BULL. AUSTRAL. MATH. SOC. 60F15, 60F17, 60G42
VOL. 67 (2003) [459–465]

THE UNIFORM LAW OF LARGE NUMBERS FOR THE KAPLAN-MEIER INTEGRAL PROCESS

JONGSIG BAE AND SUNGYEUN KIM

Let $U_n(f) = \int f d(\widehat{F}_n - \widehat{F})$ be the function-indexed Kaplan-Meier integral process constructed from the random censorship model. We study a uniform version of the law of large numbers of Glivenko-Cantelli type for $\{U_n\}$ under the bracketing entropy condition. The main result is that the almost sure convergence and convergence in the mean of the process U_n holds uniformly in $\mathcal{F}$. In proving the result we shall employ the bracketing method which is used in the proof of the uniform law of large numbers for the complete data of the independent and identically distributed model.

1. INTRODUCTION

In this paper, we obtain a uniform version of the law of large numbers of Glivenko-Cantelli type for the function-indexed Kaplan-Meier integral process based on the incomplete data of the random censorship model.

In obtaining the uniform law of large numbers, we observe the role played by bracketing of the indexed class of functions of the process and slightly modify the underlying metric. Then we employ the idea of DeHardt [1] of the bracketing method to the Kaplan-Meier integral process.

The uniform law of large numbers of the present paper extends the one dimensional law of large numbers for random censoring that was established by Stute and Wang [4] and the DeHardt's uniform law of large numbers for independent and identically distributed random variables [1]. Among others our results are stated not only for almost sure convergence but also for convergence in the mean. The results may be used in nonparametric statistical inference in verifying uniform consistency. See Van de Geer [6] for applications.

We begin by 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 functions 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 $\{f(X_i) : i \geqslant 1\}$

Received 17th October, 2002

forms 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

$$S_n(f) = \int f(x)d(F_n - F)(x) \quad \text{for } f \in \mathcal{F}. \tag{1}$$

Throughout the paper events are identified with their indicator functions. So, for example, the summand of the empirical distribution function means the indicator functions of the events $\{X_i \leqslant x\}$.

In a classical probability theory, one is often interested in establishing three famous limit theorems: the law of large numbers, the central limit theorem, and the law of the iterated logarithm. These topics have their own importance in the classical probability theory. The results also have applications in a parametric statistical inference. In recent years, under the topic of empirical process theory, authors have become interested in the uniform analogue of the three theorems: the uniform law of large numbers of Glivenko-Cantelli type, the uniform central limit theorem for Donsker type, and the uniform law of the iterated logarithm for Strassen type. When the underlying data are independent and identically distributed and complete, these topics are well developed up to a generality of function indexed process and usefully applied in nonparametric statical inference. See for example, DeHardt [**1**] for the uniform law of large numbers, and Ossiander [**3**] for the uniform central limit theorem and the uniform law of the iterated logarithm.

In survival analysis, authors deal with statistical inference problems based on incomplete data and the random censorship model. Since Kaplan and Meier introduced the model [**2**], this topic has been studied in various directions. The Kaplan–Meier estimator has been generalised and refined up to a Kaplan–Meier integral and some limit theorems for this integral have been obtained. See, for example, Stute and Wang [**4**] for the law of large numbers and Stute [**5**] for the central limit theorem.

In the present paper, we deal with the uniform law of large numbers problem for the function-indexed Kaplan–Meier integral process.

Developing a uniform law of large numbers for a function-indexed process such as S_n has usually meant that $\sup_{F\in\mathcal{F}}\left|S_n(f)\right|$ converges to zero in a certain sense under a certain entropy condition on the class $\mathcal{F}$. The process is indexed by $\mathcal{F}$ and is treated as random elements in $B(\mathcal{F})$, the space of bounded real-valued function on $\mathcal{F}$, taken with the sup norm $\|\cdot\|_{\mathcal{F}}$. It is known that $\left(B(\mathcal{F}), \|\cdot\|_{\mathcal{F}}\right)$ forms a Banach space.

We use the following defintion of almost sure convergence and convergence in the mean. See, for a recent reference, Van der Vaart and Wellner [**7**].

DEFINITION 1: A sequence of $B(\mathcal{F})$-valued random functions $\{Y_n\}$ converges with probability 1 to a constant c if

$$P^*\{\sup_{f\in\mathcal{F}} Y_n(f) \to c\} = P\{\sup_{f\in\mathcal{F}} Y_n(f)^* \to c\} = 1.$$

The sequence $\{Y_n\}$ converges in the mean to a constant c if

$$E^* \sup_{f\in\mathcal{F}} Y_n(f) = E \sup_{f\in\mathcal{F}} Y_n(f)^* \to c.$$

Here E^* denotes the upper expectation with respect to the outer probability P^*, and $\sup_{f\in\mathcal{F}} Y_n(f)^*$ is the measurable cover function of $\sup_{f\in\mathcal{F}} Y_n(f)$.

In 1971, DeHardt [**1**] obtained the uniform law of large numbers for the sequence of independent and identically distributed random variables under bracketing entropy. DeHardt's result states that if $\mathcal{F}$ has a bracketing entropy then $\sup_{\mathcal{F}} |S_n| \to 0$ with probability 1.

In 1993, taking a statistical point of view, Stute and Wang [**4**] proved, using a super martingale method, a one dimensional law of large numbers for the Kaplan–Meier integral based on the incomplete data of the random censorship model.

The aim of our work is to extend Stute and Wang's law of large numbers for the Kaplan-Meier integral to a function indexed process version by employing DeHardt's idea of the bracketing method.

In Section 2, we introduce the Kaplan-Meier integral process, establish the uniform law of large numbers for the process, and discuss some corollaries of the main result.

2. THE MAIN RESULTS

We introduce 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 leads us to consider a subdistribution function $\widetilde{F}$ defined by

$$\text{(2.2)} \qquad \widetilde{F}(x) = F(x)\{x < \tau_H\} + \left[F(\tau_H-) + \{\tau_H \in A\}F\{\tau_H\}\right]\{x \geqslant \tau_H\}.$$

https://doi.org/10.1017/S0004972700037254 Published online by Cambridge University Press

Let $\mathcal{F} \subset \mathcal{L}_1(F) := \left\{f : \int |f(x)| F(dx) < \infty\right\}$ be a class of functions which are real-valued measurable defined on $\mathbf{R}$. In this paper, we shall use the metric defined by

$$d(f,g) := \int |f(x) - g(x)| \widetilde{F}(dx).$$

REMARK 1. In the case $\tau_H < \infty$, such as independent and identically distributed model of the complete data, $\widetilde{F}$ boils down to F so that the metric d becomes the usual $\mathcal{L}_1$ metric.

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 [**7**] for a recent reference.

DEFINITION 2: Given two functions l and u, the bracket $[l,u]$ is the set of all functions f with $l \leqslant f \leqslant u$. An ε-bracket is a bracket $[l,u]$ with $\int (u-l)(x)\widetilde{F}(dx) < \varepsilon$. The bracketing number $N_{[\,]}(\varepsilon) := N_{[\,]}(\varepsilon, \mathcal{F}, d)$ is the minimum number of ε-brackets needed to cover $\mathcal{F}$. We say that $\mathcal{F}$ has a bracketing entropy if

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

We now consider the Kaplan–Meier integral process $\{U_n\}$ defined by

$$U_n(f) = \int f d\big(\widehat{F}_n - \widetilde{F}\big) \quad \text{for } f \in \mathcal{F}, \tag{2.3}$$

where $\widehat{F}_n$ is the Kaplan–Meier product limit estimator of the underlying distribution F in the random censorship model. See Kaplan and Meier [**2**].

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.

Our goal is to show that the almost sure and mean convergence of the process $\{U_n\}$, given in (2.3), to 0 holds uniformly for a class $\mathcal{F} = \{f\}$ of f's that have a bracketing entropy condition.

We are ready to state the main result of the paper.

THEOREM 2.1. *Suppose that F and G do not have jumps in common. Suppose that $\mathcal{F}$ has a bracketing entropy. Then, with the probability 1 and in the mean,*

$$\sup_{f\in\mathcal{F}} \left| \int f(x)\big(\widehat{F}_n - \widetilde{F}\big)(dx) \right| \to 0. \tag{2.4}$$

REMARK 2. We note that the bracketing proof of a uniform law of large numbers, due to DeHardt [**1**], is relatively short and concise comparing with the proof of a uniform central limit theorem. On the other hand, to the best of our knowledge, it is unfortunate that the bracketing approach for the uniform central limit theorem must depend on a complicated chaining argument with stratifications. See Ossiander [**3**].

We shall use the following proposition that appears in Stute and Wang [**4**].

PROPOSITION 1. *Suppose that F and G do not have jumps in common. Then, with probability 1 and in the mean,*

$$\int f(x)\big(\widehat{F}_n - \widetilde{F}\big)(dx) \to 0.$$

We are ready to give the proof of Theorem 2.1. We find that the idea of DeHardt [**1**] still works in getting the uniform law of large numbers for the Kaplan–Meier integral process, in both sense of the almost sure convergence and the mean convergence.

PROOF: Fix $\varepsilon > 0$. Choose finitely many ε-brackets $[l_i, u_i]$ whose union contains $\mathcal{F}$ and such that $\int (u_i - l_i)(x)\widetilde{F}(dx) < \varepsilon$ for every $i = 1, \dots, N_{[\,]}(\varepsilon)$. Then, for every $f \in \mathcal{F}$, there is a bracket such that

$$\begin{aligned} U_n(f) &= \int f(x)\widehat{F}_n(dx) - \int f(x)\widetilde{F}(dx) \\ &= \int f(x)\widehat{F}_n(dx) - \int u_i(x)\widetilde{F}(dx) + \int u_i(x)\widetilde{F}(dx) - \int f(x)\widetilde{F}(dx) \\ &\leqslant \int u_i(x)\big(\widehat{F}_n - \widetilde{F}\big)(dx) + \int (u_i - l_i)(x)\widetilde{F}(dx). \end{aligned}$$

Consequently

$$\sup_{f\in\mathcal{F}} U_n(f) \leqslant \max_{1\leqslant i\leqslant N_{[\,]}(\varepsilon)} \int u_i(x)\big(\widehat{F}_n - \widetilde{F}\big)(dx) + \varepsilon.$$

The right hand side converges almost surely and in the mean to ε by Proposition 1. Combining with a similar argument for $\inf_{f\in\mathcal{F}} U_n(f)$ yields that

$$\limsup_{n\to\infty} \sup_{f\in\mathcal{F}} \big|U_n(f)\big|^* \leqslant \varepsilon,$$

almost surely, for every $\varepsilon > 0$. Take a sequence $\varepsilon_m \downarrow 0$ to see that the lim sup must actually be zero almost surely. The proof of Theorem 2.1 is completed. □

We consider the $\mathbf{R}$-indexed Kaplan–Meier integral process as a random element of $D(\mathbf{R})$, the space of *cadlag* functions on *infinite finite time scale non-compact* interval $\mathbf{R}$.

Let $\varphi : \mathbf{R} \to \mathbf{R}$ be a measurable function such that $\int |\varphi| dF < \infty$. Consider the Kaplan-Meier integral process $\{U_n(t) : t \in \mathbf{R}\}$ defined by

$$U_n(t) := \int_{-\infty}^{t} \varphi(x) d\big(\widehat{F}_n - \widetilde{F}\big)(x) \quad \text{for } t \in \mathbf{R}. \tag{2.5}$$

https://doi.org/10.1017/S0004972700037254 Published online by Cambridge University Press

The following result can be considered as the uniform law of large numbers for the $\mathbf{R}$-indexed Kaplan–Meier integral process.

COROLLARY 1. *Let $\varphi : \mathbf{R} \to \mathbf{R}$ be a measurable function such that $\int |\varphi| dF < \infty$. Suppose that F and G do not have jumps in common. Then, with probability 1 and in the mean,*

$$\sup_{t\in\mathbf{R}} |U_n(t)| = \sup_{t\in\mathbf{R}} \left| \int_{-\infty}^{t} \varphi(x)\big(\widehat{F}_n - \widetilde{F}\big)(dx) \right| \to 0.$$

PROOF: Apply Theorem 2.1 to $\mathcal{F} = \{\varphi \cdot 1_{(-\infty,t]} : t \in \mathbf{R}\}$ which certainly satisfies the bracketing entropy condition. ∎

The following corollary deals with the case that the underlying distribution F is continuous. In this case, the additional assumption that F and G do not have jumps in common is unnecessary.

COROLLARY 2. *Suppose that F is continuous. Suppose that $\mathcal{F}$ has a bracketing entropy. Then, with the probability 1 and in the mean*

$$\sup_{f\in\mathcal{F}} \left| \int f(x)\big(\widehat{F}_n - F\big)(dx) \right| \to 0.$$

PROOF: When F is continuous, we observe from (2.2) that $\widetilde{F}$ boils down to F. The result follows from Theorem 2.1.

Applying Theorem 2.1 we obtain the following corollary which is the well known uniform laws of large numbers under finite entropy with bracketing. See the recent references Van de Geer [**6**] and Van der Vaart and Wellner [**7**]. ∎

COROLLARY 3. *Suppose that $\mathcal{F}$ has a bracketing entropy. If no censoring is present then, with probability 1 and in the mean,*

$$\sup_{f\in\mathcal{F}} \left| \int f(x)(F_n - F)(dx) \right| \to 0. \tag{2.6}$$

PROOF: If no censoring is present, $\widehat{F}_n = F_n$ and $H = F$, so that (2.4) leads to (2.6). The proof is completed. ∎

REFERENCES

[1] J. DeHardt, 'Generalizations of the Glivenko-Cantelli theorem', *Ann. Math. Statist.* **42** (1971), 2050–2055.

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

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

[4] W. Stute and J.L. Wang, 'The strong law under random censorship', *Ann. Statist.* **21** (1993), 1591–1607.

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

[6] S. Van de Geer, *Applications of empirical process theory*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, U.K., 2000).

[7] A. 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
sykim@math.skku.ac.kr