Research Article

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

Youyou Chen and Zaichen Qian

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

One of the sources of “invariance principle” is that the limit properties of the uniform empirical process coincide with that of a Brownian bridge. The deep discussion of limit theorem of the uniform empirical process gathered wild interest of the researchers. In this paper, the precise convergence rate of the uniform empirical process is considered. As is well-known, when \( \varepsilon \) tends to 0, the precise asymptotic theorems can be demonstrated by referring to the classical method of Gut and Spătaru, by using some nice probability inequalities and so on. However, if \( \varepsilon \) tends to a positive constant, other powerful methods and tools are needed. The method of strong approximation is used in this paper. The main theorems are proved by using the Brownian bridge \( B(t) \) to approximate the uniform empirical process \( \alpha_n(t) \). The relevant results for the uniform sample quantile process are also presented.

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.

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

This holds if \( \varepsilon X = 0 \), and \( \varepsilon X^2 < \infty \). 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 \( \varepsilon \to 0 \).

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

\[
\lim_{\varepsilon \to 0} \varepsilon^2 \sum_{n=1}^{\infty} P\left|S_n \geq \varepsilon n \right| = \varepsilon X^2,
\]

if \( \varepsilon X = 0 \), and \( \varepsilon X^2 < \infty \). 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 $\{U_1, U_2, \cdots, U_n\}$ 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 $\log x = \ln(x\varepsilon)$. 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

\[
\lim_{n \rightarrow \infty} \frac{\varepsilon}{\sqrt{n}} \sum_{n+1}^{\infty}(\log n)^b P \left\{ \| \alpha_n \| \geq 2 \log n \right\} = 2 \cdot (\log n)^b P \left\{ \| \alpha_n \| \geq \frac{\varepsilon}{\sqrt{n}} \right\}.
\]

The proof of Theorem 1 is based on the classical method introduced by Gut and Spătaru [12]. In this paper, we consider the situation “$\varepsilon \searrow \varepsilon_0$” where $c_0$ is a positive constant, and the classical argument for the case of “$\varepsilon \searrow 0$” 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 $0 = U_0^{(n)} \leq U_1^{(n)} \leq \cdots \leq U_{n}^{(n)} \leq U_{n+1}^{(n)} = 1$ 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 $U_n(y) = \begin{cases} U_k^{(n)} & \text{if } (k-1)/n \leq y < k/n, k = 1, 2, \ldots, n \setminus 0 \text{ if } y = 0 \end{cases}$. 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 $d_n(\varepsilon)$ be a function of $\varepsilon$ such that $d_n(\varepsilon) \log n \longrightarrow \tau$ as $n \longrightarrow \infty, \varepsilon \searrow \sqrt{a + 1/2}$. Then,

\[
\lim_{n \rightarrow \infty} \frac{\varepsilon}{\sqrt{n}} \sum_{n+1}^{\infty}(\log n)^b P \left\{ \| U_n \| \geq 2 \log n \varepsilon \right\} = 2 \cdot (\log n)^b P \left\{ \| U_n \| \geq 2 \varepsilon \right\}.
\]

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

\[
\lim_{n \rightarrow \infty} \frac{\varepsilon}{\sqrt{n}} \sum_{n+1}^{\infty}(\log n)^b P \left\{ \| U_n \| \geq 2 \log n \varepsilon \right\} = 2 \cdot (\log n)^b P \left\{ \| U_n \| \geq 2 \varepsilon \right\}.
\]

**Remark 1.** We define the general empirical process as

\[
\beta_n(x) = \sqrt{n}(F_n(x) - F(x)), \quad -\infty < x < \infty,
\]

where $F_n(x) = 1/n \sum_{i=1}^n I_{(-\infty, x)}(X_i)$. If $F(\cdot)$ is a continuous distribution function since $\alpha_n(F(x)) = \beta_n(x),$ 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 based 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 Csörgő and Révész [23, 24].

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

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

(7)

In particular,

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

(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

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

(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 \( n = 1, 2, \cdots \), and for all \( |z| < c \sqrt{n} \) and \( c > 0 \), we have

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

(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 \( db_n(e) \) be a function of \( s \) such that

\[
b_n(e) \log n \to \tau \text{ as } e^{-\sqrt{2} (\log n)\epsilon}.
\]

(11)

Then,

\[
\lim_{e^{-\sqrt{2} (\log n)\epsilon}} \left[ 4e^{-2} - (a + 1) \right] \sum_{n=1}^{\infty} n^{\delta} \log n \Gamma(b + 1).
\]

Proof. By Lemma 1 and (11), we have \( P[\|B\| \geq \sqrt{2} \log n (e + b_n(e))] \sim 2 \exp[-4 \log n (e + b_n(e))] \exp[-8b_n(e) \log n] \) as \( n \to \infty \), uniformly in \( e \in (\sqrt{\alpha + 1}/2, \sqrt{\alpha + 1}/2 + \delta) \) for some \( \delta > 0 \). Therefore, for any \( 0 < \theta < 1 \), there exist \( \delta > 0 \) and \( n_0 \) such that for all \( n \geq n_0 \) and \( e \in (\sqrt{\alpha + 1}/2, \sqrt{\alpha + 1}/2 + \delta) \),

\[
2 \exp[-4 \log n \exp[-4 \sqrt{\alpha + 1} \tau - \theta]] \leq P[\|B\| \geq 2 \log n (e + b_n(e))] \leq 2 \exp[-4 \log n \exp[-4 \sqrt{\alpha + 1} \tau + \theta]].
\]

(12)

We calculate that

\[
\lim_{e^{-\sqrt{2} (\log n)\epsilon}} \left[ 4e^{-2} - (a + 1) \right] \sum_{n=1}^{\infty} n^{\delta} \log n \exp[-4 \log n] = \lim_{e^{-\sqrt{2} (\log n)\epsilon}} \left[ 4e^{-2} - (a + 1) \right] \int_{e^{-\sqrt{2} (\log n)\epsilon}}^{\infty} \log n \exp[-4 \log n] \frac{dx}{dx}.
\]

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

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

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

(13)

From Lemma 2, we have \( P\left( \sup_{0 \leq t \leq 1} \left| \alpha_n(t) - B(t) \right| \geq (\log n)^p \right) \leq P\left( \sup_{0 \leq t \leq 1} \left| \alpha_n(t) - B(t) \right| \geq K \log n + (a + 2) \log n \right) \leq Le^{-Kx}. \) and then \( \sum_{n=1}^{\infty} n^{\delta} \log n \exp[-4 \log n] \leq C \log n \exp[-4 \log n] < \infty. \) Furthermore, it follows

\[
\lim_{e^{-\sqrt{2} (\log n)\epsilon}} \left[ 4e^{-2} - (a + 1) \right] \sum_{n=1}^{\infty} n^{\delta} \log n \Gamma(b + 1) = 0.
\]

(14)
With Proposition 1, it follows

\[
\lim_{\epsilon \sqrt{n/(\log n)}} \left[ 4e^2 - (a + 1) \right]^{b+1} \sum_{n=1}^{\infty} \frac{n^a (\log n)^b}{n} \mathbb{P} \left\{ \sup_{0 \leq t \leq 1} B(t) \geq \sqrt{2 \log n} (e + b_n(e)) \right\} = 2 \exp \left[ -4r\sqrt{\lambda} + 1 \right] \Gamma(b + 1).
\]  

(15)

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

**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 \(e \in (\sqrt{\lambda} + 1/2, \sqrt{\lambda} + 1/2 + \delta)\), \(2 \exp \left[ -4e^2 \log \log n \right] \exp \left[ -4r\sqrt{\lambda} + 1 - \theta \right] \leq P[\|B\| \geq \sqrt{2 \log \log (e + d_n(e))}] \leq 2 \exp \left[ -4e^2 \log \log n \right] \exp \left[ -4r\sqrt{\lambda} + 1 + \theta \right]

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

Therefore, we have

\[
\lim_{\epsilon \sqrt{n/(\log n)}} \left[ 4e^2 - (a + 1) \right]^{b+1} \sum_{n=1}^{\infty} \frac{(\log n)^a (\log \log n)^b}{n} \left\{ \sup_{0 \leq t \leq 1} \left| B_n(t) - B_n(\epsilon) \right| \right\} = 2 \exp \left[ -4r\sqrt{\lambda} + 1 \right] \Gamma(b + 1).
\]  

(16)

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. In the future, the asymptotic properties of the test statistics 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**

[1] M. Li and W. Zhao, “Golden Ratio Phenomenon of random data obeying von Karman spectrum,” Mathematical Problems in Engineering, vol. 2013, pp. 1–6, 2013.

[2] L. Li, B. Lei, and C. Mao, “Digital Twin in Smart Manufacturing,” Journal of Industrial Information Integration, vol. 26, no. 9, p. 100289, Article ID 100289, 2022.
[3] L. Wang, J. X. Chen, G. Jiang, and B Zheng, “Robust spectrum sensing algorithm based on free probability theory,” Wireless Communications and Mobile Computing, vol. 16, no. 13, pp. 1668–1679, 2016.

[4] I. Alam, “Limiting probability Measures,” JOURNAL OF LOGIC AND ANALYSIS, vol. 12, 2019.

[5] W. Li and M. K. Ng, “On the limiting probability distribution of a transition probability tensor,” Linear and Multilinear Algebra, vol. 62, no. 3, pp. 362–385, 2014.

[6] P. L. Hsu and H. Robbins, “Complete convergence and the law of large numbers,” Proceedings of the National Academy of Sciences, vol. 33, no. 2, pp. 25–31, 1947.

[7] P. Erdos, “On a theorem of Hsu and Robbins,” The Annals of Mathematical Statistics, vol. 20, no. 2, pp. 286–291, 1949.

[8] C. C. Heyde, “A supplement to the strong law of large numbers,” Journal of Applied Probability, vol. 12, no. 01, pp. 173–175, 1975.

[9] L. X. Zhang, “Precise rates in the law of the iterated logarithm,” 2001, https://arxiv.org/abs/math/0610519.

[10] W. Huang and L. X. Zhang, “Precise rates in the law of the logarithm in the Hilbert space,” Journal of Mathematical Analysis and Applications, vol. 304, no. 2, pp. 734–758, 2005.

[11] Y. Zhang and X. Y. Yang, “Precise asymptotics in the law of the iterated logarithm and the complete convergence for uniform empirical process,” Statistics & Probability Letters, vol. 78, no. 9, pp. 1051–1055, 2008.

[12] A. Gut and A. Spătaru, “Precise asymptotics in the Baum-Katz and Davis laws of large numbers,” Journal of Mathematical Analysis and Applications, vol. 248, no. 1, pp. 233–246, 2000.

[13] J. A. Višek, “Empirical distribution function under heteroscedasticity,” Statistics, vol. 45, no. 5, pp. 497–508, 2011.

[14] D. Blanke and D. Bosq, “Polygonal smoothing of the empirical distribution function,” Statistical Inference for Stochastic Processes, vol. 21, no. 2, pp. 263–287, 2018.

[15] G. Hesamian and S. M. Taheri, “Fuzzy empirical distribution function: properties and application,” Kybernetika, vol. 49, no. 6, pp. 962–982, 2013.

[16] A. Munteanu and M. Wornowizki, “Correcting statistical models via empirical distribution functions,” Computational Statistics, vol. 31, no. 2, pp. 465–495, 2016.

[17] G. G. Hu, S. S. Gao, Y. Zhong, and C. Gu, “Asymptotic properties of random Weighted empirical distribution function,” Communications in Statistics - Theory and Methods, vol. 44, no. 18, pp. 3812–3824, 2015.

[18] N. Y. Li, Y. Li, and Y. Liu, “Empirical Bayes Inference for the parameter of power distribution based on Ranked set sampling,” Discrete Dynamics in Nature and Society, vol. 2015, Article ID 760768, 2015.

[19] S. Tang, J. X. Zhang, and F. Q. Niu, “Spatial-temporal Evolution Characteristics and Countermeasures of Urban Innovation space distribution: an empirical study based on data of Nanjing High-Tech Enterprises,” Complexity, vol. 2020, Article ID 2905482, 2020.

[20] G. Gao, G. T. Shi, H. Zou, and S. Zhou, “Characterizing the statistical properties of SAR Clutter by using an empirical distribution,” International Journal of Antennas and Propagation, vol. 2013, pp. 1–8, 2013.

[21] C. S. Marange and Y. Qin, “A Simple empirical Likelihood Ratio test for Normality based on the Moment Constraints of a Half-normal distribution,” JOURNAL OF PROBABILITY AND STATISTICS, vol. 10, pp. 1–10, 2018.

[22] P. Billingsley, Convergence of Probability Measure, Wiley, New York, 1968.