{n}\Big) + \mathbf{M}_n\Big(\frac{k+1}{n}\Big) - \mathbf{M}_\infty\Big(\frac{k+1}{n}\Big) \\ &\quad + \frac{1}{n}\mathbb{E}\Big[f\Big(\frac{S_k}{n}\Big)\Big] - \frac{1}{n}f\Big(\mathbf{M}_n\Big(\frac{k}{n}\Big)\Big). \qquad (2.6)\end{aligned}$$

Using the decomposition recursively, we then have

$$\Big|\mathbb{E}\Big(\frac{S_{k+1}}{n}\Big) - \mathbf{M}_\infty\Big(\frac{k+1}{n}\Big)\Big| \leq \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3,$$

where

$$\begin{aligned}\mathbf{A}_1 &= \Big|\mathbb{E}\Big(\frac{S_1}{n}\Big) - \mathbf{M}_\infty\Big(\frac{1}{n}\Big)\Big|, \\ \mathbf{A}_2 &= \Big|\mathbf{M}_n\Big(\frac{k+1}{n}\Big) - \mathbf{M}_\infty\Big(\frac{k+1}{n}\Big)\Big| + \Big|\mathbf{M}_n\Big(\frac{1}{n}\Big) - \mathbf{M}_\infty\Big(\frac{1}{n}\Big)\Big|, \\ \mathbf{A}_3 &= \sum_{i=1}^{k}\frac{1}{n}\Big|\mathbb{E}\Big[f\Big(\frac{S_i}{n}\Big)\Big] - f\Big(\mathbf{M}_n\Big(\frac{i}{n}\Big)\Big)\Big|.\end{aligned}$$

It is easy to show that both $\mathbf{A}_1$ and $\mathbf{A}_2$ are $O(1/n)$ for every k. Thus, we only need to show that $\mathbf{A}_3$ converges to 0 as n increases.

The summand in $\mathbf{A}_3$ is smaller than each of

$$\frac{1}{n}\Big|\mathbb{E}\Big[f\Big(\frac{S_i}{n}\Big)\Big] - f\Big(\mathbb{E}\Big(\frac{S_i}{n}\Big)\Big)\Big| + \frac{1}{n}\Big|f\Big(\mathbb{E}\Big(\frac{S_i}{n}\Big)\Big) - f\Big(\mathbf{M}_n\Big(\frac{i}{n}\Big)\Big)\Big|$$

and

$$\frac{2\lambda}{2n-i}\Big\{\mathbb{E}\Big|\frac{S_i}{n} - \mathbb{E}\Big(\frac{S_i}{n}\Big)\Big| + \Big|\mathbb{E}\Big(\frac{S_i}{n}\Big) - \mathbf{M}_n\Big(\frac{i}{n}\Big)\Big|\Big\},$$

which is an easy consequence of the fact that

$$\frac{1}{n}\big|\ f(y) - f(x)\ \big| \leq \frac{2\lambda}{2n-k}\big|y-x\big|.$$

Therefore, for $k = 1, 2, \ldots, 2n-1$,

$$\begin{aligned}\mathbf{A}_3 \leq \sum_{i=1}^{k} \frac{2\lambda}{2n-i}\Big\{\mathbb{E}\ \Big|\ \frac{S_i}{n} - \mathbb{E}\Big(\frac{S_i}{n}\Big)\ \Big| + \Big|\ \mathbb{E}\Big(\frac{S_i}{n}\Big) - \mathbf{M}_\infty\Big(\frac{i}{n}\Big)\ \Big| \\ + \Big|\ \mathbf{M}_\infty\Big(\frac{i}{n}\Big) - \mathbf{M}_n\Big(\frac{i}{n}\Big)\ \Big|\Big\}. \qquad (2.7)\end{aligned}$$

In (2.7), from the negative correlation between Y_i's, we have

$$\begin{aligned}\mathbb{E}\ \Big|\frac{S_i}{n} - \mathbb{E}\Big(\frac{S_i}{n}\Big)\Big| &\leq \Big\{\mathbb{E}\ \Big[\frac{S_i}{n} - \mathbb{E}\Big(\frac{S_i}{n}\Big)\ \Big]^2\Big\}^{1/2} \\ &= \Big\{\frac{1}{n^2}\mathrm{Var}\Big(\sum_{j=1}^{i} Y_j\Big)\Big\}^{1/2} \approx O\Big(n^{-1/2}\Big), \qquad (2.8)\end{aligned}$$

and

$$\mathbf{A}_3 \leq 2\lambda\frac{1}{\sqrt{n}}\sum_{i=1}^{2n-1}\frac{1}{i} \approx O\Big(\frac{\log n}{n^{1/2}}\Big).$$

Thus, for every $k = 1, 2, \ldots, 2n-1$,

$$\mathbb{E}\Big(\frac{S_k}{n}\Big) - \mathbf{M}_\infty\Big(\frac{k}{n}\Big) \approx O\Big(n^{-1/2}\Big).$$

This leads to the assertion in Lemma 2.2. □

2.2. Proof of the general p-sample PHM

We will show that for each $q = 1, 2, \ldots, p$,

$$\lim_{n\to\infty}\sup_{t\in[0,p]}\Big|\frac{1}{n}S^q_{[nt]} - \mathbf{M}^q_\infty(t;\lambda)\Big| = 0 \quad \text{almost surely.} \qquad (2.9)$$

Let $\mathbf{M}^q_n(t)$ and $\mathbf{M}_n = (\mathbf{M}^1_n(t), \ldots, \mathbf{M}^p_n(t))$ be the solution to the finite difference equation as in the two sample PHM. Again, we decompose the proof into two parts:

$$\lim_{n\to\infty}\sup_{t\in[0,p]}\Big|\frac{1}{n}S^q_{[nt]} - \mathbb{E}\Big(\frac{1}{n}S^q_{[nt]}\Big)\Big| = \max_{k=1}^{pn}\Big|\frac{1}{n}S^q_k - \mathbb{E}\Big(\frac{1}{n}S^q_k\Big)\Big| = 0 \quad \text{almost surely} \qquad (2.10)$$

and

$$\lim_{n\to\infty}\sup_{t\in[0,p]}\left|\mathbb{E}\Big(\frac{1}{n}S^q_{[nt]}\Big)-\mathbf{M}^q_\infty(t)\right|=\max_{1\le k\le pn}\left|\mathbb{E}\Big(\frac{1}{n}S^q_k\Big)-\mathbf{M}^q_\infty\Big(\frac{k}{n}\Big)\right|=0. \quad (2.11)$$

The arguments for the proof of (2.10) are very similar to those for the proof of Lemma 2.1, so we will focus only on (2.11) below.

As in (2.6), we can see that

$$\begin{aligned}\left|\mathbb{E}\Big(\frac{S^q_{k+1}}{n}\Big)-\mathbf{M}^q_\infty\Big(\frac{k+1}{n}\Big)\right| &\le \left|\mathbb{E}\Big(\frac{S^q_1}{n}\Big)-\mathbf{M}^q_\infty\Big(\frac{1}{n}\Big)\right|\\ &\quad+\left|\mathbf{M}^q_n\Big(\frac{k+1}{n}\Big)-\mathbf{M}^q_\infty\Big(\frac{k+1}{n}\Big)\right|\\ &\quad+\left|\mathbf{M}^q_n\Big(\frac{1}{n}\Big)-\mathbf{M}^q_\infty\Big(\frac{1}{n}\Big)\right|\\ &\quad+\sum_{i=1}^{k}\frac{1}{n}\left|\mathrm{E}\Big[f_q\Big(\frac{S_i}{n}\Big)\Big]-f_q\Big(\mathbf{M}_n\Big(\frac{i}{n}\Big)\Big)\right|,\end{aligned}$$

where the function $f_q:\mathbb{R}^p\to\mathbb{R}$ is defined by

$$f_q(x)=\frac{(1-x_q)\lambda_q}{\sum_{l=1}^p(1-x_l)\lambda_l}.$$

Here, let us note that

$$\begin{aligned}\frac{1}{n}\Big|\ f_q(y)-f_q(x)\ \Big| &\le \left|\ \frac{(1-x_q)\lambda_q\sum_{l=1}^p(y_l-x_l)\lambda_l-(y_q-x_q)\lambda_q\sum_{l=1}^p(1-x_l)\lambda_l}{\big[\sum_{l=1}^p(1-x_l)\lambda_l\big]\big[\sum_{l=1}^p(1-y_l)\lambda_l\big]}\right|\\ &\le\Big(\frac{p(p+1)\lambda_p^2}{pn-n\sum_{l=1}^p x_l}\Big)\max_{1\le l\le p}\Big|y_l-x_l\Big|, \qquad (2.12)\end{aligned}$$

where $1=\lambda_1\le\lambda_2\le\cdots\le\lambda_p$. The remaining arguments are very similar to those used to show Lemma 2 and hence are omitted. □

ACKNOWLEDGEMENTS

We are grateful to the editor and an anonymous referee for helpful comments.

REFERENCES

BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed., John Wiley & Sons, New York.

COX, D. R. (1972). "Regression models and life-tables (with discussion)", *Journal of the Royal Statistical Society*, Ser. B, **34**, 187–220.

COX, D. R. (1975). "Partial likelihood", *Biometrika*, **62**, 269–276.

KALBFLEISCH, J. D. AND PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, New York.

THERNEAU, T. M. AND GRAMBSCH, P. M. (2000). *Modeling Survival Data: Extending the Cox Model* (Statistics for Biology and Health), Springer-Verlag, New York.