Research Article

Precise Asymptotics for the Uniform Empirical Process and the Uniform Sample Quantile Process

Youyou Chen¹ and Zaichen Qian²

¹School of Science, Zhejiang Sci-Tech University, Hangzhou 310000, China
²Ant Group, Hangzhou 310000, China

Correspondence should be addressed to Zaichen Qian; cyyqzc385@163.com

Received 15 March 2022; Revised 11 April 2022; Accepted 17 June 2022; Published 31 July 2022

1. Introduction and Main Results

Random phenomena exist in almost every branch of science and engineering and permeate every aspect of ordinary people’s modern life [1, 2]. Probability theory is a subject that studies the quantitative regularity of random phenomena everywhere. Probability is a method of thinking about the world [3].

Probability limit theory is one of the main branches of probability theory [4, 5]. The famous probability scientists Kolmogorov and Gnedenko once said, “the epistemological value of probability theory can be revealed only through the limit theorem. Without the limit theorem, it is impossible to understand the real meaning of the basic concepts of probability theory.” Probability limit theory is also an important basis of statistical large sample theory [4]. People are very concerned about whether the estimator approximates the real parameter when the sample size tends to infinity, that is, the so-called consistency in statistical large sample theory. Furthermore, we need to consider the speed at which the estimator approximates the real parameters and how to solve these statistical large sample problems. The solution of these problems must rely on the probability limit theorem.

Let \( \{X, X_n; n \geq 1\} \) be a sequence of independent and identically distributed (i.i.d.) random variables with the common distribution function \( F \), and set \( S_n = \sum_{i=1}^{n} X_i \) for \( n \geq 1 \). Hsu and Robbins [6] introduced the following complete convergence.

\[
\sum_{n=1}^{\infty} P(\{|S_n| \geq \varepsilon n\} < \infty, \quad \varepsilon > 0, \quad (1)
\]

This holds if \( \mathbb{E}X = 0 \), and \( \mathbb{E}X^2 < \infty \). The converse part was proved by Erdős [7]. The complete convergence is stronger than the almost sure convergence. Obviously, the sum in (1) tends to infinity as \( \varepsilon \downarrow 0 \).

The first result on the convergence rate of this kind was given by Heyde [8]. It is proved that

\[
\lim_{\varepsilon \downarrow 0} \varepsilon^2 \sum_{n=1}^{\infty} P(\{|S_n| \geq \varepsilon n\} = \mathbb{E}X^2, \quad (2)
\]

if \( \mathbb{E}X = 0 \), and \( \mathbb{E}X^2 < \infty \). Heyde [8], Alam [4] got general conclusions and termed them “precise asymptotics.”
The precise asymptotics for “$S_n$” have been extensively studied. One can refer to Zhang [9], Huang [10], and so on. Now, we consider the relevant results for the uniform empirical process. Let $\{U_1, U_2, \ldots, U_n\}$ be a sequence of i.i.d. $U[0,1]$-distributed random variables. Define the uniform empirical process as $\alpha_n(t) = n^{-1/2} \sum_{i=1}^{n} (I[U_i \leq t] - t)$, $0 \leq t \leq 1$. Denote the norm of a function $f(t)$ on $[0, 1]$ by $\|f\| = \sup_{0 \leq t \leq 1} |f(t)|$, and log $x = \ln(x)$. The following is one conclusion provided by Zhang and Yang [11].

**Theorem 1.** Let $\{B(t); 0 \leq t \leq 1\}$ be a Brownian bridge, and for any $\delta > 1$, we have $\lim \epsilon^{2\delta+2} \sum_{n=1}^{\infty} (\log n)^{2\delta+1} \epsilon^{\gamma}$

The proof of Theorem 1 is based on the classical method introduced by Gut and Spătaru [12]. In this paper, we consider the situation “$c \sim c_0$” where $c_0$ is a positive constant, and the classical argument for the case of “$c \sim 0$” does not work anymore. We will use more powerful tools, such as strong approximation. Besides the uniform empirical process, we also consider the uniform sample quantile process. Let $0 = U_0^{(n)} \leq U_1^{(n)} \leq \cdots \leq U_n^{(n)} = 1$ denote the order statistics of the random sample $U_1, U_2, \ldots, U_n$, for each $n \geq 1$. Define the uniform quantile function as $U_n(y) = \begin{cases} U_k^{(n)} & \text{if } (k-1)/n < y \leq k/n, \ k = 1, 2, \cdots, n. \\ 0 & \text{if } y = 0 \end{cases}$

The uniform sample quantile process should be defined as $u_n(y) = n^{1/2} (U_n(y) - y)$, $0 \leq y \leq 1$. The following are our main results.

**Theorem 2.** Let $a > -1$, $b > -1$, and $d_n(\epsilon)$ be a function of $\epsilon$ such that $d_n(\epsilon) \log n \rightarrow \tau$ as $n \rightarrow \infty, \epsilon \sim \sqrt{a} + 1/2$. Then,

\[ \lim_{\epsilon \rightarrow 0} \epsilon^{2(1-a)}(\log n)^{b-1} \sum_{n=1}^{\infty} (\log n^2/\log n)^b \left\{ \|u_n\| \geq \sqrt{2 \log n (\epsilon + b_n(\epsilon))} \right\} = 2 \exp[-4\tau \sqrt{a + 1}] \Gamma(b + 1), \]

and

\[ \lim_{\epsilon \rightarrow 0} \epsilon^{2(1-a)}(\log n)^{b-1} \sum_{n=1}^{\infty} (\log n^2/\log n)^b \left\{ \|u_n\| \geq \sqrt{2 \log n (\epsilon + b_n(\epsilon))} \right\} = 2 \exp[-4\tau \sqrt{a + 1}] \Gamma(b + 1), \]

**Theorem 3.** Let $a > -1$, $b > -1$, and $d_n(\epsilon)$ be a function of $\epsilon$ such that $d_n(\epsilon) \log n \rightarrow \tau$ as $n \rightarrow \infty, \epsilon \sim \sqrt{a} + 1/2$. Then,

\[ \lim_{\epsilon \rightarrow 0} \epsilon^{2(1-a)}(\log n)^{b-1} \sum_{n=1}^{\infty} (\log n^2/\log n)^b \left\{ \|u_n\| \geq \sqrt{2 \log \log n (\epsilon + d_n(\epsilon))} \right\} = 2 \exp[-4\tau \sqrt{a + 1}] \Gamma(b + 1), \]

and

\[ \lim_{\epsilon \rightarrow 0} \epsilon^{2(1-a)}(\log n)^{b-1} \sum_{n=1}^{\infty} (\log n^2/\log n)^b \left\{ \|u_n\| \geq \sqrt{2 \log \log n (\epsilon + d_n(\epsilon))} \right\} = 2 \exp[-4\tau \sqrt{a + 1}] \Gamma(b + 1). \]

**Remark 1.** We define the general empirical process as $\beta_n(x) = \sqrt{n} \left( F_n(x) - F(x) \right)$, $-\infty < x < \infty$, where $F_n(x) = 1/n \sum_{i=1}^{n} I_{(-\infty,x]}(X_i)$. If $F(\cdot)$ is a continuous distribution function since $\alpha_n(F(\cdot)) = \beta_n(x)$, the results for $\beta_n(x)$ can be obtained immediately from the uniform case. But we cannot handle the quantile process in the same way.

2. **Proofs**

The starting point of this paper is the empirical distribution function. The empirical distribution function plays a very important role in statistics [13–18]. Although it is not a beautiful piecewise function, as a nonparametric estimation
of the distribution function, it is unbiased, consistent, and asymptotically obeys the normal distribution. The empirical process is constructed on the basis of the empirical distribution function. The uniform empirical process is a special and important one [19–21].

We lay out some lemmas which will be used in the proofs later. Lemma 1 is well known (cf. [22]). Lemma 2 and 3 are from Csörgö and Révész [23, 24].

**Lemma 1.** Let \( \{B(t); 0 \leq t \leq 1\} \) be a Brownian bridge. Then, for all \( x > 0 \),

\[
P[\|B(t)\| \geq x] = 2 \sum_{k=1}^{\infty} (-1)^{k+1} e^{-2k^2 x^2}. \tag{7}
\]

In particular,

\[
P[\|B(t)\| \geq x] \sim 2e^{-2x^2} \text{ as } x \to +\infty. \tag{8}
\]

**Lemma 2.** There exists a sequence of Brownian bridges \( \{B_n(t); 0 \leq t \leq 1\} \) such that for all \( n \) and \( x \) we have

\[
P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - B_n(t)| > n^{-1/2} (K \log n + x)\right) \leq L e^{-\lambda x}, \tag{9}
\]

where \( K, L, \lambda \) are positive absolute constants.

**Lemma 3.** There exists a sequence of Brownian bridges \( \{B_n(t); 0 \leq t \leq 1\} \) such that for each \( n = 1, 2, \ldots \), and for all \( |z| < c \sqrt{n} \) and \( c > 0 \), we have

\[
P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - B_n(t)| > n^{-1/2} (A \log n + z)\right) \leq Be^{-Cz}, \tag{10}
\]

where \( A, B, C, c \) are positive absolute constants.

First, we obtain the conclusion for the Brownian bridge \( \{B(t); 0 \leq t \leq 1\} \).

**Proposition 1.** Let \( a > -1, b > -1, \) and \( b_n(\epsilon) \) be a function of \( \epsilon \) such that

\[
b_n(\epsilon) \log n \to \tau \quad \text{as} \quad \epsilon \sqrt{\alpha + 1} / 2. \tag{11}
\]

Then,

\[
\lim_{\epsilon \to \sqrt{\alpha + 1}/2} 4\epsilon^2 - (a + 1)^b + \sum_{n=1}^{\infty} n^b (\log n)^b
\]

\[
P[\|B\| \geq \sqrt{2 \log n} (\epsilon + b_n(\epsilon))] = 2 \exp[-4\epsilon^2 \sqrt{\alpha + 1}] \Gamma(b + 1). \]

**Proof.** By Lemma 1 and (11), we have \( P[\|B\| \geq \sqrt{2 \log n} (\epsilon + b_n(\epsilon))] \sim 2 \exp[-4\epsilon^2 \sqrt{\alpha + 1}] \Gamma(b + 1). \)

We calculate that

\[
\lim_{\epsilon \to \sqrt{\alpha + 1}/2} 4\epsilon^2 - (a + 1)^b + \sum_{n=1}^{\infty} n^b (\log n)^b = \lim_{\epsilon \to \sqrt{\alpha + 1}/2} 4\epsilon^2 - (a + 1)^b + \int_0^\infty x^b (\log n)^b \exp[-4\epsilon^2 \log x] dx
\]

\[
= \lim_{\epsilon \to \sqrt{\alpha + 1}/2} 4\epsilon^2 - (a + 1)^b + \int_0^\infty x^b (\log n)^b \exp[-4\epsilon^2 (a + 1) x] dx
\]

\[
= \int_0^\infty x^b (\log n)^b \exp[-4\epsilon^2 (a + 1) x] dx \Gamma(b + 1).
\]

From (12), and noting that \( \theta \) is arbitrary, we get the proposition immediately. \[\square\]

**Proof of Theorem 2.** Here, we only present the proof for (3) since the argument for (4) is similar. It is obvious, for \( p < -1/2, \)

\[
P\left(\sup_{0 \leq t \leq 1} |B(t)| \geq \sqrt{2 \log n} (\epsilon + b_n(\epsilon)) + (\log n)^p\right)
\]

\[-P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - B(t)| \geq (\log n)^p\right)
\]

\[
\leq P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t)| \geq 2 \log n (\epsilon + b_n(\epsilon))\right) \tag{13}
\]

\[
\leq P\left(\sup_{0 \leq t \leq 1} |B(t)| \geq 2 \log n (\epsilon + b_n(\epsilon)) - (\log n)^p\right)
\]

\[+ P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - B(t)| \geq (\log n)^p\right).
\]

From Lemma 2, we have \( P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - B(t)| \geq (\log n)^p\right) \leq P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - B(t)| \geq K \log n + (a + 2) \log n / \sqrt{\alpha + 1}/2\right) \leq L e^{-\alpha (p + 2)} \), and then \( \sum_{n=1}^{\infty} n^b (\log n)^b \leq L \sum_{n=1}^{\infty} n^b (\log n)^b \leq e^{-\alpha (p + 2)} \). Furthermore, it follows

\[
\lim_{\epsilon \to \sqrt{\alpha + 1}/2} 4\epsilon^2 - (a + 1)^b + \sum_{n=1}^{\infty} n^b (\log n)^b P\left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - B(t)| \geq (\log n)^p\right) = 0. \tag{14}
\]

On the other hand, since \( p < -1/2, \) we have \( P[\|B\| \geq \sqrt{2 \log n} (\epsilon + b_n(\epsilon)) \pm (\log n)^p] = P[\|B\| \geq \sqrt{2 \log n} (\epsilon + b_n(\epsilon)) \pm (\log n)^p] \sim 2 \exp[-4\epsilon^2 \sqrt{\alpha + 1}] \exp[-8b_n(\epsilon) \log n] \sim P[\|B\| \geq \sqrt{2 \log n (\epsilon + b_n(\epsilon))}, \text{ as } n \to \infty]. \)
With Proposition 1, it follows

$$
\lim_{c_n \rightarrow \infty} \left[ 4c^2 - (a + 1) \right]^{b+1} \sum_{n=1}^{\infty} n^a (\log n)^b P \left\{ \sup_{0 \leq t \leq 1} |B(t)| \geq \sqrt{2 \log n \left( \epsilon + b_n(\epsilon) \right)} \right\} = 2 \exp\left[-4r\sqrt{\alpha + 1}\right] \Gamma(b + 1).
$$

From (13) to (15), we get the result of Theorem 2. \(\square\)

**Proof of Theorem 3.** In this part, we only present the outline of the proof for the uniform sample quantile process, so the arguments for Theorem 2 and 3 are mutually complementary.

Follow the proof of Proposition 1 closely, we can get the following conclusion. For any \(0 < \theta < 1\), there exist \(\delta > 0\) and \(n_0\) such that for all \(n \geq n_0\) and \(\epsilon \in (\sqrt{\alpha + 1}/2, \sqrt{\alpha + 1} + \delta)\), \(2 \exp\left[-4r\log n\right] \exp\left[-4r\sqrt{\alpha + 1} - \theta \right] \leq P\left( \sup_{0 \leq t \leq 1} |B(t)| \geq \sqrt{2 \log n \left( \epsilon + d_n(\epsilon) \right)} \right] \leq 2 \exp\left[-4r\log n\right] \exp\left[-4r\sqrt{\alpha + 1} + \theta \right]

On the other hand, \(\lim_{c_n \rightarrow \infty} \left[ 4c^2 - (a + 1) \right]^{b+1} \sum_{n=1}^{\infty} (\log n)^a (\log \log n)^b n \cdot \exp\left[-4r\log n\right] = \Gamma(b + 1).

Therefore, we have

$$
\lim_{c_n \rightarrow \infty} \left[ 4c^2 - (a + 1) \right]^{b+1} \sum_{n=1}^{\infty} (\log n)^a (\log \log n)^b P \left\{ \sup_{0 \leq t \leq 1} |B(t)| \geq \sqrt{2 \log n \left( \epsilon + d_n(\epsilon) \right)} \right\} = 2 \exp\left[-4r\sqrt{\alpha + 1}\right] \Gamma(b + 1).
$$

3. Conclusion

The empirical process theory plays an important role in large sample theory in statistics. The researchers are very much interested in the large sample properties of the statistical estimator. As long as the sample size tends to infinity, the estimator converges to the true value of the parameter. In the procedure of demonstration of large sample properties, especially for the estimators in the semiparameter models, this study on convergence rates for the uniform empirical process and the uniform sample quantile process can provide a series of effective methods and tools.

The limitation of this study may lie in the lack of consideration of the exact asymptotic properties of uniform empirical processes; in addition, in the study of the convergence rate of the uniform empirical process and uniform sample quantile process, the influence of the exact asymptotic behavior of self-regularity and logarithmic law on the convergence rate should also be taken into account.

Due to the needs of practical applications, dependent random samples are often of more interest to statisticians. Positive and negative cocontributes also widely exist in real life and engineering, such as reliability testing, statistical mechanics, and so on. The limit properties of the sequences of associated random variables, such as the law of iterated logarithm and the law of large numbers of the sequences of associated random variables, will be a hot topic in the future. In the future, the asymptotic properties of the test statistics of the model and parameters will be studied by parameter estimators.

**Data Availability**

The data set can be accessed upon request.

**Conflicts of Interest**

The author declares that there are no conflicts of interest.

**Acknowledgments**

This work was supported by the Natural Science Foundation of Zhejiang Province (Grant no. LQ18A010009).

**References**


