THE SEQUENTIAL UNIFORM LAW OF LARGE NUMBERS

Jongsig Bae and Sungyeun Kim

Abstract. Let $Z_n(s,f) = n^{-1} \sum_{i=1}^{\lfloor ns \rfloor} (f(X_i) - Pf)$ be the sequential empirical process based on the independent and identically distributed random variables. We prove that convergence problems of $\sup_{s,f} |Z_n(s,f)|$ to zero boil down to those of $\sup_f |Z_n(1,f)|$. We employ Ottaviani’s inequality and the complete convergence to establish, under bracketing entropy with the second moment, the almost sure convergence of $\sup_{s,f} |Z_n(s,f)|$ to zero.

1. Introduction and the main result

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables. Let $\{D_n\}$ be the sequential process defined by

$$D_n(s) = n^{-1/2} \sum_{i=1}^{\lfloor ns \rfloor} X_i \text{ for } 0 \leq s \leq 1,$$

where $\lfloor x \rfloor$ denotes the integral part of $x$. It is well known that, under the topic of Donsker’s invariance principle, the process $D_n$ in (1.1) converges weakly to a Gaussian process. See, for example, Billingsley [1].

We have encountered the sequential process $\{G_n\}$ defined by

$$G_n(s) = n^{-1} \sum_{i=1}^{\lfloor ns \rfloor} X_i \text{ for } 0 \leq s \leq 1$$

in mathematical finance literatures. See Shreve [6].

The fact that $G_n(s)$ converges almost surely to $sEX_1$ for each fixed $s \in [0,1]$ is well known as a strong law of large numbers.

It is surprising that the following question is not yet settled down.

Received April 7, 2005.
2000 Mathematics Subject Classification: Primary 60F15; Secondary 62G20.
Key words and phrases: sequential Glivenko-Cantelli class, Ottaviani’s inequality, complete convergence, almost sure convergence, uniform law of large numbers.
QUESTION. What are the asymptotic behaviors of 

$$\sup_{0 \leq s \leq 1} |G_n(s)|?$$

In this paper we provide a solution to the question by Ottaviani’s inequality and complete convergence.

We begin by formulating the above question.

Let \( \{X_i : i \geq 1\} \) be a sequence of independent and identically distributed random variables defined on a probability space \((\Omega, \mathcal{A}, P)\). Given a Borel measurable function \( f : \mathbb{R} \to \mathbb{R} \), we see that \( \{f(X_i) : i \geq 1\} \) forms a sequence of independent and identically distributed random variables that are more flexible in applications than \( \{X_i : i \geq 1\} \).

We state the following sequential strong law of large numbers.

**Theorem 1.1.** Suppose that \( \int f^2(x)P(dx) < \infty \). Then,

$$\sup_{0 \leq s \leq 1} \left| \frac{1}{n} \sum_{i=1}^{[ns]} (f(X_i) - Pf) \right| \to 0 \text{ almost surely.}$$

The proof of Theorem 1.1 will be given at the end of the paper.

2. Sequential Glivenko-Cantelli classes

In addition to the setting in previous section, we consider a class \( \mathcal{F} \) of real-valued Borel measurable functions defined on \( \mathbb{R} \). Define a sequential empirical process \( Z_n \) by

$$Z_n(s, f) = \frac{1}{n} \sum_{i=1}^{[ns]} (f(X_i) - P f) \text{ for } (s, f) \in [0, 1] \otimes \mathcal{F}. $$

Define the empirical measure by

$$P_n(f) := \frac{1}{n} \sum_{i=1}^{n} f(X_i) \text{ for } f \in \mathcal{F}. $$

Observe that

$$Z_n(s, f) = \frac{[ns]}{n} (P_{[ns]} - P)(f)$$

and

$$Z_n(1, f) = (P_n - P)(f). $$

For a process \( \{Y_t\} \) indexed by an arbitrary set, we denote \( \|Y\| \) to mean \( \sup_t |Y_t| \). Let \( P^* \) denote the outer probability with respect to the underlying probability \( P \).
We need the following Ottaviani's inequality.

**Lemma 2.1.** Let \( X_1, \ldots, X_n \) be independent stochastic processes indexed by an arbitrary set. Let \( S_n := X_1 + \cdots + X_n \). Then for \( \lambda, \mu > 0 \),

\[
P^* \left( \max_{1 \leq k \leq n} ||S_k|| > \lambda + \mu \right) \leq \frac{P^* (||S_n|| > \lambda)}{1 - \max_{1 \leq k \leq n} P^* (||S_n - S_k|| > \mu)}.
\]

**Proof.** See Proposition A.1.1 in Van der Vaart and Wellner [8]. □

We obtain the following maximal inequality that illuminates the role of \( s \in [0, 1] \) is negligible in the problem of law of large numbers. Compare with the role of \( s \in [0, 1] \) in the problem of central limit theorem in section 2.12 in Van der Vaart and Wellner [8]. Let \( S := [0, 1] \otimes \mathcal{F} \).

**Theorem 2.2.** There exists a universal constant \( C \) such that

\[
P^* (||Z_n||_S > 2\epsilon) \leq CP^* (||Z_n(1, \cdot)||_F > \epsilon) \quad \text{for every } \epsilon > 0.
\]

**Proof.** For \( s \in [0, 1] \) and \( f \in \mathcal{F} \), we see that \( |Z_n(s, f)| \leq ||Z_n(s, \cdot)||_F \). Take the sup over \( S \) on both sides to obtain

\[
(2.1) \quad ||Z_n||_S \leq \sup_{0 \leq s \leq 1} ||Z_n(s, \cdot)||_F.
\]

In the right hand side of (2.1) the parameter \( s \) may be restricted to the points \( k/n \) with \( k \) ranging over \( 1, 2, \ldots, n \). Since \( Z_n(s, f) = \frac{\lfloor ns \rfloor}{n} (P_{\lfloor ns \rfloor} - P)(f) \), we have

\[
||Z_n||_S \leq \max_{1 \leq k \leq n} \frac{k}{n} ||P_k - P||_F.
\]

Ottaviani's inequality in Lemma (2.1) gives

\[
P^* \left( \max_{1 \leq k \leq n} \frac{k}{n} ||P_k - P||_F > 2\epsilon \right) \leq \frac{P^*(||P_n - P||_F > \epsilon)}{1 - \max_{1 \leq k \leq n} P^*(||P_k - P||_F > \epsilon)}.
\]

The term \( \max_{1 \leq k \leq n} P^*(\frac{k}{n} ||P_k - P||_F > \epsilon) \) indexed by \( k \leq n_0 \) can be controlled with the help of the inequality

\[
k ||P_k - P||_F \leq 2 \sum_{i=1}^{n_0} F(X_i) + 2n_0 P^* F
\]

for an envelope function \( F \). For sufficiently large \( n_0 \) the term indexed by \( k > n_0 \) are bounded away from 1 by the uniform weak law of large
numbers for \( P_n \). Conclude that the denominator is bounded away from zero. The proof is completed.

Let \( T^* \) denote the measurable cover function for any mapping \( T : \Omega \to \mathbb{R} \).

**Definition 2.3.** A class \( F \) of measurable functions is a sequential weak Glivenko-Cantelli if

\[
\|Z_n\|_F^* \to 0 \text{ in probability.}
\]

The class \( F \) is a weak Glivenko-Cantelli if

\[
Z_n(1, \cdot)^* \to 0 \text{ in probability.}
\]

Sequential strong Glivenko-Cantelli class and strong Glivenko-Cantelli class are defined in a similar fashion by using almost sure convergence.

**Corollary 2.4.** A class \( F \) is a sequential weak Glivenko-Cantelli if and only if it is a weak Glivenko-Cantelli.

**Proof.** Observe that \( \|Z_n(1, \cdot)\|_F \leq \|Z_n\|_S \). By Theorem 2.2, we get

\[
P^*(\|Z_n(1, \cdot)\|_F > 2\epsilon) \leq P^*(\|Z_n\|_S > 2\epsilon) \leq CP^*(\|Z_n(1, \cdot)\|_F > \epsilon).
\]

The proof is completed.

**Corollary 2.5.** Let \( f \) be such that \( \int |f(x)|P(dx) < \infty \). Then, singleton set \( \{f\} \) is a sequential weak Glivenko-Cantelli class.

**Proof.** For a singleton set \( \{f\} \), since \( \int |f(x)|P(dx) < \infty \), we have by weak law of large numbers

\[
\|Z_n(1, \cdot)\|_{\{f\}} = (P_n - P)(f) \to 0 \text{ in probability.}
\]

Apply Corollary 2.4 to finish the proof.

**Definition 2.6.** A sequence of random variables \( \{Y_n\} \) converges completely to constant \( c \) if the series \( \sum_{n=1}^{\infty} P(|Y_n - c| > 1) \) converges. A class \( F \) is a sequential complete Glivenko-Cantelli if

\[
\|Z_n\|_S^* \to 0 \text{ completely.}
\]

It is a complete Glivenko-Cantelli class if

\[
\|Z_n(1, \cdot)\|_F^* \to 0 \text{ completely.}
\]

**Remark 2.7.** It is clear that if \( \sum_{n=1}^{\infty} P(|Y_n - c| > 1) \) converges then \( \sum_{n=1}^{\infty} P(|Y_n - c| > \epsilon) \) converges for every \( \epsilon > 0 \). See Erdos [3].

We will use the following result on complete convergence in proving the results on almost sure convergence.
Corollary 2.8. A class of measurable functions is a sequential complete Glivenko-Cantelli if and only if it is complete Glivenko-Cantelli.

Proof. Recall that
\[ P^*(||Z_n(1, \cdot)||_F > 2\epsilon) \leq P^*(||Z_n||_S > 2\epsilon) \leq CP^* (||Z_n(1, \cdot)||_F > \epsilon). \]
The proof is completed by taking the summation on each side. \[\square\]

3. A sequential strong Glivenko-Cantelli class

Let \( F \subset L_2(P) := \{ f : \int f^2(x)P(dx) < \infty \} \) be a class of real-valued measurable functions defined on \( \mathbb{R} \). In this paper, we will use the \( L^2 \) 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 [8] and Van der Geer [7] for recent references.

Definition 3.1. Given two functions \( l \) and \( u \), the bracket \([l, u]\) is the set of all functions \( f \) with \( l \leq f \leq u \). An \( \epsilon \)-bracket is a bracket \([l, u]\) with \( \left[ \int (u - l)^2(x)P(dx) \right]^{1/2} < \epsilon \). The bracketing number \( N_{[1]}(\epsilon) := N_{[1]}(\epsilon, F, d) \) is the minimum number of \( \epsilon \)-brackets needed to cover \( F \).

We say that \( F \) has a bracketing entropy if \( \int_0^\infty [\ln N_{[1]}(\epsilon, F, d)]^{1/2} d\epsilon < \infty \).

Our goal is to find a condition on \( F \) that suffices to be a sequential strong Glivenko-Cantelli class.

We are ready to state the following.

Theorem 3.2. Suppose that \( F \) has a bracketing entropy. Then, it is a sequential strong Glivenko-Cantelli class. That is,
\[ \sup_S \left\{ \frac{1}{n} \sum_{i=1}^{[ns]} (f(X_i) - Pf) \right\}^* \rightarrow 0 \text{ almost surely.} \]

We will use the following complete law of large numbers that appears in Hsu and Robbins [4].

Proposition 3.3. Suppose that \( \int f^2(x)P(dx) < \infty \). Then,
\[ \int f(x)(P_n - P)(dx) \rightarrow 0 \]
completely. That is, the series
\[ \sum_{n=1}^\infty P \left( \left| \int f(x)(P_n - P)(dx) \right| > 1 \right) \]
converges.

Remark 3.4. It is known that the assumption \( \int f^2(x)P(dx) < \infty \) is essential. See Theorem 2 and their conjecture in Hsu and Robbins [4]. See also Erdos [3] for the affirmative answer to the conjecture.

We are ready to perform the proof of Theorem 3.2. We find that the idea of DeHardt [2] is still worked in getting Theorem 3.2.

Proof. In order to get the almost sure convergence, it suffices to show the complete convergence. See Proposition 5.7 in [5]. In view of Corollary 2.8, to obtain complete convergence of \( \sup_S |\frac{1}{n} \sum_{i=1}^{[ns]} (f(X_i) - Pf)|^* \) to zero, it is enough to prove \( \sup_{f \in \mathcal{F}} |\frac{1}{n} \sum_{i=1}^{n} (f(X_i) - Pf)|^* \to 0 \) completely. Represent \( \frac{1}{n} \sum_{i=1}^{n} (f(X_i) - Pf) \) as an integral form \( \int f(x)(P_n - P)(dx) \). We name the integral process as \( U_n(f) \). Fix \( \epsilon > 0 \). Choose finitely many \( \epsilon \)-brackets \([l_i, u_i]\) whose union contains \( \mathcal{F} \) and such that \( \int (u_i - l_i)^2(x)P(dx) < \epsilon^2 \) for every \( i = 1, \ldots, N_{[\cdot]}(\epsilon) \). Then, for every \( f \in \mathcal{F} \), there is a bracket such that

\[
U_n(f) = \int f(x)P_n(dx) - \int f(x)P(dx) = \int f(x)P_n(dx) - \int u_i(x)P(dx) + \int u_i(x)(P_n - P)(dx) \leq \int u_i(x)(P_n - P)(dx) + \int (u_i - l_i)(x)P(dx).
\]

Now, observe that

\[
\int (u_i - l_i)(x)P(dx) \leq \left[ \int (u_i - l_i)^2(x)P(dx) \right]^{1/2} < \epsilon.
\]

Consequently,

\[
\sup_{f \in \mathcal{F}} U_n(f) \leq \max_{1 \leq i \leq N_{[\cdot]}(\epsilon)} \int u_i(x)(P_n - P)(dx) + \epsilon.
\]

The right hand side converges completely to \( \epsilon \) by Proposition 3.3. Combination with a similar argument for \( \inf_{f \in \mathcal{F}} U_n(f) \) yields that

\[
\lim \sup_{n \to \infty} \sup_{f \in \mathcal{F}} |U_n(f)|^* \leq \epsilon,
\]

completely, for every \( \epsilon > 0 \). Take a sequence \( \epsilon_m \downarrow 0 \) to see that the limsup must actually be zero completely. The proof of Theorem 3.2 is completed. \( \square \)
Corollary 3.5. Suppose that $\mathcal{F}$ has a bracketing entropy. Then, it is a sequential complete Glivenko-Cantelli class. That is,

$$\sup_s \left| \frac{1}{n} \sum_{i=1}^{[ns]} (f(X_i) - P f) \right|_* \to 0 \text{ completely.}$$

Proof. This is not a corollary to Theorem 3.2 itself but a corollary to the proof of Theorem 3.2.

We are ready to finish the proof of Theorem 1.1.

Proof. The singleton set $\mathcal{F} = \{ f \}$ certainly satisfies the bracketing entropy condition. The result follows from Theorem 3.2.

Remark 3.6.

1. To the best of our knowledge, we cannot weaken the second moment assumption of $\int f^2(x) P(dx) < \infty$. See Remark 3.4.

2. Convergence of $\sup_{0 \leq s \leq 1} \left| n^{-1} \sum_{i=1}^{[ns]} (f(X_i) - P f) \right|$ to zero in probability is valid under the first moment assumption. See Corollary 2.5.

3. It is our opinion that considering $G_n$ in (1.2) without accompanying $D_n$ in (1.1) is not natural. When we consider $G_n$ together with $D_n$, the second moment condition is presumably given. See Shreve [6].

4. Under a second moment condition, the weak laws can be obtained as a result of central limit theorem by an application of Slutsky's theorem. However, strong laws have to be dealt with separately.

References


**DEPARTMENT OF MATHEMATICS AND INSTITUTE OF BASIC SCIENCE, SUNKYUN-KWAN UNIVERSITY, SUWON 440-746, KOREA**

_E-mail:_ jsbae@skku.edu  
sykim@math.skku.ac.kr