The central limit theorem under censoring

MICHAEL G. AKRITAS

Department of Statistics, Penn State University, 326 Thomas Bldg., University Park PA 16802, USA.
E-mail: mga@stat.psu.edu

The central limit theorem for integrals of the Kaplan–Meier estimator is obtained. The basic tools are the martingale methods developed by Gill and the identities and inequalities of Efron and Johnstone. The assumptions needed are both weaker and more transparent than those in the recent literature, and the resulting variance expression is simpler, especially for distributions with atoms.

Keywords: distributions with atoms; i.i.d. representation; Kaplan–Meier integrals; martingales for counting processes

1. Introduction

A number of recent papers have been putting the final touches to the asymptotic theory of the Kaplan–Meier estimator (Kaplan and Meier 1958) and functionals based on it (see Wang 1987; Gijbels and Veraverbeke 1991; Einmahl and Koning 1992; Gill 1994; Stute 1995). The last paper cited establishes a central limit theorem (CLT) for a Kaplan–Meier integral by first expressing it as a sum of independent and identically distributed (i.i.d.) random variables plus an asymptotically negligible remainder term. Stute's result allows both discontinuous populations and a general class of functions, thus generalizing other CLT results (Gill 1983; Schick *et al*. 1988; Yang 1994). However, it is obtained using a delicate (and computation-intensive) approach based on U-statistic approximations. He justifies this approach by citing difficulties in the application of the counting processes techniques, and requires stronger assumptions than those used with martingale methods. In addition, the expression for the terms in his i.i.d. representation (and consequently for the asymptotic variance) is quite complicated, especially for distributions with atoms.

The main purpose of the present paper is to prove the CLT and provide an alternative i.i.d. representation with simpler terms and under weaker conditions. This is made possible by using the martingale methods developed by Gill (1980; 1983), and the identities and inequalities of Efron and Johnstone (1990). With the present approach the CLT is established directly, not as a consequence of the i.i.d. representation. These techniques require that the Kaplan–Meier integral be re-expressed as an integral in terms of the cumulative hazard function. Efron and Johnstone (1990) studied extensively the conse-quences of such re-expressions in the uncensored data context, and their results are central to understanding the relation between the expressions given here and those of Stute (1995).

It will be seen that the variance expression in the present paper is related to Efron and Johnstone's 'advance time' transformation A, while the terms in our i.i.d. representation are related to its adjoint transformation B. Surprisingly, expressions resulting from these transformations do not change much under random censoring, while the traditional

expressions which Stute (1995) worked with do. This explains why the present i.i.d. representation and formula for the asymptotic variance are simpler under censoring.

The next section sets the notation and formulates the main result. The longer proofs are given in Section 3.

2. Central limit theorem and i.i.d. representation

2.1. Notation and assumptions

Let T_i, $i = 1, \ldots, n$, be i.i.d. random variables on the real line and let F denote their common distribution function. The survival function of the T_i will be denoted by S, and the cumulative hazard function by Λ; thus $S = 1 - F$ and $\Lambda(t) = \int_{-\infty}^{t} [S(x-)]^{-1} \, dF(x)$. The observed data consist of

$$X_i = \min(T_i, C_i) \quad \text{and} \quad \Delta_i = I(X_i = T_i), \qquad i = 1, \ldots, n, \tag{1}$$

where $C_1, \ldots, C_n$ are i.i.d. random variables which are also independent of the Ts, and $I(E)$ denotes the indicator of the event E. The common distribution of the C_i will be denoted by G. The distribution function of the X_i will be denoted by H. Thus, $1 - H = (1 - F)(1 - G)$.

Let $N_i(t) = I(X_i \leq t, \Delta_i = 1)$, $N.(t) = \sum_{i=1}^{n} N_i(t)$, $Y_i(t) = I(X_i \geq t)$, $Y.(t) = \sum_{i=1}^{n} Y_i(t)$. We will assume that all random variables are defined on the probability space $(\Omega, \mathscr{F}, P)$ and we will consider the filtration

$$\mathscr{F}_t = \mathscr{N} \vee \sigma\{(X_i, \Delta_i)I(X_i \leq s), I(X_i > s) : -\infty < s \leq t, i = 1, \ldots, n\}, \tag{2}$$

where $\mathscr{N}$ consists of all P-null sets of $\mathscr{F}$. Then, $M_i(t) = N_i(t) - \int_{-\infty}^{t} Y_i(s) d\Lambda(s)$ is a martingale with respect to the filtration in (2). By analogy with the notation introduced above, $M.(t) = \sum_{i=1}^{n} M_i(t)$. The Kaplan and Meier (1958) product-limit estimator (PLE) of S based on the observations (1) will be denoted by $\hat{S}$, while $\hat{F} = 1 - \hat{S}$, $\hat{\Lambda}$ will denote the corresponding estimators of F, Λ, respectively. $A_-(s)$ or $A(s-)$ will denote the left-continuous version of a right-continuous function A, and $\Delta A(s) = A(s) - A(s-)$. Unless otherwise explicitly indicated, the domain of integration includes the upper and lower integration limits. Finally, define $\tau_n = \max(X_1, \ldots, X_n)$, and let $\tau_F = \sup\{x : F(x) < 1\}$, for any distribution function F.

Let $\phi : \mathbb{R} \to \mathbb{R}$ be any measurable function such that $\int \phi^2 \, dF < \infty$. The objective of this paper is to obtain a CLT for the Kaplan–Meier integral $\int \phi \, d\hat{F}$ and to derive an i.i.d. representation for it. The only assumption needed for the CLT and i.i.d. representation is the following:

Assumption 1. Let $\tau = \tau_H$. Then

$$\int_{-\infty}^{\tau} \frac{\phi(s)^2}{1 - G(s-)} \, dF(s) < \infty.$$

Note that assumption (1.5) of Stute (1995) is analogous to the present Assumption 1, and

apparently of comparable strength. Indeed, in the uncensored case both reduce to $\int \phi(s)^2 \, dF(s) < \infty$, while if both F and G are continuous Stute shows that his assumption (1.5) reduces to the simple and transparent form of the present Assumption 1 (see his relation (1.9)). Therefore, his assumption (1.6) is an additional assumption which is not needed in the present derivation of the CLT.

For any integrable function ϕ, we define

$$\bar{\phi}(s) \equiv \frac{1}{S(s)} \int_{(s,\tau]} \phi(t) \, dF(t). \tag{3}$$

The asymptotic variance of $\sqrt{n} \int \phi \, d\hat{F}$ will be

$$\sigma^2 = \int_{-\infty}^{\tau} \frac{S(s)}{1 - H(s-)} [\phi(s) - \bar{\phi}(s)]^2 \, dF(s), \tag{4}$$

provided that $\tau_n < \tau_F$ almost surely, or $\phi(\tau_F) = 0$.

Remark 1. The assumption that $\tau_n < \tau_F$ a.s. is also used, for example, in Shorack and Wellner (1986, p. 301), but we remark here that this entails no loss of generality. Indeed, if $\tau_n = \tau_F$, then $\int_{-\infty}^{\tau} \phi \, d(\hat{F} - F) = \int_{-\infty}^{\tau-} (\phi - \phi(\tau)) d(\hat{F} - F) = \int_{-\infty}^{\tau} (\phi - \phi(\tau)) d(\hat{F} - F)$, and the asymptotic distribution in this case follows from the first since, by redefining ϕ, the condition $\phi(\tau_F) = 0$ is satisfied.

Remark 2. Set $\psi(s) = \phi(s)(1 - G(s-))^{-1/2}$, so that, according to Assumption 1, $\int_{-\infty}^{\tau} \psi(s)^2 \, dF(s) < \infty$. Assume for the moment that $\tau = \tau_F$. Then, the expression for the variance of $\psi(T)$ obtained in Efron and Johnstone (1990) (see their relations (1.8), (3.16) and Section 4 for the continuous, discrete and general case, respectively) is

$$\int_{-\infty}^{\tau} \frac{S(s)}{S(s-)} (\psi(s) - \bar{\psi}(s))^2 \, dF(s). \tag{5}$$

Therefore the expression in (5) is finite, and since $S(s)/S(s-) \leqslant 1$, $\psi \in \mathscr{L}_2(F)$ and $\mathscr{L}_2$ spaces are linear, we obtain $\int_{-\infty}^{\tau} S(s)/S(s-) \bar{\psi}(s)^2 \, dF(s) < \infty$. Repeating the same argument for $|\psi|$, it follows that $\int_{-\infty}^{\tau} (S(s)/S(s-)) |\overline{\psi(s)}|^2 \, dF(s) < \infty$. If in this last relation we evaluate $(1 - G_-)^{-1/2}$ at the lower limit of the inside integral (the one defining $\overline{|\psi|}$ and pull it out of that integral, we obtain

$$\int_{-\infty}^{\tau} \frac{S(s)}{1 - H(s-)} \left(\frac{1}{S(s)} \int_{(s,\tau]} |\phi(x)| \, dF(x) \right)^2 \, dF(s) < \infty. \tag{6}$$

If $\tau < \tau_F$, the above arguments applied to the conditional distribution of T given that it is at most τ, imply that (6) holds with F replaced by $(F(s)/F(\tau))I(s \leqslant \tau) + I(s \geqslant \tau)$, and thus also for F. Relation (6) will be useful for bounding certain quantities in the proofs. In particular, Assumption 1 and relation (6) imply that $\sigma^2 < \infty$.

Remark 3. It is easy to see that with uncensored data the variance expression given in Corollary 1.2 of Stute (1995) reduces to $\mathrm{var}(\phi(T))$, while the variance expression in (4)

reduces to (5). Thus, (4) is related to the 'advance time' transformation A of Efron and Johnstone (1990). While the latter variance expression is practically never used with uncensored data (due to its unusual form), its advantage is that it undergoes a very minor modification under censoring. Furthermore, Corollary 1.2 of Stute (1995) shows that the traditional variance expression becomes considerably more complicated under censoring.

An alternative expression for σ^2 in (4), which is perhaps more familiar in the biostatistical literature, is

$$\sigma^2 = \int_{-\infty}^{\tau} h(s)(1 - \Delta\Lambda(s)) \mathrm{d}\Lambda(s), \tag{7}$$

where

$$h(s) \equiv h(s; \tau) = \tilde{\phi}(s)^2 \frac{I(s \leqslant \tau)}{S(s-)(1 - G(s-))}, \tag{8}$$

$$\tilde{\phi}(s) \equiv \tilde{\phi}(s; \tau) = S(s-)\left[\phi(s) - \frac{1}{S(s)} \int_{(s,\tau]} \phi(t) \, \mathrm{d}F(t)\right]. \tag{9}$$

The function $\tilde{\phi}$ will appear again later in the re-expression of the Kaplan–Meier integral as an integral in terms of the cumulative hazard function.

2.2. The main results

The CLT and i.i.d. representation will be a consequence of the following four propositions. The first establishes an asymptotic equivalence using a truncated version of the function ϕ.

Proposition 1. *Under Assumption* 1 *there exists a sequence of constants* $K_n \to \infty$ *such that the function* ϕ *truncated at* K_n, $\phi_n(s) = \phi(s)I(|\phi(s)| \leqslant K_n)$, *satisfies*

$$\int_{(-\infty,\tau]} \phi(s)\mathrm{d}(\hat{F}(s) - F(s)) = \int_{(-\infty,\tau]} \phi_n(s)\mathrm{d}(\hat{F}(s) - F(s)) + o_p(n^{-1/2}).$$

The proof is given in Section 3.

The second proposition replaces the upper limit of the integral by τ_n.

Proposition 2. *Let* $\phi_n(s)$ *be as defined in Proposition* 1. *Then, under Assumption* 1,

$$\int_{(-\infty,\tau]} \phi_n(s)\mathrm{d}(\hat{F}(s) - F(s)) = \int_{(-\infty,\tau_n]} \phi_n(s)\mathrm{d}(\hat{F}(s) - F(s)) + o_p(n^{-1/2}).$$

Proof. It will be shown that $\sqrt{n} \int_{(\tau_n, \tau]} |\phi_n(s)| dF(s) \xrightarrow{P} 0$. Indeed,

$$\int_{(\tau_n, \tau]} |\phi_n(s)| dF(s) \leqslant \int_{(\tau_n, \tau]} |\phi(s)| dF(s)$$

$$= S(\tau_n) \int_{(\tau_n, \tau]} |\phi(s)| \frac{dF(s)}{S(\tau_n)} \leqslant S(\tau_n) \left(\int_{(\tau_n, \tau]} |\phi(s)|^2 \frac{dF(s)}{S(\tau_n)} \right)^{1/2}$$

$$\leqslant (1 - H(\tau_n))^{1/2} \left(\int_{(\tau_n, \tau]} \frac{|\phi(s)|^2}{1 - G(s-)} dF(s) \right)^{1/2}.$$

Thus, by Assumption 1 and the fact that $\tau_n \xrightarrow{P} \tau$, the integral in the last relation goes to zero as $n \to \infty$ while $n(1 - H(\tau_n)) = O_p(1)$ (cf. Yang 1994). $\square$

The third proposition expresses the centred Kaplan–Meier integral, $\int_{-\infty}^{\tau} \phi_n(s) d(\hat{F}(s) - F(s))$ as an integral in terms of $\hat{\Lambda} - \Lambda$.

Proposition 3. *Suppose that either* $\tau_n < \tau_F$ *a.s., or that* $\phi(\tau_F) = 0$. *Define*

$$\tilde{\phi}_n(s) \equiv \tilde{\phi}_n(s; \tau) = S(s-) \left[\phi_n(s) - \frac{1}{S(s)} \int_{(s,\tau]} \phi_n(t) dF(t) \right].$$

Then, under Assumption 1,

$$\int_{-\infty}^{\tau_n} \phi_n(s) d(\hat{F}(s) - F(s)) = \int_{-\infty}^{\tau_n} \tilde{\phi}_n(s) d(\hat{\Lambda}(s) - \Lambda(s)) + o_p(n^{-1/2}).$$

The proof is given in Section 3.

The final proposition gives an i.i.d. representation for $\int_{-\infty}^{\tau_n} \tilde{\phi}_n(s) d(\hat{\Lambda}(s) - \Lambda(s))$.

Proposition 4. *Under Assumption* 1,

$$\int_{-\infty}^{\tau_n} \tilde{\phi}_n(s) d(\hat{\Lambda}(s) - \Lambda(s)) = \frac{1}{n} \int_{-\infty}^{\tau_n} \frac{\tilde{\phi}_n(s)}{1 - H(s-)} dM.(s) + o_p(n^{-1/2}).$$

The proof is given in Section 3.

We now present the main result.

Theorem 5. *Let Assumption* 1 *hold. Then if* $\tau_n < \tau_F$ *a.s. or* $\phi(\tau_F) = 0$ *(see Remark* 1*),*

$$n^{1/2} \int_{-\infty}^{\tau} \phi(s) d(\hat{F}(s) - F(s)) \xrightarrow{\mathscr{L}} N(0, \sigma^2),$$

with σ^2 *defined in* (4).

Proof. In view of Propositions 1–4, it suffices to show that

$$\tilde{M}.(\tau) = n^{-1/2} \int_{-\infty}^{\tau} \frac{\tilde{\phi}_n(s)I(s \leqslant \tau_n)}{1 - H(s-)} \, \mathrm{d}M.(s) = \int_{-\infty}^{\tau} \tilde{H}_1(s)\mathrm{d}M.(s) \overset{\mathscr{L}}{\to} N(0, \sigma^2),$$

where $\tilde{M}.$, $\tilde{H}_1$ are defined by the above equation. We will first show that this is true under the assumption that F is continuous. Using the CLT of Rebolledo (1980) (see also Gill 1980, p. 17), it suffices to show that, for all $t \leqslant \tau$ and $\varepsilon > 0$,

$$\langle \tilde{M}. \rangle(t) = \int_{-\infty}^{t} \tilde{H}_1(s)^2 Y.(s) \frac{\mathrm{d}F(s)}{1 - F(s-)} \overset{P}{\to} \int_{-\infty}^{t} \frac{\tilde{\phi}(s)^2}{1 - H(s-)} \mathrm{d}F(s), \tag{10}$$

$$\langle \tilde{M}.^{\epsilon} \rangle(t) = \int_{-\infty}^{t} \tilde{H}_1(s)^2 I(|\tilde{H}_1(s)| \geqslant \varepsilon) Y.(s) \frac{\mathrm{d}F(s)}{1 - F(s-)} \overset{P}{\to} 0. \tag{11}$$

Using Lemma 2.6 of Gill (1983), it follows that the integrand on the left-hand side of (10) is bounded by

$$K \frac{S(s-)}{1 - H(s-)} \left[|\phi(s)| + \frac{1}{S(s)} \int_{(s,\tau]} |\phi(x)|\mathrm{d}F(x) \right]^2,$$

for some constant K, with probability as high as desired. By Assumption 1 and relation (6) the above expression is integrable. Thus, (10) follows by an application of the dominated convergence theorem. The same argument and the fact that $I(|\tilde{H}_1(s)| \geqslant \varepsilon) \to 0$, $\forall s$, a.s.-[F], implies that (11) is true. Finally, the extension to an arbitrary F can be done via the construction presented by Gill (1980, pp. 74–78). □

Next we present an i.i.d. representation for the Kaplan–Meier integral.

Theorem 6. *Let Assumption* 1 *hold, and let* $\tilde{\phi}$ *be given by* (9). *Assume that* $\tau_n < \tau_F$ *a.s., or that* $\phi(\tau_F) = 0$ (*see Remark* 1), *and define*

$$Z_i = \frac{\tilde{\phi}(X_i)\Delta_i}{1 - H(X_i-)} - \int_{-\infty}^{X_i} \frac{\tilde{\phi}(s)}{1 - H(s-)} \, \mathrm{d}\Lambda(s). \tag{12}$$

Then, $\mathrm{E}(Z_i) = 0$, $\mathrm{var}(Z_i) = \sigma^2$, *where* σ^2 *is defined in* (4), *and*

$$n^{1/2} \int_{-\infty}^{\tau} \phi(s)\mathrm{d}(\hat{F}(s) - F(s)) = n^{-1/2} \sum_{i=1}^{n} Z_i + o_p(1). \tag{13}$$

Proof. Note that

$$Z_i = \int_{\infty}^{\tau_n} \frac{\tilde{\phi}(s)}{1 - H(s-)} \mathrm{d}M_i,$$

so, by the properties of martingales, $\mathrm{E}(Z_i) = 0$, $\mathrm{var}(Z_i) = \sigma^2$, with σ^2 defined in (4). Let now $\tilde{\phi}_n$ be as defined in Proposition 3 and write

$$\frac{1}{\sqrt{n}} \int_{-\infty}^{\tau_n} \frac{\tilde{\phi}_n(s) - \tilde{\phi}(s)}{1 - H(s-)} \, \mathrm{d}M.(s) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{n,i},$$

where, for each n, $Z_{n,1}, \ldots, Z_{n,n}$ are i.i.d. random variables and, by the properties of martingales, it can be seen that $E(Z_{n,i}) = 0$, and $\text{var}(Z_{n,i}) \to 0$, as $n \to \infty$. By the Chebyshev inequality, it follows that

$$\frac{1}{\sqrt{n}} \int_{-\infty}^{\tau_n} \frac{\tilde{\phi}_n(s) - \tilde{\phi}(s)}{1 - H(s-)} \, \mathrm{d}M.(s) = o_p(1). \tag{14}$$

In view of Propositions 1–4, the above implies the stated i.i.d. representation and thus the proof of the theorem is complete. □

Remark 4. In the uncensored case the random variable Z_i of Theorem 6 is related to the 'backward time' transformation B of Efron and Johnstone (1990). On the other hand, the terms in the i.i.d. representation of Stute (1995) reduce to $\phi(T_i)$. As in Remark 3, the present simplification is due to the fact that this transformation undergoes very minor change under censoring.

3. Proofs

The proofs that follow make repeated use of Lemma 2.6 of Gill (1983); though not stated there, we remark that this result does not require continuous distributions (see Gill 1980, Theorem 3.2.1).

Proof of Proposition 1. First, it will be shown that for any sequence K_n so that $K_n n^{-1/2} \to \infty$,

$$\int_{-\infty}^{\tau} \phi(s) I(|\phi(s)| > K_n) \, \mathrm{d}F(s) = o(n^{-1/2}). \tag{15}$$

Set $\phi(s) = 0$ for $s > \tau$, and write

$$\int_{-\infty}^{\tau} \phi(s) I(|\phi(s)| > K_n) \, \mathrm{d}F(s) = \int_{-\infty}^{\infty} x I(|x| > K_n) \mathrm{d}F_\phi(x)$$

$$= K_n(1 - F_\phi(K_n)) + \int_{K_n}^{\infty} (1 - F_\phi(x)) \, \mathrm{d}x + K_n F_\phi(-K_n)$$

$$- \int_{-\infty}^{-K_n} F_\phi(c) \, \mathrm{d}x, \tag{16}$$

where $F_\phi(x) = \int I(\phi(s) \le x) \mathrm{d}F(s)$. We will show that the first and second terms on the right-hand side of (16) are $o_p(n^{-1/2})$; similar arguments apply for the other two terms. For the first term, write (for large n) $\sqrt{n} K_n(1 - F_\phi(K_n)) \le K_n^2(1 - F_\phi(K_n)) \le \int_{K_n}^{\infty} x^2 \, \mathrm{d}F_\phi(x) \to 0$ as $n \to \infty$, where we use $K_n n^{-1/2} \to \infty$ for the first inequality and Assumption 1 for

the convergence to zero. Replacing K_n by x in the middle inequality, we obtain $x^2(1 - F_\phi(x)) \to 0$ as $x \to \infty$, implying that, for large x, $(1 - F_\phi(x)) \leqslant x^{-2}$. From this it follows that $\sqrt{n} \int_{K_n}^\infty (1 - F_\phi(x))\mathrm{d}x \leqslant \sqrt{n} \int_{K_n}^\infty x^{-2}\,\mathrm{d}x = \sqrt{n}K_n^{-1} \to 0$. Thus (16) is shown, and so, consequently, is (15). Next, it will be shown that there is a sequence $K_n \to \infty$ such that

$$\int_{-\infty}^\tau \phi(s)I(|\phi(s)| > K_n)\mathrm{d}\hat{F}(s) = o_p(n^{-1/2}). \tag{17}$$

Let $X_{(n_2)}$ denote the largest uncensored observation and write

$$\left| \sqrt{n} \int_{-\infty}^\tau \phi(s)I(|\phi(s)| > K_n)\mathrm{d}\hat{F}(s) \right| \leqslant \max_{1 \leqslant i \leqslant n} \{|\phi(X_i)\Delta_i|\} \sqrt{n} \int_{-\infty}^\tau I(|\phi(s)| > K_n)\mathrm{d}\hat{F}(s)$$

$$\leqslant \max_{1 \leqslant i \leqslant n} \{|\phi(X_i)\Delta_i|\} \sqrt{n} \sum_{i=1}^n I(|\phi(X_i)\Delta_i| > K_n)\Delta\hat{F}(X_{(n_2)})$$

$$= \max_{1 \leqslant i \leqslant n} \{|\phi(X_i)\Delta_i|\} \sqrt{n} \sum_{i=1}^n I(|\phi(X_i)\Delta_i| > K_n)(1 - \hat{F}(X_{(n_2)}-)) \frac{\Delta N.(X_{(n_2)})}{Y.(X_{(n_2)})} \tag{18}$$

$$\leqslant B \max_{1 \leqslant i \leqslant n} \{|\phi(X_i)\Delta_i|\} \sqrt{n} \sum_{i=1}^n I(|\phi(X_i)\Delta_i| > K_n) \frac{1 - F(X_{(n_2)}-)}{1 - H(X_{(n_2)}-)} \frac{\Delta N.(X_{(n_2)})}{n},$$

where, by Lemmas 2.6 and 2.7 of Gill (1983) the positive constant B can be chosen large enough for the last inequality in (18) to hold with probability as high as desired. Using Markov's inequality and Assumption 1, it can be seen that, for any sequence $K_n \to \infty$,

$$R_{1,n} = \frac{K_n^2}{n} \sum_{i=1}^n I(|\phi(X_i)\Delta_i| > K_n) = o_p(1), \tag{19}$$

$$R_{2,n} = n^{-1/2} \max_{1 \leqslant i \leqslant n} \{|\phi(X_i)\Delta_i|\} = o_p(1). \tag{20}$$

By (18), (19) and (20) we can write

$$\sqrt{n} \int_{-\infty}^\tau \phi(s)I(|\phi(s)| > K_n)\mathrm{d}\hat{F}(s) \leqslant BR_{1,n}R_{2,n} \frac{n^2}{K_n^2} \frac{1 - F(X_{(n_2)}-)}{1 - H(X_{(n_2)}-)} \frac{\Delta N.(X_{(n_2)})}{n}, \tag{21}$$

with the inequality holding for n large enough and with probability as high as desired by choosing B large enough. This relation implies that if τ is an atom for H, (17) holds by choosing $K_n n^{-1} \to \infty$. If τ is not an atom for H, we need some further arguments. Set $\omega(x) = (1 - F(x))^{-1}$ and pick K_n so that $n/\omega(m_n)$ remains bounded as $n \to \infty$, where $m_n^{-1} = n^2/K_n^2$. It can be seen that, for any $t > 0$,

$$P\left(m_n^{-1} \frac{1}{1 - G(X_{(n_2)}-)} > t \right) = 1 - \left(P\left(\frac{\Delta_i}{1 - G(X_i-)} \leqslant m_n t \right) \right)^n$$

$$= 1 - \left(1 - P\left(\frac{\Delta_i}{1 - G(X_i-)} > m_n t \right) \right)^n = 1 - \left(1 - \frac{1}{\omega(m_n t)} \right)^n,$$

which, by the choice of K_n (and hence of m_n) converges to an arbitrarily small positive constant by choosing t large enough. Thus $m_n^{-1}(1/(1 - G(X_{(n_2)}-)))$ remains bounded in probability. This shows that the expression in (21) is $o_p(1)$ which implies that relation (17), and hence the proposition, is shown. □

Proof of Proposition 3. First, using the identity

$$\hat{F}(t) - F(t) = S(t) \int_{-\infty}^{t} \hat{S}(x-)[S(x)]^{-1} \mathrm{d}(\hat{\Lambda}(x) - \Lambda(x))$$

(Gill 1980) and a straightforward calculation, it follows that

$$\int_{-\infty}^{\tau_n} \phi_n(s) \mathrm{d}(\hat{F}(s) - F(s)) = \int_{-\infty}^{\tau_n} \tilde{\phi}_n^*(s; \tau_n) \mathrm{d}(\hat{\Lambda}(s) - \Lambda(s)), \tag{22}$$

where $\tilde{\phi}_n^*(s; t) = \hat{S}(s-)[\phi_n(s) - [S(s)]^{-1} \int_{(s,t]} \phi_n(x) \, \mathrm{d}F(x)]$, for all t. Next, it will be shown that

$$\int_{-\infty}^{\tau_n} (\tilde{\phi}_n^*(s; \tau_n) - \tilde{\phi}_n^*(s; \tau)) \mathrm{d}(\hat{\Lambda}(s) - \Lambda(s)) = o_p(n^{-1/2}). \tag{23}$$

Clearly, if τ is a point of discontinuity of H, then $\tau_n = \tau$ almost surely for some n large enough, and so (23) holds. Suppose that τ is a continuity point of H. Showing (23) reduces to showing

$$\sqrt{n} \int_{(\tau_n, \tau]} \phi_n(t) \, \mathrm{d}F(t) \frac{\hat{F}(\tau_n) - F(\tau_n)}{S(\tau_n)} \xrightarrow{P} 0. \tag{24}$$

If $F(\tau) < 1$, (24) follows by direct calculations. Consider now the case where τ is a point of continuity of F and $F(\tau) = 1$. By Remark 2.2 in Gill (1983), the expression in (24) is zero at $\tau_n = \tau$. Using Theorem 4.2 in Billingsley (1968), to show (24) it suffices to show that the process

$$\psi_n(t) Z_{1n}(t) := \int_{(t,\tau]} \phi_n(s) \, \mathrm{d}F(s) \sqrt{n} \frac{\hat{F}(t) - F(t)}{S(t)} \tag{25}$$

is 'tight at τ', that is to say, that

$$\lim_{s \uparrow \tau} \limsup_{n \to \infty} P\left(\sup_{s \leqslant t \leqslant \tau_n} |\psi_n(t) Z_{1n}(t) - \psi_n(s) Z_{1n}(s)| > \varepsilon \right) = 0, \tag{26}$$

for all $\varepsilon > 0$. We will follow the arguments used by Gill (1983) to prove his relation (2.7). Note, however, that his function h is now replaced by ψ_n which depends on n and, moreover, is only right-continuous (as opposed to the continuous h function), and is not necessarily non-increasing. Therefore we briefly repeat the arguments he uses to prove his relation (2.7). Write

$$\sup_{s \leqslant t \leqslant \tau_n} |\psi_n(t) Z_{1n}(t) - \psi_n(s) Z_{1n}(s)| \leqslant \sup_{s \leqslant t \leqslant \tau_n} |\psi_n(t)(Z_{1n}(t) - Z_{1n}(s))| + |(\psi_n(t) - \psi_n(s)) Z_{1n}(s)|. \tag{27}$$

By the weak convergence of $Z_{1n}(s)$ to a Gaussian process with variance function

$$C(t) = \int_0^t \frac{\mathrm{d}\Lambda(s)}{1 - H(s-)},$$

it follows that

$$\limsup_{n\to\infty} P(|(\psi_n(t) - \psi_n(s))Z_{1n}(s)| > \varepsilon) \leq \limsup_{n\to\infty} P\left(\left|\int_{(s,t]} |\phi(x)|\mathrm{d}F(x)Z_{1n}(s)\right| > \varepsilon\right)$$

$$\leq \frac{\left(\int_{(s,t]} |\phi(x)|\mathrm{d}F(x)\right)^2 C(s)}{\varepsilon^2}.$$

The right-hand side of the above expression is seen to converge to zero, as $s \to \tau$, with $s \leq t < \tau$, from Assumption 1 and the following set of inequalities:

$$\left(\int_{(s,t]} |\phi(x)|\mathrm{d}F(x)\right)^2 C(s) \leq \left[S(s)\int_s^\tau |\phi(x)| \frac{\mathrm{d}F(x)}{S(s)}\right]^2 \int_{-\infty}^s \frac{\mathrm{d}F(x)}{S(x-)^2(1 - G(x-)}$$

$$\leq S(s)^2 \int_s^\tau \phi(x)^2 \frac{\mathrm{d}F(x)}{S(s)} \int_{-\infty}^s \frac{\mathrm{d}F(x)}{S(x-)^2(1 - G(x-))}$$

$$\leq S(s)\int_s^\tau \phi(x)^2 \,\mathrm{d}F(x) \frac{1}{1 - G(s-)} \int_{-\infty}^s \frac{S(x)}{S(x-)} \frac{-\mathrm{d}S(x)}{S(x-)S(x)}$$

$$\leq S(s)\int_s^\tau \frac{\phi(x)^2}{1 - G(x-)} \,\mathrm{d}F(x) \int_{-\infty}^s \frac{-\mathrm{d}S(x)}{S(x-)S(x)}$$

$$= S(s)\int_s^\tau \frac{\phi(x)^2}{1 - G(x-)} \,\mathrm{d}F(x) \int_{-\infty}^s \mathrm{d}\left(\frac{1}{S(x)}\right).$$

The first inequality is due to replacing t by τ, the second follows by the Cauchy–Schwarz inequality, the third and fourth by the monotonicity of distribution functions. For the first term on the right-hand side of (27), write

$$\sup_{s\leq t\leq \tau_n} |\psi_n(t)(Z_{1n}(t) - Z_{1n}(s))| \leq \sup_{s\leq t\leq \tau_n}\left|\int_{(t,\tau]} |\phi(x)|\mathrm{d}F(x)(Z_{1n}(t) - Z_{1n}(s))\right|.$$

Next, note that the inequality in Lemma 2.9 of Gill (1983) holds also for right-continuous functions, and thus (26) will follow from

$$\lim_{s\uparrow\tau}\limsup_{n\to\infty} P\left(\sup_{s\leq t\leq \tau_n}\left|\int_s^t \int_{(x,\tau]} |\phi(y)|\mathrm{d}F(y)\mathrm{d}Z_{1n}(x)\right| > \varepsilon\right) = 0.$$

This follows by the use of the inequality of Lenglart (1977) as in the proof of Theorem 2.1 of Gill (1983) by noting that the first of the series of two inequalities is valid for non-continuous

distributions by the fact that $0 \leqslant 1 - \Delta\Lambda \leqslant 1$. This completes the proof of (24) and thus of (23).

Finally, it will be shown that

$$\int_{-\infty}^{\tau_n} (\tilde{\phi}_n^*(s; \tau) - \tilde{\phi}_n(s; \tau)) \mathrm{d}(\hat{\Lambda}(s) - \Lambda(s)) = o_p(n^{-1/2}). \tag{28}$$

This can be rewritten as

$$\int_{-\infty}^{\tau} \tilde{H}_2(s) \mathrm{d}M.(s) \xrightarrow{P} 0, \tag{29}$$

where

$$\tilde{H}_2(s) = \frac{\sqrt{n}(\tilde{\phi}_n^*(s; \tau) - \tilde{\phi}_n(s; \tau))}{Y.(s)} I(s \leqslant \tau_n).$$

Using the CLT of Rebolledo (1980) (see also Gill 1980, p. 17), it suffices to show that

$$\int_{-\infty}^{\tau} \tilde{H}_2(s)^2 (1 - \Delta\Lambda(s)) Y.(s) \frac{\mathrm{d}F(s)}{1 - F(s-)} \xrightarrow{P} 0. \tag{30}$$

The integrand on the left-hand side of (30) is bounded by

$$\frac{(\hat{S}(s-) + S(s-))|\hat{S}(s-) - S(s-)|(\phi_n(s) - \bar{\phi}_n(s))^2}{n^{-1} Y.(s)} \frac{1}{1 - F(s-)}$$

$$\leqslant K \frac{|\hat{S}(s-) - S(s-)|(|\phi(s)| + |\overline{\phi(s)}|)^2}{1 - H(s-)} \tag{31}$$

$$\leqslant K \frac{S(s-)(|\phi(s)| + |\overline{\phi(s)}|)^2}{1 - H(s-)}, \tag{32}$$

where the inequalities in (31), (32) hold (with appropriate choices of K) with probability as high as desired by Lemmas 2.6 and 2.7 of Gill (1983). By Assumption 1 and relation (6), (32) is F-integrable. Also, (31) and the uniform consistency result of Wang (1987) imply that the dominated convergence theorem applies on a set of probability as high as desired. This implies (30) and thus (28) is shown. The proof of the proposition is a result of (22), (23) and (28). □

Proof of Proposition 4. First, note that

$$\sqrt{n} \int_{-\infty}^{\tau_n} \tilde{\phi}_n(s) \mathrm{d}(\hat{\Lambda}(s) - \Lambda(s)) = \frac{1}{\sqrt{n}} \int_{-\infty}^{\tau_n} \frac{\tilde{\phi}_n(s)}{n^{-1} Y.(s)} \mathrm{d}M.(s).$$

Thus it suffices to show that

$$\int_{-\infty}^{\tau} \tilde{H}_3(s) \mathrm{d}M.(s) \xrightarrow{P} 0, \tag{33}$$

where

$$\tilde{H}_3(s) = \frac{1}{\sqrt{n}} \left(\frac{\tilde{\phi}_n(s)}{n^{-1} Y.(s)} - \frac{\tilde{\phi}_n(s)}{1 - H(s-)} \right) I(s \leq \tau_n).$$

This, however, can be shown by the same arguments used for showing (29). □

Acknowledgement

References

Billingsley, P. (1968) *Convergence of Probability Measures*. New York: Wiley.

Efron, B. and Johnstone, I.M. (1990) Fisher's information in terms of the hazard rate. *Ann. Statist.*, **18**, 38–62.

Einmahl, H.J. and Koning, A.J. (1992) Limit theorems for a general weighted process under random censoring. *Canad. J. Statist.*, **20**, 77–89.

Gijbels, I. and Veraverbeke, N. (1991) Almost sure asymptotic representation for a class of functionals of the Kaplan–Meier estimator. *Ann. Statist.*, **19**, 1457–1470.

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

Gill, R.D. (1983) Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.*, **11**, 49–58.

Gill, R.D. (1994) Glivenko–Cantelli for Kaplan–Meier. *Math. Methods Statist.*, **3**, 76–87.

Kaplan, E.L. and Meier, P. (1958) Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.*, **53**, 457–481.

Lenglart, E. (1977) Relation de domination entre deux processus. *Ann. Inst. H. Poincaré (n. Sér.) Sect. B*, **13**, 171–179.

Rebolledo, R. (1980) Central limit theorems for local martingales. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **51**, 269–286.

Schick, A., Susarla, V. and Koul, H. (1988) Efficient estimation with censored data. *Statist. Decisions*, **6**, 349–360.

Shorack, G. and Wellner, J. (1986) *Empirical Processes*. New York: Wiley.

Stute, W. (1995) The central limit theorem under random censorship. *Ann. Statist.*, **23**, 422–439.

Wang, J.-G. (1987) A note on the uniform consistency of the Kaplan–Meier estimator. *Ann. Statist.*, **15**, 1313–1316.

Yang, S. (1994) A Central limit theorem for functionals of the Kaplan–Meier estimator. *Statist. Probab. Lett.*, **21**, 337–345.