THE UNIFORM CENTRAL LIMIT THEOREM FOR THE KAPLAN-MEIER INTEGRAL PROCESS

JONGSIG BAE AND SUNGYEUN KIM

Let $U_n(f) = \sqrt{n} \int f d(\widehat{F}_n - \widetilde{F})$ be the Kaplan-Meier integral process constructed from a random censorship model. We prove a uniform central limit theorem for $\{U_n\}$ under the bracketing entropy condition and mild conditions due to the censoring effects. We also prove a sequential version of the uniform central limit theorem that will give a functional law of the iterated logarithm of Strassen type.

1. INTRODUCTION

In this paper, we develop the uniform central limit theorem for the function indexed Kaplan–Meier integral process based on the incomplete data of the random censorship model. The main goal is to investigate a tightness for the process under the metric entropy with $\mathcal{L}^2$ bracketing condition for the indexed class of functions of the process and mild assumptions due to censoring effects as in Stute [7]. In the achieving the goal, because of the lack of exponential inequalities for the random censorship model, we do not attempt to use the usual chaining argument that depends on the use of exponential inequalities on the tail probabilities. Instead we investigate the uniform order of convergence of the remainder terms in the representation of Stute [7] on the Kaplan-Meier integral and use the result of Ossiander [5] on the complete data of independent and identically distributed random variables.

The uniform central limit theorem of the present paper extends the one dimensional central limit theorem under random censoring that was established by Stute in 1995 [7] and the Ossiander's uniform central limit theorem for independent and identically distributed random variables that appeared in 1987 ([5]). Among others a sequential integral process and an invariance principle of the Kaplan-Meier integral will be produced as a corollary of the main result. The results may be used in nonparametric statistical inference. We begin with introducing the integral version of the usual empirical process based on the complete data of independent and identically distributed random variables.

Let X be a random variable defined on a probability space $(\Omega, \mathcal{T}, P)$ whose distribution function is F. Consider a sequence $\{X_i : i \geqslant 1\}$ of independent copies of X. Given a Borel measurable function $f : \mathbf{R} \to \mathbf{R}$, we see that $\{f(X_i) : i \geqslant 1\}$ forms of a sequence of independent and identically distributed random variables that are more flexible in applications than the sequence $\{X_i : i \geqslant 1\}$. Consider a class $\mathcal{F}$ of real-valued Borel measurable functions defined on $\mathbf{R}$. Introduce the usual empirical distribution function F_n defined by $F_n(x) = n^{-1} \sum_{i=1}^{n} \{X_i \leqslant x\}$ for $x \in \mathbf{R}$. Define a function indexed integral process S_n by

$$(1.1) \qquad S_n(f) = n^{1/2} \int f(x) d(F_n - F)(x) \ \text{ for } f \in \mathcal{F}.$$

Throughout the paper events are identified with their indicator functions when there is no risk of ambiguity. So, for example, the summand of the empirical distribution function means the indicator functions of the events $\{X_i \leqslant x\}$.

Developing a uniform central limit theorem for the processes such as S_n usually means that $\mathcal{L}(S_n(f) : f \in \mathcal{F}) \to \mathcal{L}(W(f) : f \in \mathcal{F})$, where the processes are indexed by $\mathcal{F}$ and are considered as random elements in $B(\mathcal{F})$, the space of the bounded real-valued functions on $\mathcal{F}$, taken with the sup norm $\|\cdot\|_{\mathcal{F}}$. It is known that $(B(\mathcal{F}), \|\cdot\|_{\mathcal{F}})$ forms a Banach space. The process $(W(f) : f \in \mathcal{F})$ will be Gaussian which is uniformly continuous in f with respect to a metric. The metric we shall use is the $\mathcal{L}^2$ metric defined by $d(f, g) = \left[\int (f - g)^2 dF \right]^{1/2}$, for $f, g \in \mathcal{F}$, where F is the underlying distribution associated with X.

In order to measure the size of the function space, we define the following version of metric entropy with bracketing. See, for example, Van der Vaart and Wellner [8] for the recent reference.

DEFINITION 1: For a metric space $(\mathcal{F}, d)$ and $\delta > 0$ we define the covering number with bracketing $N_{[]}(\delta, \mathcal{F}, d)$, or $N_{[]}(\delta)$, as the smallest n for which there exists $\{f_{0,\delta}^l, f_{0,\delta}^u, \ldots, f_{n,\delta}^l, f_{n,\delta}^u\}$ so that for every $f \in \mathcal{F}$ there exist some $0 \leqslant i \leqslant n$ satisfying $f_{i,\delta}^l \leqslant f \leqslant f_{i,\delta}^u$ and $d\left(f_{i,\delta}^l, f_{i,\delta}^u\right) < \delta$. Define the metric entropy with bracketing to be $\ln N_{[]}(\delta, \mathcal{F}, d)$. We also define the associated integral of the metric entropy with bracketing to be

$$J(\delta) := \int_0^{\delta} \left[\ln N_{[]}(u, \mathcal{F}, d) \right]^{1/2} du \ \text{ for } 0 < \delta \leqslant 1.$$

We use the following definition of weak convergence which is originally due to Hoffman-Jørgensen [2].

DEFINITION 2: A sequence of $B(\mathcal{F})$-valued random functions $\{Y_n\}$ converges in law to a $B(\mathcal{F})$-valued random function Y whose law concentrates on a separable subset of $B(\mathcal{F})$ if

$$Eg(Y) = \lim_{n \to \infty} E^* g(Y_n) \ \forall \ g \in U\big(B(\mathcal{F}), \|\cdot\|_{\mathcal{F}}\big),$$

where $U\big(B(\mathcal{F}), \|\cdot\|_{\mathcal{F}}\big)$ is the set of all bounded, uniformly continuous functions from $\big(B(\mathcal{F}), \|\cdot\|_{\mathcal{F}}\big)$ into $\mathbf{R}$. Here E^* denotes the upper expectation with respect to the outer probability P^*. We denote this convergence by $Y_n \Rightarrow Y$.

In 1987, using a delicate chaining argument with stratifications, Ossiander [5] developed the uniform central limit theorem for the sequence of independent and identically distributed random variables for which $J(1) < \infty$. The Ossiander result states that if $J(1) < \infty$, then $S_n \Rightarrow W$ as random elements of $B(\mathcal{F})$ where $\big\{W(f) : f \in \mathcal{F}\big\}$ is a mean zero Gaussian process with the covariance structure $EW(f)W(g) = Ef(X)g(X)$.

In 1995, from a statistical point of view, Stute [7] developed a one dimensional central limit theorem for the Kaplan–Meier integral based on the incomplete data of the random censorship model.

The aim of our work is to extend Stute's one dimensional central limit theorem to a function indexed process version, as was done in Ossiander's setting.

In Section 2, the random censorship model and notations are introduced and the main results are stated. Proofs of the results including the uniform order of convergence of remainder terms are contained in Section 3. Finally, in Section 4, we consider the sequential Kaplan–Meier integral process as an application of the main results.

2. THE MAIN RESULTS

We consider the random censorship model where one observes the incomplete data $\{Z_i, \delta_i\}$. The $\{Z_i\}$ are independent copies of Z whose distribution is H. The $\{Z_i, \delta_i\}$ are obtained by the equations $Z_i = \min(X_i, Y_i)$ and $\delta_i = \{X_i \leqslant Y_i\}$ where the $\{Y_i\}$ are independent copies of the censoring random variable Y with distribution G which is also assumed to be independent of F, the distribution of independent and identically distributed random variables $\{X_i\}$ of original interest in a statistical inference. Let $F\{a\} = F(a) - F(a-)$ denote the jump size of F at a and let A be the set of all atoms of H which is an empty set when H is continuous. Let $\tau_H = \inf\{x : H(x) = 1\}$ denote the least upper bound of the support of H. The fact that the τ_H is not necessarily finite provides one of the reasons why we need the theory of weak convergence of the *infinite time scale* stochastic processes such as the invariance principle in the corollary below.

Consider a subdistribution function $\widetilde{F}$ defined by

$$\widetilde{F}(x) = F(x)\{x < \tau_H\} + \big[F(\tau_H -) + \{\tau_H \in A\}\big]F\{\tau_H\}\{x \geqslant \tau_H\}.$$

Again let $\mathcal{F}$ be a class of functions which are real-valued measurable defined on $\mathbf{R}$. We consider the Kaplan–Meier integral process $\{U_n\}$ defined by

$$(2.2) \qquad U_n(f) = n^{1/2} \int f d(\widehat{F}_n - \widetilde{F}) \ \text{ for } f \in \mathcal{F},$$

where $\widehat{F}_n$ is the usual Kaplan–Meier estimator of the underlying distribution F in the random censorship model. See Kaplan and Meier [3]. The process $\{U_n(f) : f \in \mathcal{F}\}$ will be the proper extension of the process $\{S_n(f) : f \in \mathcal{F}\}$ given in (1.1) to the random censorship model. As mentioned before our goal is to consider the weak convergence of the Kaplan–Meier integral process U_n to a Gaussian process as random elements of $B(\mathcal{F})$ under $J(1) < \infty$ and the minimal assumptions due to censoring effect. In order to describe the last assumptions, we need to consider the following subdistribution functions $\widetilde{H}^0(z) = P(Z \leqslant z, \delta = 0)$ and $\widetilde{H}^1(z) = P(Z \leqslant z, \delta = 1)$ for $z \in \mathbf{R}$. Define

$$\gamma(x) = \exp\left\{ \int_{-\infty}^{x-} \frac{\widetilde{H}^0(dz)}{1 - H(x)} \right\}, \ \text{ and } \ C(x) = \int_{-\infty}^{x-} \frac{G(dy)}{[1 - H(y)][1 - G(y)]}.$$

The following two assumptions, which trivially reduce to square integrability of functions when there is no censoring, will be imposed on the main results of the paper.

$$(2.3) \qquad \int f^2(x)\gamma^2(x)\widetilde{H}^1(dx) = \int \big[f(Z)\gamma(Z)\delta \big]^2 dP < \infty \ \text{ for each } f \in \mathcal{F}$$

and

$$(2.4) \qquad \int |f(x)| C^{1/2}(x)\widetilde{F}(dx) < \infty \ \text{ for each } f \in \mathcal{F}.$$

The assumptions (2.3) and (2.4) will be called "the minimal pointwise property" for $\mathcal{F}$. Before stating the results, we need more notation:

$$\gamma_1^f(x) = \frac{1}{1 - H(x)} \int \{x < w\} f(w)\gamma(w)\widetilde{H}^1(dw),$$

$$\gamma_2^f(x) = \iint \{v < x, v < w\} \frac{f(w)\gamma(w)}{[1 - H(v)]^2} \widetilde{H}^0(dv)\widetilde{H}^1(dw).$$

Write, for each $f \in \mathcal{F}$,

$$(2.5) \qquad \xi(f) = f(Z)\gamma(Z)\delta - \int f d\widetilde{F} + \gamma_1^f(Z)(1 - \delta) - \gamma_2^r(Z).$$

Let $\{W(f) : f \in \mathcal{F}\}$ be the mean zero Gaussian process with

$$(2.6) \qquad Cov\big(W(f), W(g)\big) = Cov\big(\xi(f), \xi(g)\big).$$

We are ready to state the uniform central limit theorem for the Kaplan–Meier integral process.

THEOREM 2.1. *Suppose $J(1) < \infty$. Assume $\mathcal{F}$ has the minimal pointwise property. Then $U_n \Rightarrow W$ as elements of $B(\mathcal{F})$ where $\{W(f) : f \in \mathcal{F}\}$ is a Gaussian process with the mean $EW(f) = 0$ and the covariance function is given by (2.6). The limiting process W has continuous sample paths with repsect to the $\mathcal{L}^2$ metric d.*

THEOREM 2.2. *Assume $\mathcal{F}$ has the minimal pointwise property. Then the finite dimensional distributions of U_n converge to those of W.*

THEOREM 2.3. *Suppose $J(1) < \infty$. Assume $\mathcal{F}$ has the minimal pointwise property. Then for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\limsup_{n \to \infty} P^* \left\{ \sup_{d(f,g)<\delta} |U_n(f) - U_n(g)| > \varepsilon \right\} < \varepsilon.$$

3. PROOF OF THE RESULTS

We begin with stating a decomposition of the integral $\int f d\widehat{F}_n$ into a sum of independent and identically distributed random variables that can be controlled by the result of Ossiander [5] and remainder terms that can be shown *uniformly* negligible by extensive calculation. The proof of the following Proposition 1 appears in Stute [7].

PROPOSITION 1. *Assume $\mathcal{F}$ has the minimal pointwise property. Then for each fixed $f \in \mathcal{F}$, we have*

$$(3.7) \qquad n^{1/2} \int f d(\widehat{F}_n - \widetilde{F}) = n^{-1/2} \sum_{i=1}^{n} \xi_i(f) + n^{1/2} R_n(f),$$

where $n^{1/2}|R_n(f)| \to^P 0$ and $\xi_i(f)$ are independent and identically distributed copies of the random variables $\xi(f)$ given by (2.5).

The following Proposition 2 will also be used to verify the finite dimensional distribution convergence of the process $\{U_n\}$.

PROPOSITION 2. *Assume $\mathcal{F}$ has the minimal pointwise property. Then, for each fixed $f \in \mathcal{F}$, $n^{1/2} \int f d(\widehat{F}_n - \widetilde{F}) \to N\left(0, var(\xi(f))\right)$ in distribution.*

For the family $\mathcal{F}$ we define an envelope by $\Phi(\cdot) := \sup_{f \in \mathcal{F}} |f(\cdot)|$. We need the following regularity results on the envelope Φ.

LEMMA 1. *Suppose $J(1) < \infty$. Then $\int \Phi^2(x)dF(x) < \infty$. Assume further that $\mathcal{F}$ has the minimal pointwise property. Then $\int [\Phi(Z)\gamma(Z)\delta]^2 dP < \infty$ and $\int \Phi(x)C^{1/2}(x)\widetilde{F}(dx) < \infty$.*

PROOF OF LEMMA 1: Since $N_{[]}(1) < \infty$, square integrability of Φ follows from the observation $\Phi(\cdot) \leqslant \sum_{i=0}^{N_{[]}(1)} \left(|f_{i,1}^l(\cdot)| + |f_{i,1}^u(\cdot)| \right)$. The other two results easily follow

from the assumptions of the minimal pointwise property of $\mathcal{F}$. The proof of Lemma 1 is completed.   □

Next for completeness we introduce the following temporary assumption on the envelope $\Phi(\cdot)$ that will be imposed on lemmas below and finally be removed using Lemma 1 and the assumptions of Theorem 1.

$$(3.8) \qquad \Phi(x) = 0 \text{ for all } x > T, \text{ for some } T < \tau_H.$$

The assumption (3.8) will be called the "tail property" for Φ. Write $V_n(f) = n^{-1/2} \sum_{i=1}^{n} \xi_i(f)$ and write $\mathcal{G} := \{\xi(f) : f \in \mathcal{F}\}$.

LEMMA 2. *If $J(1) < \infty$, if $\mathcal{F}$ has the minimal pointwise property, and if Φ has the tail property then the metric entropy integrability condition*

$$\int_0^1 \left[\ln N_{[]}(\varepsilon, \mathcal{G}, d)\right]^{1/2} d\varepsilon < \infty$$

holds so that, in particular Ossiander [5, Theorem 3.1 and Theorem 3.3] can be applied to the process $\{V_n(f) : f \in \mathcal{F}\}$.

PROOF OF LEMMA 2: Let $\varepsilon > 0$. By definition of $N_{[]}(\varepsilon)$, there exists $\left\{[f_{0,\varepsilon}^l, f_{0,\varepsilon}^u],\right.$ $\left.\ldots, [f_{N_{[]}(\varepsilon),\varepsilon}^l f_{N_{[]}(\varepsilon),\varepsilon}^u]\right\}$ so that for every $f \in \mathcal{F}$ there exists $0 \leqslant i \leqslant N_{[]}(\varepsilon)$ satisfying $f_{i,\varepsilon}^l \leqslant f \leqslant f_{i,\varepsilon}^u$ and $d(f_{i,\varepsilon}^l, f_{i,\varepsilon}^u) < \varepsilon$. Let $g \in \mathcal{G}$. Then $g = \xi(f)$ for some $f \in \mathcal{F}$. Now define the brackets for the class $\mathcal{G}$ by the equations

$$g_{j,\varepsilon}^l := f_{j,\varepsilon}^l(Z)\gamma(Z)\delta - \int f_{j,\varepsilon}^u d\widetilde{F} + \gamma_1^{f_{j,\varepsilon}^l}(Z)(1-\delta) - \gamma_2^{f_{j,\varepsilon}^u}(Z)$$

and

$$g_{j,\varepsilon}^u := f_{j,\varepsilon}^u(Z)\gamma(Z)\delta - \int f_{j,\varepsilon}^l d\widetilde{F} + \gamma_1^{f_{j,\varepsilon}^u}(Z)(1-\delta) - \gamma_2^{f_{j,\varepsilon}^l}(Z),$$

for $j = 0, \ldots, N_{[]}(\varepsilon)$. Simplify the notations by writing $f^l = f_{j,\varepsilon}^l$, $f^u = f_{j,\varepsilon}^u$, $g^l = g_{j,\varepsilon}^l$ and $g^u = g_{j,\varepsilon}^u$. Clearly, we have $g^l \leqslant g \leqslant g^u$.

CLAIM 1. There exists a constant C satisfyng $d^2(g^l, g^u) \leqslant C \cdot d^2(f^l, f^u)$.

From the constructions of g^u and g^l, using Jensen inequality, we see that $d^2(g^l, g^u) = \int (g^u - g^l)^2 dP$ is bounded by $4 \cdot (I_1 + I_2 + I_3 + I_4)$, where

$$I_1 = \int (f^u - f^l)^2(Z)\gamma^2(Z)\delta dP, \qquad I_2 = \iint (f^u - f^l)^2 d\widetilde{F} dP,$$

$$I_3 = \int \left(\gamma_1^{f^u} - \gamma_1^{f^l}\right)^2(Z)(1-\delta)dP, \quad I_4 = \int \left(\gamma_2^{f^u} - \gamma_2^{f^l}\right)^2(Z)dP.$$

A straightforward calculation shows $I_1 \leqslant d^2(f^l, f^u)$ and $I_2 \leqslant d^2(f^l, f^u)$. Let T be such that $\Phi(x) = 0$ for all $x > T$, $T < \tau_H$. In order to obtain a bound of I_3, we put $M := \sup\limits_{w \leqslant T} \gamma(w)$. Then we see that

$$I_3 \leqslant M^2 \int \left[\int \frac{\{Y < w\}}{(1 - G(Y))^2} dP \right] (f^l - f^u)^2(w)(1 - G(w-)) F(dw).$$

Now, change the variable, $y = G(v)$, to see that $\int \left(\{Y < w\}/(1 - G(Y))^2 \right) dP \leqslant 1/(1 - G(w-))$. Therefore, the last equation is bounded by $M^2 d^2(f^l, f^u)$. Similarly, letting $M := \sup\limits_{w \leqslant T} \gamma(w) \sup\limits_{w \leqslant T} \sup\limits_{v \in (-\infty, w)} (1 - H(v))^{-2}$, we conclude that $I_4 \leqslant M^2 d^2(f^l, f^u)$. The proof of Claim 1 is completed. Finally the result follows from assumption $J(1) < \infty$. $\square$

PROOF OF THEOREM 1: Observe that condition $J(1) < \infty$ implies the total boundedness of the metric space $(\mathcal{F}, d)$. Using the finite dimensional distribution convergence result of Theorem 2, a tightness result of Theorem 3, we apply [**6**, Theorem 10.2] to complete the proof of Theorem 1. $\square$

PROOF OF THEOREM 2: Recall thet $U_n(f) = V_n(f) + n^{1/2} R_n(f)$. With the assumptions we can apply Proposition 2 for each fixed $f \in \mathcal{F}$ and Slutsky Theorem to get the one dimensional central limit theorem for U_n. Then use the usual Cramer Wold device to get the finite dimensional distribution convergence of U_n. $\square$

On the class of functions $\mathcal{F}$ we shall work with the $\mathcal{L}^2$-metric d to prove a tightness of the process U_n. We need the following Lemma whose proof will be given in the next section. For a function $\psi : \mathcal{F} \to \mathbf{R}$, we let $\|\psi\|_{\mathcal{F}} = \sup\limits_{f \in \mathcal{F}} |\psi(f)|$ denote the sup of $|\psi|$ over $\mathcal{F}$. Also let $\|\psi\|_{\delta} = \sup\limits_{d(f,g) < \delta} |\psi(f) - \psi(g)|$ denote the sup of $|\psi(f) - \psi(g)|$ over the pairs (f, g) in the set of diameter δ about the diagonal of the space $\mathcal{F} \otimes \mathcal{F}$.

LEMMA 3. *If $J(1) < \infty$, if $\mathcal{F}$ has the minimal pointwise property, and if Φ has the tail property then $\{\sqrt{n} R_n(f) : f \in \mathcal{F}\}$ is tight. That is, for every $\varepsilon > 0$ we have*

$$\limsup_{n \to \infty} P^* \{ n^{1/2} \|R_n\|_{\mathcal{F}} > \varepsilon \} = 0.$$

PROOF OF THEOREM 3: Notice that with probability 1,

$$|U_n(f) - U_n(g)| \leqslant |V_n(f) - V_n(g)| + 2n^{1/2} \|R_n\|_{\mathcal{F}}.$$

Now from [**5**, Theorem 3.3] and Lemma 3 we conclude that

$$P^* \{ \|U_n\|_{\delta} > 3\varepsilon \| \leqslant P^* \{ \|V_n\|_{\delta} > \varepsilon \| + P^* \{ 2n^{1/2} \|R_n\|_{\mathcal{F}} > 2\varepsilon \} < 3\varepsilon$$

eventually. Therefore the proof of the Theorem 3 will be completed when we have Lemma 3 without the assumption of tail property of Φ. ▯

Now we make our efforts to prove Lemma 3 which gives the uniform order of convergence of the remainder $R_n(f)$. For the purpose we examine the representations of R_n of Stute [7]. Let $H_n, \widetilde{H}_n^0$, and $\widetilde{H}_n^1$ be the empirical (sub-) distribution functions of H, $\widetilde{H}^0$, and $\widetilde{H}^1$, respectively. In order to describe the specific form of the remainder terms R_n we need the following form of $\int f d\widetilde{F}_n$. Stute [7, Lemma 2.1] states that, under continuity of H,

$$(3.9) \qquad \int f d\widetilde{F}_n = \int f(w) \exp\left\{ n \int_{-\infty}^{w-} L_n(z) \widetilde{H}_n^0(dz) \right\} \widetilde{H}_n^1(dw),$$

where $L_n(z) := \ln\left[1 + \left(1/n(1 - H_n(z)) \right) \right]$. The exponential term in (3.9) can be expanded as

$$\exp\left\{ \int_{-\infty}^{Z_i-} \frac{\widetilde{H}^0(dz)}{1 - H(z)} \right\} \left[1 + n \int_{-\infty}^{Z_i-} L_n(z) \widetilde{H}_n^0(dz) - \int_{-\infty}^{Z_i-} \frac{\widetilde{H}^0(dz)}{1 - H(z)} \right]$$
$$+ \frac{1}{2} e^{\Delta_i} \left\{ n \int_{-\infty}^{Z_i-} L_n(z) \widetilde{H}_n^0(dz) - \int_{-\infty}^{Z_i-} \frac{\widetilde{H}^0(dz)}{1 - H(z)} \right]^2,$$

where Δ_i is between the two terms in brackets. Write

$$A_{in} := n \int_{-\infty}^{Z_i-} L_n(z) \widetilde{H}_n^0(dz) - \int_{-\infty}^{Z_i-} \frac{\widetilde{H}^0(dz)}{1 - H(z)} := B_{in} + C_{in},$$

with

$$B_{in} := n \int_{-\infty}^{Z_i-} L_n(z) \widetilde{H}_n^0(dz) - \int_{-\infty}^{Z_i-} \frac{\widetilde{H}_n^0(dz)}{1 - H_n(z)}, \text{ and}$$
$$C_{in} := \int_{-\infty}^{Z_i-} \frac{\widetilde{H}_n^0(dz)}{1 - H_n(z)} - \int_{-\infty}^{Z_i-} \frac{\widetilde{H}^0(dz)}{1 - H(z)}.$$

Denote for simplicity

$$H_n(u, v, w) := H_n(u) \widetilde{H}_n^0(v) \widetilde{H}_n^1(w),$$
$$\widetilde{H}_n(u, v, w) := H_n(u) \widetilde{H}^0(v) \widetilde{H}^1(w) + H(u) \widetilde{H}_n^0(v) \widetilde{H}^1(w) + H(u) \widetilde{H}^0(v) \widetilde{H}_n^1(w)$$
$$- 2H(u) \widetilde{H}^0(v) \widetilde{H}^1(w),$$
$$H_n(v, w) := \widetilde{H}_n^0(v) \widetilde{H}_n^1(w), \text{ and}$$
$$\widetilde{H}_n(v, w) := \widetilde{H}^0(v) \widetilde{H}_n^1(w) + \widetilde{H}_n^0(v) \widetilde{H}^1(w) - \widetilde{H}^0(v) \widetilde{H}_n^1(w).$$

Then $R_n(f)$ can be written as

$$R_n(f) = S_{n1}(f) + S_{n2}(f) + R_{n1}(f) - R_{n2}(f) + 2R_{n3}(f),$$

where

$$S_{n1}(f) = n^{-1} \sum_{i=1}^{n} f(Z_i)\gamma(Z_i)\delta_i B_{in},$$

$$S_{n2}(f) = n^{-1} \sum_{i=1}^{n} \frac{1}{2}|f(Z_i)|\delta_i e^{\Delta_i}(B_{in} + C_{in})^2,$$

$$R_{n1}(f) = \iint f(w)\frac{\gamma(w)\{z < w\}[H_n(z) - H(z)]^2}{[1 - H(z)]^2[1 - H_n(z)]}\widetilde{H}_n^0(dz)\widetilde{H}_n^1(dw),$$

$$R_{n2}(f) = \iiint f(w)\frac{\gamma(w)\{v < u, v < w\}}{(1 - H(v))^2}\left(H_n - \widetilde{H}_n\right)(du, dv, dw), \quad \text{and}$$

$$R_{n3}(f) = \iint f(w)\frac{\gamma(w)\{v < w\}}{1 - H(v)}\left(H_n - \widetilde{H}_n\right)(dv, dw).$$

REMARK 1. An observation in a derivation of the uniform order of convergence of $R_n(f)$ is that $A_{in}, B_{in}, C_{in}, \Delta_i$, and γ do not depend on f.

REMARK 2. The fact that, under the assumption of tail property of Φ and the assumption of our main Theorem 1, $\|S_{n1}\|_{\mathcal{F}}$, $\|S_{n2}\|_{\mathcal{F}}$, and $\|R_{n1}\|_{\mathcal{F}}$ are $O(\ln n/n)$, with probability 1, is an immediate consequence of Stute [7, Lemma (2.6), Lemma (2.5), and Lemma (2.7)].

The uniform order of convergence of R_{n2} and R_{n3} need more work.

LEMMA 4. *If* $J(1) < \infty$, *if* $\mathcal{F}$ *has the minimal pointwise property, and if* Φ *has the tail property then* $\left\{n^{1/2}R_{n2}(f) : f \in \mathcal{F}\right\}$ *is tight.*

PROOF OF LEMMA 4: Notice that $\left(H_n - \widetilde{H}_n\right)$ can be factored into $H_n \cdot \left(\widetilde{H}_n^0 - \widetilde{H}^0\right) \cdot \left(\widetilde{H}_n^1 - \widetilde{H}^1\right) + (H_n - H) \cdot \widetilde{H}^0 \cdot \left(\widetilde{H}_n^1 - \widetilde{H}^1\right) + (H_n - H) \cdot \left(\widetilde{H}_n^0 - \widetilde{H}^0\right) \cdot \widetilde{H}^1$. Write $\lambda_n := H_n - H$, $\nu_n := \widetilde{H}_n^0 - \widetilde{H}^0$, and $\mu_n := \widetilde{H}_n^1 - \widetilde{H}^1$. Then $R_{n2}(f)$ can be written as $R_{n2}(f) := R_{n21}(f) + R_{n22}(f) + R_{n23}(f)$ where

$$R_{n21}(f) := \iiint f(w)\frac{\gamma(w)\{v < u, v < w\}}{(1 - H(v))^2}H_n(du)\nu_n(dv)\mu_n(dw),$$

$$R_{n22}(f) := \iiint f(w)\frac{\gamma(w)\{v < u, v < w\}}{(1 - H(v))^2}\lambda_n(du)\widetilde{H}^0(dv)\mu_n(dw), \quad \text{and}$$

$$R_{n23}(f) := \iiint f(w)\frac{\gamma(w)\{v < u, v < w\}}{(1 - H(v))^2}\lambda_n(du)\nu_n(dv)\widetilde{H}^1(dw).$$

Simply denote

$$Z^0 := Z\{\delta = 0\},$$

$$\xi_{ni}(w) := \frac{\{Z_i^0 < w\}\big(1 - H_n\big(Z_i^0\big)\big)}{[1 - H(Z_i^0)]^2}$$

Let T be such that $\Phi(x) = 0$ for all $x > T$, $T < \tau_H$. Note that $\sup\limits_{w \leqslant T} \gamma(w) < \infty$, and $\sup\limits_{w \leqslant T} \sup\limits_{v \in (-\infty, w)} (1 - H(v))^{-2} < \infty$. Put $M := \sup\limits_{w \leqslant T} \gamma(w) \sup\limits_{w \leqslant T} \sup\limits_{v \in (-\infty, w)} \big(1 - H(v)\big)^{-2}$. Notice that

$$\begin{aligned}
R_{n21}(f) &= \iiint f(w) \frac{\gamma(w)\{v < u, v < w\}}{(1 - H(v))^2} H_n(du)\nu_n(dv)\mu_n(dw) \\
&= \int_{-\infty}^{T} f(w)\gamma(w) \int \frac{\{v < w\}}{(1 - H(v))^2} \int_v^{\infty} H_n(du)\nu_n(dv)\mu_n(dw) \\
&= \int_{-\infty}^{T} f(w)\gamma(w) \left[\int \frac{\{v < w\}(1 - H_n(v))}{(1 - H(v))^2}\nu_n(dv)\right]\mu_n(dw) \\
&= \int_{-\infty}^{T} f(w)\gamma(w) \left[\frac{1}{n}\sum_{i=1}^{n}\Big(\xi_{ni}(w) - \int \xi_{ni}(w)d\widetilde{H}^0\Big)\right]\mu_n(dw) \\
&= \int_{-\infty}^{T} f(w)\gamma(w) \left[\frac{1}{n}\sum_{i=1}^{n}\Big(\xi_{ni}(w) - \int \xi_{ni}(w)d\widetilde{H}^0\Big)\right]\widetilde{H}_n^1(dw) \\
&\quad - \int_{-\infty}^{T} f(w)\gamma(w) \left[\frac{1}{n}\sum_{i=1}^{n}\Big(\xi_{ni}(w) - \int \xi_{ni}(w)d\widetilde{H}^0\Big)\right]\widetilde{H}^1(dw) \\
&:= R_{n211}(f) - R_{n212}(f).
\end{aligned}$$

CLAIM 2. $\big\{\sqrt{n}R_{211}(f) : f \in \mathcal{F}\big\}$ is tight.

Notice that

$$\frac{1}{n}\sum_{i=1}^{n}\Big(\xi_{ni}(w) - \int \xi_{ni}(w)d\widetilde{H}^0\Big)$$

is bounded since $\xi_{ni}(w)$ is bounded.

$$\sqrt{n}\sup_{f \in \mathcal{F}}\big|R_{n211}(f)\big| \leqslant \quad \text{Constant} \cdot \sqrt{n} \int_{-\infty}^{T} \Phi(w)\ (w)\widetilde{H}_n^1(dw)$$

which is tight by the central limit theorem.

CLAIM 3. $\big\{\sqrt{n}R_{n212}(f) : f \in \mathcal{F}\big\}$ is tight.

Notice that

$$\sqrt{n}R_{n212}(f) = \sqrt{n}\int_{-\infty}^{T} f(w)\gamma(w)\left[\frac{1}{n}\sum_{i=1}^{n}\left(\xi_{ni}(w) - \int \xi_{ni}(w)d\widetilde{H}^0\right)\right]\widetilde{H}^1(dw)$$

$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_{-\infty}^{T} f(w)\gamma(w)\left(\xi_{ni}(w) - \int \xi_{ni}(w)d\widetilde{H}^0\right)\widetilde{H}^1(dw).$$

Notice also that for fixed $f \in \mathcal{F}$ and for each fixed n

$$\left\{\int_{-\infty}^{T} f(w)\gamma(w)\left(\xi_{ni}(w) - \int \xi_{ni}(w)d\widetilde{H}^0\right)\widetilde{H}^1(dw) : i = 1, \ldots, n\right\}$$

are independent and identically distributed random variables. Therefore we have that

$$\sqrt{n}R_{n212}(f) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(g_{n,f}(Z_i^0) - \int g_{n,f}(Z_i^0)d\widetilde{H}^0\right)$$

where

$$g_{n,f}(x) = \frac{1 - H_n(x)}{[1 - H(x)]^2}\int_{-\infty}^{T} f(w)\gamma(w)\{x < w\}\widetilde{H}^1(dw).$$

Use [**8**, Theorem 2.11.23] to obtain the result. The same reasoning applies to $\left\{\sqrt{n}R_{n22}(f) : f \in \mathcal{F}\right\}$ and $\left\{\sqrt{n}R_{n23}(f) : f \in \mathcal{F}\right\}$. The proof of Lemma 4 is completed.   ☐

LEMMA 5. *If $J(1) < \infty$, if $\mathcal{F}$ has the minimal pointwise property, and if Φ has the tail property then $\left\{n^{1/2}R_{n3}(f) : f \in \mathcal{F}\right\}$ is tight.*

PROOF OF LEMMA 5: This follows from the same argument as in the proof of the last Lemma.   ☐

PROOF OF LEMMA 3: The result immediately follows from the last two Lemmas, Lemma 1 and Remark 2.

Finally, it remains to prove that the assumption of tail property for Φ can be removed using the assumptions of Theorem 1. Indeed, under the assumptions of Theorem 1, Lemma 1 gives us that $\int\left[\Phi(Z)\gamma(Z)\delta\right]^2 dP < \infty$ and $\int \Phi(x)C^{1/2}(x)\widetilde{F}(dx) < \infty$. Now, exactly the same arguments of [**7**, pp. 435–438] apply with Φ instead of φ. This completes the proof.   ☐

4. THE SEQUENTIAL KAPLAN-MEIER INTEGRAL PROCESS

Let $\mathcal{F}$ be a class of functions which are real-valued measurable defined on $\mathbf{R}$. Consider now the sequential Kaplan–Meier integral process $\left\{U_n(t,f) : (t,f) \in \mathbf{R} \otimes \mathcal{F}\right\}$ defined by

$$(4.10) \qquad U_n(t,f) = n^{1/2}\int f(x)\{x \leqslant t\}d\left(\widehat{F}_n - \widetilde{F}\right)(x) \text{ for } (t,f) \in \mathbf{R} \otimes \mathcal{F}.$$

Denote simply $f(\cdot)1_{(-\infty,t]}(\cdot)$ by $f_t(\cdot)$. Write $\xi(t,f) := \xi(f_t)$.

Let $W = \{W(t,f) : (t,f) \in \mathbf{R} \otimes \mathcal{F}\}$ be the mean zero Gaussian process with the covariance function

$$(4.11) \qquad\qquad Cov\big(W(t,f), W(s,g)\big) = Cov\big(\xi(f_t), \xi(g_s)\big).$$

The following uniform central limit theorem for the sequential Kaplan-Meier integral process U_n will be obtained as a corollary of Theorem 2.1.

THEOREM 4.1. *Under the assumptions of Theorem 2.1, $U_n \Rightarrow W$ as random elements of $B(\mathbf{R} \otimes \mathcal{F})$. The limiting process $W = \{W(t,f) : (t,f) \in \mathbf{R} \otimes \mathcal{F}\}$ is the mean zero Gaussian process and the covariance function is given by (4.11). The sample paths of W are bounded and uniformly continuous with respect to the metric $\rho\big((t,f),(s,g)\big) := |s - t| + d(f,g)$.*

PROOF OF THEOREM 4.1: Let $\varepsilon > 0$. Fix $f \in \mathcal{F}$. Choose a grid of points $-\infty = t_0 < t_{1(f)} < \cdots < t_{k(f)} = \infty$ with the property $\int 1_{(t_{i(f)}, t_{i(f)+1}]} dF < \varepsilon$ for each $i(f)$. Since $J(\varepsilon) < \infty$, there exists $\{[f_i^l, f_i^u], i = 0, 1, \ldots, N_{[]}(\varepsilon)\}$ so that for $f \in \mathcal{F}$ there exists $0 \leqslant i \leqslant N_{[]}(\varepsilon)$ satisfying $f_i^l \leqslant f \leqslant f_i^u$ and $d(f_i^l, f_i^u) < \varepsilon$. Now us the brackets $\mathcal{G}_\varepsilon := \big\{[f_i^l 1_{(-\infty, t_{i(f)}]}, f_i^u 1_{(-\infty, t_{i(f)+1}]}]\big\}$, whose elements are finite. Then for $f_t \in \mathcal{G} = \{f_t : (t,f) \in \mathbf{R} \otimes \mathcal{F}\}$, there exists some bracket $[f_i^l 1_{(-\infty, t_{i(f)}]}, f_i^u 1_{(-\infty, t_{i(f)+1}]}] \in \mathcal{G}_\varepsilon$ such that $f_i^l 1_{(-\infty, t_{i,(f)}]} \leqslant f_t \leqslant f_i^u 1_{(-\infty, t_{i(f)+1}]}$ and

$$\rho\big(f_i^l 1_{(-\infty, t_{i(f)})}, f_i^u 1_{(-\infty, t_{i,(f)+1}]}\big) = d(f_i^l, f_i^u) + \int 1_{(t_{i(f)}, t_{i(f)+1}]} dF < 2\varepsilon.$$

This verifies that $\int_0^l [\ln N_{[]}(2\varepsilon, \mathcal{G}, \rho)]^{1/2} d\varepsilon < \infty$. Apply Theorem 1 to complete the proof. □

As the first application of Theorem 4 we consider the invariance principle of the Kaplar–Meier integral as a random elements of $D(\mathbf{R})$, the space of *cadlag* functions on *infinite time scale non-compact* interval $\mathbf{R}$. Let $\varphi : \mathbf{R} \to \mathbf{R}$ be a measurable function such that $\int \varphi^2 dF < \infty$. Consider the Kaplan-Meier integral process $\{U_n(t) : t \in \mathbf{R}\}$ defined by

$$(4.12) \qquad U_n(t) := U_n(\varphi_t) = n^{1/2} \int \varphi(x)\{x \leqslant t\} d(\widehat{F}_n - \widetilde{F})(x) \text{ for } t \in \mathbf{R}.$$

The following result can be considered as the invariance principle of the Kaplan-Meier integral.

COROLLARY 1. *Assume that $\int \big[\varphi(Z)\gamma(Z)\delta\big]^2 dP < \infty$ and $\int |\varphi(x)|C^{1/2}(x)\widetilde{F}(dx)$ $< \infty$. Then $U_n \Rightarrow W$ as random elements of $D(\mathbf{R})$. The process $\big\{W(t) : t \in \mathbf{R}\big\}$ is the mean zero Gaussian with continuous sample paths and the covariance structure is given by $Cov\big(W(s), W(t)\big) = Cov\big(\xi(\varphi_s), \xi(\varphi_t)\big)$.*

PROOF: Apply Theorem 4.1 to $\mathcal{F} = \{\varphi\}$, a singleton class.   ☐

As the second application of Theorem 4 we consider the sequential empirical process based on a complete data of independent and identically distributed random variables. Consider a class $\mathcal{F}$ of square integrable functions which are real-valued measurable defined on $\mathbf{R}$. Consider the process $\big\{S_n(t,f) : (t,f) \in \mathbf{R} \otimes \mathcal{F}\big\}$ defined by

$$(4.13) \qquad S_n(t,f) = n^{1/2}\int f(x)\{x \leqslant t\}d(F_n - F)(x) \ \text{ for } (t,f) \in \mathbf{R} \otimes \mathcal{F}$$

Write for each $(t,f) \in \mathbf{R} \otimes \mathcal{F}$,

$$\xi(t,f) = f(X)\{X \leqslant t\} - \int f(x)\{x \leqslant t\}dF(x).$$

Notice that the covariance of $\xi(t,f)$ and $\xi(s,g)$ is given by

$$(4.14) \qquad \int_{-\infty}^{t\wedge s} f(x)g(x)dF(x) - \int_{-\infty}^{t} f(x)dF(x)\int_{-\infty}^{s} g(x)dF(x).$$

Let $\mathbf{Z} = \big\{\mathbf{Z}(t,f) : (t,f) \in \mathbf{R} \otimes \mathcal{F}\big\}$ be the mean zero Gaussian process with covariance function.

$$(4.15) \qquad Cov\big(\mathbf{Z}(t,f), \mathbf{Z}(s,g)\big) = Cov\big(\xi(f_t), \xi(g_s)\big)$$

The following uniform central limit theorem for the sequential integral process S_n will be obtained as a corrollary of Theorem 4.1. See, for example, [1, Theorem 1.3].

COROLLARY 2. *Suppose $J(1) < \infty$. Then $S_n \Rightarrow \mathbf{Z}$ as random elements of $B(\mathbf{R} \otimes \mathcal{F})$. The limiting process $\mathbf{Z} = \big\{\mathbf{Z}(t,f) : (t,f) \in \mathbf{R} \otimes \mathcal{F}\big\}$ is the mean zero Gaussian process and the covariance function is given by (4.15). The sample paths of $\mathbf{Z}$ are bounded and uniformly continuous with respect to the metric $\rho\big((t,f),(s,g)\big)$ $:= |t - s| + d(f,g)$.*

PROOF: Apply Theorem 4.1 to the class of square integrable functions of $\mathcal{F}$ by specialising the assumptions Theorem 4.1 to the complete data of independent and identically distributed random variables.   ☐

REMARK 3. A functional law of the iterated logarithm for the Kaplan-Meier integral process, under the assumptions of Theorem 2.1, of Strassen type (see, for example, Keulbs and Dudley [4]) readily follows from Theorem 4.1 by a method adapted from Ossiander [5].

References

[1] R.M. Dudley and W. Philipp, 'Invariance principles for sums of Banach space valued random elements and empirical processes', *Z. Wahrsch. Verw. Gebitete* **62** (1983), 509–552.

[2] J. Hoffmann-Jørgensen, *Stochastic processes on Polish spaces* (Aahus Universite, Matematisk Institut, Aarhus, 1991).

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

[4] J. Kuelbs and R.M. Dudley, 'Log log law for Empirical measures', *Ann. Probab.* **8** (1980), 405–418.

[5] M. Ossiander, 'A central limit theorem under metric entropy with L_2 bracketing', *Ann. Probab.* **15** (1987), 897–919.

[6] D. Pollard, *Empirical processes: theory and applications*, Regional conference series in Probability and Statistics **2** (Inst. Math. Statist, Hayward CA, 1990).

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

[8] A.W. Van der Vaart and J.A. Wellner, *Weak convergence and empirical processes with applications to statistics*, Springer Series in Statistics (Springer-Verlag, New York, 1996).

Department of Mathematics and Institute of Basic Science
SungKyunKwan University
Suwon 440-746
Korea
e-mail: jsbae@yurim.skku.ac.kr
          coke@math.skku.ac.kr