J. Korean Math. Soc. **58** (2021), No. 1, pp. 237–264
https://doi.org/10.4134/JKMS.j200064
pISSN: 0304-9914 / eISSN: 2234-3008

WEAK CONVERGENCE FOR STATIONARY BOOTSTRAP EMPIRICAL PROCESSES OF ASSOCIATED SEQUENCES

Eunju Hwang

Abstract. In this work the stationary bootstrap of Politis and Romano [27] is applied to the empirical distribution function of stationary and associated random variables. A weak convergence theorem for the stationary bootstrap empirical processes of associated sequences is established with its limiting to a Gaussian process almost surely, conditionally on the stationary observations. The weak convergence result is proved by means of a random central limit theorem on geometrically distributed random block size of the stationary bootstrap procedure. As its statistical applications, stationary bootstrap quantiles and stationary bootstrap mean residual life process are discussed. Our results extend the existing ones of Peligrad [25] who dealt with the weak convergence of non-random blockwise empirical processes of associated sequences as well as of Shao and Yu [35] who obtained the weak convergence of the mean residual life process in reliability theory as an application of the association.

1. Introduction, notations and assumptions

Let $\{X_i, i \in \mathbb{Z}\}$ be a stationary sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let F be the common distribution function of $\{X_i, i \in \mathbb{Z}\}$. The empirical distribution function F_n of $X_1, \ldots, X_n$ is defined by

$$F_n(t) \equiv F_n(t, \omega) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(X_i(\omega) \le t),\ t \in \mathbb{R},\ \omega \in \Omega,$$

where $\mathbb{I}(\cdot)$ is the indicator function. The empirical process G_n based on the observations $X_1, \ldots, X_n$ is defined by

$$G_n(t) \equiv G_n(t, \omega) := \sqrt{n}[F_n(t) - F(t)],\ t \in \mathbb{R}.$$

Received February 6, 2020; Revised July 3, 2020; Accepted July 21, 2020.
2010 *Mathematics Subject Classification.* Primary 62E20; Secondary 62F40, 62G30.
Key words and phrases. Stationary bootstrap, empirical process, weak convergence, associated random variables, mean residual life process.
This work was supported by National Research Foundation of Korea (NRF-2018R1D1A1B07048745).

Let $D[-\infty, +\infty]$ be the space of cadlag functions on $[-\infty, +\infty]$ having finite limits at $\pm\infty$. Suppose that $D[-\infty, +\infty]$ is equipped with the Skorohod topology. It is known that under some conditions the empirical process $\{G_n(t), t \in \mathbb{R}\}$ converges in distribution, as a random element of $D[-\infty, +\infty]$, to a Gaussian process G with mean zero and covariance

$$\text{(1)} \qquad \operatorname{Cov}(G(t), G(s)) = \sum_{i\in\mathbb{Z}} \operatorname{Cov}(\mathbb{I}(X_0(\omega) \le t), \mathbb{I}(X_i(\omega) \le s)),\ t, s \in \mathbb{R}.$$

Many researchers discussed weak convergence of the empirical processes of stationary sequences with dependence such as mixing, associated or more general weak dependence. Among them, [21, 35] and [37], in particular, investigated the weak convergence of the empirical processes of associated sequences.

Moreover, several authors dealt with the weak convergence of bootstrapped empirical processes of stationary sequences: for example, those in [3, 4, 22, 25, 30, 31, 34] and [36] developed the consistency of blockwise bootstrap of Künsch [19] for the empirical processes of strong mixing and associated sequences, while Politis and Romano [28] established the weak convergence of the stationary bootstrap empirical processes of strong mixing sequences. [10] studied the validity of the moving-block bootstrap for the empirical distribution of a short memory causal linear process based on [9], and [11] discussed the blockwise bootstrapping empirical distribution of a stationary process with change-point. As an adopted version of the dependent wild bootstrap of [33], [7] proposed a model-free bootstrap method for empirical processes and investigated some applications.

In this paper, we establish weak convergence for a bootstrap version of the empirical processes of associated random variables, by adopting the stationary bootstrap of Politis and Romano [27]. $\{X_i, i \in \mathbb{Z}\}$ is said to be a sequence of *associated* random variables if for every finite subcollection $X_{i_1}, \ldots, X_{i_n}$ and every pair of coordinatewise nondecreasing functions $\varphi_1, \varphi_2 : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{Cov}\left(\varphi_1(X_{i_1}, \ldots, X_{i_n}), \varphi_2(X_{i_1}, \ldots, X_{i_n})\right) \ge 0,$$

whenever the covariance is defined. The notion of the associated random variables implies $P(S_1 \le s_1, \ldots, S_k \le s_k) \ge \prod_{i=1}^k P(S_i \le s_i)$ and $P(S_1 > s_1, \ldots, S_k > s_k) \ge \prod_{i=1}^k P(S_i > s_i)$, where $S_i = \varphi_i(X_{i_1}, \ldots, X_{i_n})$, with nondecreasing φ_i, $i = 1, \ldots, k$, for all $s_1, \ldots, s_k \in \mathbb{R}$. Interesting applications may be partial sum, order statistics and multivariate exponential distribution. Time series models with natural conditions, such as AR models, LARCH(∞) models with nonnegative coefficients, integer-valued AR models, and more general integer bilinear models, belong to a class of associated processes. Besides, associated sequences are widely found in applications of reliability theory and mathematical physics. In particular, in survival or reliability theory, dependence structure is often represented by the associativity because of strong relevance of positive correlated survival data. We refer to [6] for examples of time series models and [5] for introduction to the vast areas of progresses and systems related to association.

As for the empirical processes of the stationary and associated sequences, [37] first proved an empirical central limit theorem under condition

$$\mathrm{Cov}(F(X_1), F(X_n)) = O(n^{-b}) \quad \text{for some } b > 15/2$$

and next [35] weakened the above condition to $b > (3+\sqrt{33})/2$. [21] established a weak convergence theorem for the empirical processes by taking $b > 4$. In a meanwhile, blockwise bootstrap version of the empirical processes of the associated sequences has been discussed by [25], who adopted a weaker condition of the summability $\sum_{i=n}^{\infty} \mathrm{Cov}^{1/3}(X_0, X_i) = O(n^{-\gamma})$ for some $\gamma > 0$ and applied the moving-block bootstrap of Künsch [19]. This paper extends the work of [25] to the stationary bootstrap version of the empirical processes of the associated sequences, but the proof is not quite straightforward.

For the case of mixing, a weak convergence theorem on the stationary bootstrap empirical processes of strong mixing observations has been established by [28]. However, as noted by [21], mixing sequences do not contain associated ones, and the weak convergence for the stationary bootstrap empirical processes of the associated sequences has not yet been proved so far to the best of my knowledge. This paper proves this problem by applying a random central limit theorem with random sample size, which is the length of the blocks in the stationary bootstrap procedure. In the stationary bootstrap, the weak convergence has been proved only for the strong mixing case by [28], whereas in the moving-block bootstrap it has been done by many authors in [3, 4, 22, 25, 30, 31, 34, 36], dealing with various weakly dependent stationary observations under various conditions including mixing and association, as mentioned above. Therefore, our work is a new and novel result as well as yields a significant contribution in the areas of both the stationary bootstrap and the empirical process.

The stationary bootstrap empirical process of the associated sequences is defined as

$$G_n^*(t) := \sqrt{n}(F_n^*(t) - F_n(t)),\ t \in \mathbb{R}, \tag{2}$$

where $F_n^*(t) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(X_i^* \leq t)$ is the empirical distribution function of the stationary bootstrap sample $\{X_i^*,\ i = 1, \ldots, n\}$. The stationary bootstrap of Politis and Romano [27] is a powerful block-resampling technique with random block length following a geometric distribution with parameter p, which tends to zero as sample size goes to infinity. The main goal of this paper is to establish the weak convergence of $G_n^*(t)$ in (2) of the associated sequence to the Gaussian process $G(t)$ in (1) almost surely, conditionally on the stationary observations. In this paper, to show the validity of the stationary bootstrap, a random central limit theorem is established and applied with random sample size following the geometric distribution. A different condition on the rate of parameter p of the geometric distribution from that in [28] is required. For recent references of the stationary bootstrap, see [14] for strong consistency of the stationary bootstrap mean under the ψ-weak dependence of [8], and see [13, 16–18] for statistical applications on time series.

The stationary bootstrap sample $\{X_i^*,\ i = 1, \dots, n\}$ is generated as follows. Let $\{X_1, \dots, X_n\}$ be observed with sample size n. So as to construct blocks with random length, which follows the geometric distribution in the stationary bootstrap procedure, a periodic expansion $\{X_{ni} : i \geq 1\}$ is first made by setting $X_{ni} := X_j$ where $1 \leq j \leq n$ is such that $i = qn + j$ for some integer q. Consider a block $B(i, \ell) = \{X_{ni}, \dots, X_{n(i+\ell-1)}\}$, consisting of ℓ consecutive observations starting from X_{ni} for $i, \ell \in \{1, 2, \dots\}$. Bootstrap observations by means of the stationary bootstrap are obtained by selecting random blocks $\{B(I_k, L_k),\ k = 1, 2, \dots\}$ where $I_1, I_2, \dots,$ are i.i.d. discrete uniform random variables on $\{1, \dots, n\}$ with $P^*(I_1 = i) = 1/n$, $i = 1, \dots, n$, and $L_1, L_2, \dots$ are i.i.d. geometric random variables with parameter $p \in (0, 1)$, $P^*(L_1 = \ell) = p(1-p)^{\ell-1}$, $\ell = 1, 2, \dots$. The collections $\{I_1, I_2, \dots\}$ and $\{L_1, L_2, \dots\}$ are independent of $\{X_1, \dots, X_n\}$ and as well as of each other. Arranging elements in blocks $B(I_1, L_1), \dots, B(I_\kappa, L_\kappa)$ in a series where $\kappa := \inf\{k \geq 1 : L_1 + \dots + L_k \geq n\}$, and deleting the last $L_1 + \dots + L_\kappa - n$ elements, the stationary bootstrap sample $\{X_1^*, \dots, X_n^*\}$ is chosen. Note that $\{X_1^*, \dots, X_n^*\}$ is stationary, conditionally on $\{X_1, \dots, X_n\}$. Also note that the expected block length EL_1 is $1/p$ and $L_1 + \dots + L_{\kappa-1} < n \leq L_1 + \dots + L_\kappa$, and thus $\kappa/(np) \xrightarrow{\text{p}} 1$, more specifically $\kappa = np + O_p(np)$. We assume that $p = p_n$ goes to 0 as $n \to \infty$. For notational simplicity, we suppress dependence of the variables $\{I_1, I_2, \dots\}$, $\{L_1, L_2, \dots\}$ and of the parameter p on n. Here and in the following, P^*, E^* and Var^* denote the conditional probability, the conditional expectation and the conditional variance, respectively, given $X_1, \dots, X_n$.

A main advantage of the stationary bootstrap is that the resampled pseudo-time series is stationary due to geometrically distributed random block size. This property might be a natural requirement, as pointed out by Politis and Romano [27], for reconstructing statistics or quantity of interest related to the original stationary data as well as for applying theories with conditions for stationary weakly dependent time series. Indeed, Politis and Romano [27] have proposed this remarkable resampling technique to overcome the lack of the moving-block bootstrap of Künsch [19] whose resampled pseudo-time series is not stationary. Hence we employ the stationary bootstrap procedure to reconstruct a bootstrap version of empirical process and furthermore bootstrap versions of empirical quantile and mean residual life process as applications. In particular, in the case of positive correlated data of association, the stationarity of the resampled data should be ensured to possess the original probabilistic properties.

We make the following assumptions:

(A1): $\{X_i, i \in \mathbb{Z}\}$ is a stationary and associated sequence of random variables with continuous bounded density and mean zero, and satisfying

$$\sum_{i=n}^{\infty} \mathrm{Cov}^{1/3}(X_0, X_i) \leq O(n^{-\gamma}) \text{ for some } \gamma > 0. \tag{3}$$

(A2): For parameter p of the stationary bootstrap block length, we assume $p = cn^{-\varrho}$ for $\varrho \in (\frac{1}{2}, 1)$ and for some constant $c > 0$, so that $p \to 0$, $np \to \infty$ and $\sqrt{n}p \to 0$ as $n \to \infty$. We denote it by $p \sim n^{-\varrho}$.

By the condition (3) and by the fact that for two associated random variables X and Y each having a continuous bounded density, $\sup_{s,t}(P(X \leq s, Y \leq t) - P(X \leq s)P(Y \leq t)) \leq C \cdot \mathrm{Cov}^{1/3}(X, Y)$ for some generic $C \geq 0$, (see [32]), we have

$$\sup_{n>m} \left| \sum_{i=m}^{n} \mathrm{Cov}\left(\mathbb{I}(s < X_0 \leq t), \mathbb{I}(s < X_i \leq t)\right) \right| \leq Cm^{-\gamma}. \tag{4}$$

Peligrad [25] proved the validity of the moving-block bootstrap empirical processes under condition (4) for a stationary sequence and in turn under (3) for the associated sequence. This work extends the result of [25] to the stationary bootstrap with random blocks. In the work of [25], instead of (A2), some appropriate conditions on the fixed length of blocks have been assumed: see (2.1) and (2.8) of [25].

In proving the weak convergence theorem of the stationary bootstrap empirical processes of associated sequences, a random central limit theorem, which is a modification of [15], is applied on the random block sample, whose length follows the geometric distribution. In [28] the weak convergence of the stationary bootstrap empirical process of the strong mixing sequences was established for $np^2 \to \infty$ as $n \to \infty$ under some mixing conditions, (see Theorem 3.1 of [28]). However, a strong mixing case can be verified with condition $np^2 \to 0$ in (A2) by applying the random central limit theorem of [20] in a similar way to that of the present work. In other words, the validity of the stationary bootstrap empirical process for the mixing can be shown in a wider range of the geometric parameter p by means of the technique of this work.

As its statistical applications, stationary bootstrap empirical quantiles and stationary bootstrap mean residual life processes are constructed and their asymptotic results are investigated. Quantile estimations play a key role not only in the financial risk management such as value-at-risk but are also used for prediction intervals of financial assets. Mean residual life functions and mean residual life processes are important characteristics in survival or reliability theory, in whose areas the dependence structure is frequently represented by the association. Estimates of the mean residual life functions are approached via empirical distributions and thus are, in this work, furthermore approximated by the stationary bootstrap empirical processes.

The remainder of the paper is organized as follows: In Section 2, a main result and its related lemmas are stated, and in Section 3 applications are presented. All proofs are carried out in Section 4. Some existing lemmas are given in Section 5.

2. Main results

The following theorem states the validity of the stationary bootstrap for empirical processes with stationary associated observations.

Theorem 2.1. *Assume that* (A1) *and* (A2) *hold. Then as* $n \to \infty$, *we have*

$$G_n^*(\cdot) \xrightarrow{\text{d}} G(\cdot)$$

P^**-almost surely in the Skorohod topology on* $D[-\infty, \infty]$ *where* $G_n^*(\cdot)$ *is defined in* (2) *and* $G(\cdot)$ *is the Gaussian process with mean zero and covariance structure in* (1).

As discussed in Shao and Yu [35], who established weak convergence theorems for weighted empirical processes of mixings and associated sequences, the theorem proved for the uniform empirical process denoted by $\alpha_n^*(t) := \sqrt{n}(F_n^*(Q(t)) - t)$, $0 \leq t \leq 1$, where $Q(t) = F^{-1}(t) = \inf\{x : F(x) \geq t\}$, will hold automatically for $G_n^*(\cdot)$; $\alpha_n^*(t) = \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(U_i^* \leq t) - t\right]$ by letting $U_i = F(X_i)$ and $U_i^* = F(X_i^*)$, $i = 1, 2, \ldots, n$. Thus we may assume that the marginal law is the uniform law over [0, 1], and the proof of Theorem 2.1 will be given for the uniform distribution.

The result of Theorem 2.1 follows from Lemmas 2.2–2.5 below. Contrary to the case of non-random block length bootstrapping, our Theorem 2.1 is involved with a random central limit theorem with geometrically distributed random size. The random central limit theorem provides a main tool and a novel contribution in proving the weak convergence in the present work. Lemma 2.2 below presents a random central limit theorem, which is used in proving our main results. For Lemma 2.2, we need the following discussion for notations.

For a given function $h : \mathbb{R} \to \mathbb{R}$, let h_{ni} and h_i^* denote $h_{ni} := h(X_{ni})$ and $h_i^* := h(X_i^*)$ just for simplicity, $i = 1, 2, \ldots$, and let $h_0 = E[h_{ni}]$ for each i by the stationarity. For instance, if $h(x) = x$, the identity function, then $h_0 = \mu \equiv E[X_1]$; if $h(x) = (x - \mu)^2$, then $h_0 = Var(X_1)$; and if $h(x) = \mathbb{I}(x \leq t)$ for fixed t, then $h_0 = F(t)$. Note that h_0 is the limit in probability of $\frac{1}{L_k}\sum_{i=I_k}^{I_k+L_k-1} h_{ni}$ as $n \to \infty$, $(p \to 0)$, for each $k = 1, 2, \ldots, \kappa$. In the proofs of results related to Lemma 2.2, the function h is an indicator function; for example, $h_{ni} \equiv h(X_{ni}) = \mathbb{I}(s < X_{ni} \leq t) - (t - s)$ for fixed $0 \leq s < t \leq 1$. [25] used the central limit theorem of the sequence $\{h_{ni}\}$ with a fixed bootstrap block length, instead of the geometrically distributed random length L_k in Lemma 2.2(b). This is a main difference between the blockwise bootstrap of [25] and the stationary bootstrap in this work in proving the validity of the weak convergence. In the followings, I and L are the discrete uniform and geometric random variables, that is, they have the same distributions as those of I_k and L_k, respectively.

Lemma 2.2. *Assume that* (A1) *and* (A2) *hold. We assume that h is a function such that a central limit theorem of $\{h_{ni}\}$ holds with a non-random sample size ℓ, that is,*

$$\frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell}(h_{ni}-h_0) \xrightarrow{\mathrm{d}} Z_h \quad as \quad \ell\to\infty,$$

where Z_h is a normal random variable with mean zero and variance $\sigma_h^2 := Var(h_{n1})$. Then

(a) *we have*

$$\frac{1}{L}\sum_{i=I}^{I+L-1}(h_{ni}-h_0) = \sqrt{p}Z_h + O(p) \ \textit{almost surely (a.s.)}$$

as $n\to\infty$.

(b) *Conditionally on $X_1,\ldots,X_n$, we have*

$$\frac{1}{n}\sum_{i=1}^{n}(h_i^*-\bar{h}) = \frac{1}{\kappa}\sum_{k=1}^{\kappa}\left(\frac{1}{L_k}\sum_{i=I_k}^{I_k+L_k-1}(h_{ni}-h_0)\right) + o(1/\sqrt{n}) \ a.s.$$

as $n\to\infty$, where $\bar{h} := \frac{1}{n}\sum_{i=1}^{n}h_{ni}$ and $\kappa := \inf\{\, k\geq 1 : L_1+\cdots+L_k\geq n\}$.

Lemma 2.3. *Assume that* (A1) *and* (A2) *hold. Let*

$$H_{I,L}\equiv H_{I,L}(s,t) := \sum_{j=I}^{I+L-1}\left\{\mathbb{I}(s<X_{nj}\leq t)-(F(t)-F(s))\right\}. \tag{5}$$

(a) *We have*

$$\frac{1}{L}H_{I,L} = \sqrt{p}Z_H + O(p) \ a.s. \tag{6}$$

as $n\to\infty$, where Z_H is a normal random variable with mean zero and variance

$$\sigma_H^2(s,t) := \lim_{n\to\infty}\frac{1}{n}Var\left(\sum_{i=1}^{n}\mathbb{I}(s<X_i\leq t)\right).$$

(b) *We have*

$$\sup_{s,t}\frac{1}{L}\left[H_{I,L}^2(s,t)-EH_{I,L}^2(s,t)\right] = O(n^{-g}) \ a.s. \tag{7}$$

for some $g>0$, as $n\to\infty$.

Lemma 2.4. *Assume that* (A1) *and* (A2) *hold. Let $v_n^2(s,t) := Var^*(G_n^*(t)-G_n^*(s))$. Then we have*

$$v_n^2(s,t)\leq K(|t-s|^b+n^{-c}) \tag{8}$$

for some constants $K > 0$, $0 < b < 1$ and $c > 0$, and

$$\lim_{n\to\infty} v_n^2(s,t) \text{ exists a.s..} \tag{9}$$

Lemma 2.5. *Assume that* (A1) *and* (A2) *hold. Then* (8) *in Lemma 2.4 implies that $G_n^*(\cdot)$ is tight, that is, for every $\varepsilon > 0$ and $\eta > 0$ there exist $0 < \delta < 1$ and N_0 such that for every $n \geq N_0$*

$$P^*\left(\sup_{|t-s|<\delta} |G_n^*(t) - G_n^*(s)| \geq \eta\right) \leq \varepsilon.$$

In proving the main result in Theorem 2.1, we need Lemmas above and so their proofs will be given prior to that of our main theorem. Lemma 5.1 in Section 5 is a modification of [15] as a special case of the geometrically distributed random sample size. Under assumption (A1) and (A2), Lemma 5.1 implies the result of Lemma 2.2 and in turn of Lemma 2.3(a). Arguments similar to those of [25], but with the geometric random block length, are used in proving Lemma 2.3(b). Results in Lemma 2.2(a) and Lemma 2.3 imply Lemma 2.4. (8) in Lemma 2.4 implies the tightness of Lemma 2.5, and finally, (9) in Lemma 2.4 along with the tightness implies the main result of Theorem 2.1.

Remark 2.6. As for the weak convergence of the stationary bootstrap empirical process for the strong mixing, [28] discussed the strong mixing case along with limit theorems of weakly dependent stationary Hilbert space valued random variables. In their bootstrap central limit theorem of Theorem 3.2, the geometric parameter p of the block length was assumed to be $np^2 \to \infty$ as $n \to \infty$ under strong mixing condition $\sum_{i=1}^{\infty} \alpha_i < \infty$ where α_is are the mixing coefficients. In this work with condition $np^2 \to 0$ in (A2), we can verify the weak convergence of the stationary bootstrap empirical process for the strong mixing sequences on condition $\sum_{i=n}^{\infty} \alpha_i = O(n^{-\gamma})$ for some $\gamma > 0$, as given in [25]. Its proof can be done in the same way by applying the random central limit theorem of [20] and by following the spirit of Theorem 2.3 of [25].

In next section, applications of the stationary bootstrap empirical processes are addressed in two aspects: empirical quantiles and mean residual life functions.

3. Applications

Now we give statistical applications of the main result by developing the stationary bootstrap empirical quantiles and the stationary bootstrap mean residual life functions.

3.1. Stationary bootstrap empirical quantiles

In this subsection, we discuss consistency and asymptotic normality of the stationary bootstrap quantiles of the associated random variables $\{X_i, i \in \mathbb{Z}\}$ with distribution F. Let t_q be the q-quantile of F for $q \in (0,1)$, i.e., $t_q :=$

$Q(q) = F^{-1}(q) = \inf\{x : F(x) \geq q\}$. Let $\hat{t}_{n,q} := F_n^{-1}(q) = \inf\{x : F_n(x) \geq q\}$ and $t^*_{n,q} := F_n^{*\,-1}(q) = \inf\{x : F_n^*(x) \geq q\}$.

First we have the asymptotic normality of empirical quantiles $\hat{t}_{n,q}$ of the associated sequence under an additional assumption.

(A3): F is continuously differentiable at t_q and $F'(t_q) > 0$.

Lemma 3.1. *Assume* (A1) *and* (A3) *hold. Then as $n \to \infty$ we have*

$$\hat{t}_{n,q} - t_q = \frac{q - F_n(t_q)}{F'(t_q)} + o_p(n^{-1/2}) \quad \textit{and} \quad \sqrt{n}(\hat{t}_{n,q} - t_q) \xrightarrow{\mathrm{d}} N(0, \sigma_q^2),$$

where $\sigma_q^2 = Var(G(t_q))/(F'(t_q))^2$.

By the weak convergence theorem of the empirical process for the stationary and associated sequences under (A1), the proof can be given similarly to that of sample quantiles of weakly dependent sequences. The following theorem states the validity of the stationary bootstrap quantile estimates.

Theorem 3.2. *Assume that* (A1)–(A3) *hold. Then as $n \to \infty$ we have*

(a) $t^*_{n,q} - \hat{t}_{n,q} = \frac{F_n(t_q) - F_n^*(t_q)}{F'(t_q)} + o_p(n^{-1/2})$,

(b) $\sqrt{n}(t^*_{n,q} - \hat{t}_{n,q}) \xrightarrow{\mathrm{d}^*} N(0, \sigma_q^2)$.

The quantile estimation plays an important role in financial analysis because useful measures such as value-at-risk and forecast intervals of financial assets are estimated via quantile estimates. The dependent wild bootstrap quantiles (cf. [33]) has been recently developed by [7], who have discussed the validity of the dependent wild bootstrap for empirical processes and some applications. Theorem 3.2(a) is proved by applying the main result of Theorem 2.1, and then Theorem 3.2(b) straightforwardly follows.

3.2. Stationary bootstrap mean residual life processes

Dependence structure in reliability theory is often represented by associated sequences of random variables. In this subsection the mean residual life function in reliability is discussed by means of the stationary bootstrap empirical function. Shao and Yu [35] studied the strong consistency and asymptotic normality of the mean residual life processes of reliability theory as an application of the weighted empirical processes of associated sequences. Under their assumptions, the stationary bootstrap validity of the weighted empirical processes can be obtained in a similar way as above. Indeed, the positive weight function on $(0, 1)$ that is used in the weighted empirical process of [35] is unrelated with the random sequence. So our main result in Theorem 2.1 holds for even the weighted empirical processes under appropriate conditions on the weight function. As pointed out by [35], a direct application of the weighted empirical process is to develop the weak convergence for the integral functionals of the empirical process, and in turn to obtain that of the mean residual life process. According to Theorem 2.1 we may have the consistency of the stationary bootstrap weighted empirical processes for the associated sequence of uniform distributed random variables $\{U_i, i \in \mathbb{Z}\}$.

Let $w : (0,1) \to (0,\infty)$ be a positive weight function with satisfying $w(t) \geq c(t(1-t))^{\nu_0}$ for some $c > 0,\ \nu_0 > 0$, and let $\alpha_n(t) = \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(U_i \leq t) - t\right]$, $0 \leq t \leq 1$. And its stationary bootstrap version $\alpha_n^*(t)$ is defined in the same way with $\{U_i^*,\ i = 1,\ldots,n\}$. By Theorem 2.1, we have

$$\alpha_n^*(\cdot)/w(\cdot) \xrightarrow{\text{d}} G_\alpha(\cdot)/w(\cdot), \tag{10}$$

P^*-almost surely in the Skorohod topology on $D[0,1]$ where $G_\alpha(\cdot)$ is the Gaussian process in (1) with U_i in place of X_i. Similarly to Theorem 2.4 and Corollary 2.4 of [35], we can obtain the weak convergence of the integral functionals as follows: for $0 \leq t \leq 1$,

$$\int_0^t \alpha_n^*(s)dQ(s) \xrightarrow{\text{d}} \int_0^t G_\alpha(s)dQ(s), \tag{11}$$

P^*-almost surely in the Skorohod topology on $D[0,1]$ where $Q(t) = F^{-1}(t) = \inf\{x : F(x) > t\}$. Going one step further, this weak convergence is applied to the mean residual life function in reliability theory and the Gaussian approximation of the mean residual life function is proved.

The expected additional lifetime given survival until time x is a function of x, called the mean residual life. If X is a nonnegative random variable representing the life of a component having distribution function F, called the lifetime distribution function, then the mean residual life function at age $x \geq 0$ is defined as

$$M_F(x) := E(X - x|X > x) = \frac{1}{1-F(x)}\int_x^\infty (1 - F(t))\,dt.$$

It is assumed that F is continuous and $EX < \infty$. The empirical counterpart of M_F is given by

$$M_n(x) := M_{F_n}(x) = \frac{1}{1-F_n(x)}\int_x^\infty (1 - F_n(t))\,dt$$

and its stationary bootstrap version is constructed as

$$M_n^*(x) := M_{F_n^*}(x) = \frac{1}{1-F_n^*(x)}\int_x^\infty (1 - F_n^*(t))\,dt.$$

[35] proved the strong consistency and normal approximation for the mean residual life process $M_n - M_F$ as an application of weak convergence for the weighted empirical process of the associated sequences. Here the stationary bootstrap version of the mean residual life process is developed under their conditions.

Additionally we have the following assumption for the stationary associated sequence of lifetime, $\{X_i,\ i \geq 1\}$, with the lifetime distribution function as in [35].

(A4): $\mathrm{Cov}(F(X_1), F(X_n)) = O(n^{-\nu-\epsilon})$ for some $\nu > (3+\sqrt{33})/2$ and $\epsilon > 0$.

Notice that (A4) and other related conditions, for example,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \mathrm{Cov}(X_n, \sum_{i=1}^{n} X_i) < \infty,$$

on the covariance rates in [35, 37], imply (A1). Thus under (A4) our all main results also hold, and (A4) is needed for their results about the mean residual life processes.

Let $T_F := \inf\{t : F(t) = 1\}$ and $T_{F_n} := \inf\{t : F_n(t) = 1\}$. Also, let

$$W_n(x) := \sqrt{n}(M_n(x) - M_F(x)) \text{ and } W_n^*(x) := \sqrt{n}(M_n^*(x) - M_n(x))$$

for $0 \leq x < \min\{T_F, T_{F_n}\}$. The normalized mean residual life process is written as

$$\begin{aligned} W_n(x) &= \frac{\sqrt{n}}{1 - F_n(x)} \left(-\int_x^{\infty} (F_n(t) - F(t))\, dt + M_F(x)(F_n(x) - F(x)) \right) \\ &= \frac{1}{1 - F_n(x)} \left(-\int_{F(x)}^{1} \alpha_n(t) dQ(t) + M_F(x)\alpha_n(F(x)) \right). \end{aligned}$$

Theorem 3.2 of [35] says that if $T < T_F$, then $W_n(\cdot) \xrightarrow{\mathrm{d}} W(\cdot)$ in the Skorohod topology on $D[0, T]$, where

$$W(x) := \frac{1}{1 - F(x)} \left(-\int_{F(x)}^{1} G_\alpha(t) dQ(t) + M_F(x) G_\alpha(F(x)) \right). \tag{12}$$

Following theorem states the strong consistency and the Gaussian approximation of the stationary bootstrap mean residual life processes.

Theorem 3.3. *Assume that* (A2) *and* (A4) *hold. If* $T < \min\{T_F, T_{F_n}\}$*, then as* $n \to \infty$ *we have* (a) $\sup_{0 \leq x \leq T} |M_n^*(x) - M_n(x)| \xrightarrow{\text{a.s.}} 0$ *and* (b) $W_n^*(\cdot) \xrightarrow{\mathrm{d}} W(\cdot)$, P^**-almost surely in the Skorohod topology on* $D[0, T]$.

Remark 3.4. Along with the mean residual lifetime, the mean past lifetime is also an important measure in reliability and survival analysis. The mean past lifetime function of X is defined by

$$K_F(x) := E(x - X | X < x) = \frac{1}{F(x)} \int_0^x F(t)\, dt, \; 0 < x < \infty$$

and its empirical counterpart is given by

$$K_n(x) := \frac{\int_0^x F_n(t)\, dt}{F_n(x)} \mathbb{I}(X_{1:n} \leq x), \; 0 < x < \infty,$$

where $X_{1:n}$ denotes the first order statistics among $X_1, \ldots, X_n$. Properties and estimation of the mean past lifetime function have been discussed by [1], while strong consistency and asymptotic normality of the mean past lifetime process has been established by [24]. As another statistical application of the stationary bootstrap empirical processes, we can explore the stationary bootstrap mean past lifetime process and show its Gaussian approximation.

Furthermore, one of interesting applications is bivariate or multivariate quantile residual life that [23] recently proposed. As pointed out by [23], the concept of bivariate quantile residual life for two possibly dependent components may be useful in aging analysis and classifications, and estimation of this measure is an open problem. See Section 5 of [23]. It will be also interesting to use the stationary bootstrap method for the estimation and to compare its performance with other methods.

4. Proofs

Proof of Lemma 2.2. For (a), we see that by [15] (whose result is rewritten in Lemma 5.1 of the Appendix), we have for each $j = 1, 2, \ldots, n$,

$$\frac{1}{\sqrt{L}} \sum_{i=j}^{j+L-1} (h_{ni} - h_0) \xrightarrow{\mathrm{d}} Z_h,$$

where Z_h is given in Lemma 2.2 with mean zero and variance $Var(h_{n1})$, and Z_h is independent of j, because $\{h_{ni} : i = 1, 2, \ldots\}$ is a stationary sequence in the context of stationary bootstrap with $h_{ni} = h(X_{ni})$. Thus, for $x \in \mathbb{R}$,

$$\begin{aligned} & P\left(\frac{1}{\sqrt{L}} \sum_{i=I}^{I+L-1} (h_{ni} - h_0) \le x\right) \\ &= \sum_{j=1}^{n} P(I = j) P\left(\frac{1}{\sqrt{L}} \sum_{i=j}^{j+L-1} (h_{ni} - h_0) \le x \,\middle|\, I = j\right) \\ &= \sum_{j=1}^{n} \frac{1}{n} [P(Z_h \le x) + o(1)] = P(Z_h \le x) + o(1) \to P(Z_h \le x) \text{ as } n \to \infty. \end{aligned}$$

By Lemma 5.1 we may have $P(|\sqrt{pL} - 1| \ge \delta_n^*) \to 0$ for any nonincreasing sequence δ_n^* tending to 0. Let $B_n = \{|\sqrt{pL} - 1| < \delta_n^*\}$, noting $B_n \supseteq B_{n+1}$ and $P(B_n) \to 1$ as $n \to \infty$, and let $A_m = \{|\sqrt{pL} - 1| < \epsilon_m^*, \forall n \ge m\}$, where $\epsilon_m^* = \max\{\delta_n^* : n \ge m\}$, noting $\bigcap_{n=m}^{\infty} B_n \subset A_m$. Thus we have $1 = \lim_{n\to\infty} P(B_n) \le \lim_{m\to\infty} P(A_m)$, that is, $\lim_{m\to\infty} P(A_m) = 1$. Choose $\delta_n^* = p^{1/2+\epsilon}$ for $\epsilon > 0$, then $\lim_{m\to\infty} P(|1 - \sqrt{pL}|/\sqrt{p} < p^{\epsilon},\ \forall n \ge m) = 1$. By Theorem 7.6 of [26], $(1 - \sqrt{pL})/\sqrt{p} \to 0$ a.s., and thus we have $1/\sqrt{L} = \sqrt{p} + O(p)$ a.s. as $n \to \infty$. From this, we now show

$$\frac{1}{L} \sum_{i=I}^{I+L-1} (h_{ni} - h_0) = \sqrt{p} Z_h + O(p) \text{ a.s.}. \tag{13}$$

To do this, letting $Z_n = \frac{1}{\sqrt{L}} \sum_{i=I}^{I+L-1} (h_{ni} - h_0)$, and $X_n = 1/\sqrt{L}$, $x_n = \sqrt{p} + O(p)$ with $X_n - x_n \to 0$ a.s., we verify that $|X_n Z_n - x_n Z_h| \to 0$ a.s.. For any $\epsilon_1 > 0$, let $C_m = \{|X_n Z_n - x_n Z_h| < \epsilon_1, \forall n \ge m\}$ and $D_m = C_m^{\mathrm{c}}$, complement of C_m. By Theorem 7.6 of [26], it suffices to show that $\lim_{m\to\infty} P(D_m) = 0$. Consider intervals $I_m = [-2^m, 2^m]$, and choose $\epsilon_1 = 2^m\epsilon$ for $\epsilon > 0$. We have

$$\begin{aligned}P(D_m) &\le P(|X_nZ_n - x_nZ_n| > \epsilon_1/2 \text{ for some } n \ge m)\\ &\qquad + P(|x_nZ_n - x_nZ_h| > \epsilon_1/2 \text{ for some } n \ge m)\\ &=: P(G_{m,1}) + P(G_{m,2}).\end{aligned}$$

We observe

$$\begin{aligned}P(G_{m,1}) &= P\left(G_{m,1} \cap \{|Z_n| \in I_m\}\right) + P\left(G_{m,1} \cap \{|Z_n| \in I_m^c\}\right)\\ &\le P\left(\frac{|Z_n|}{2^m}|X_n - x_n| > \epsilon/2, |Z_n| \in I_m \text{ for some } n \ge m\right)\\ &\quad + P\left(|Z_n| \in I_m^c \text{ for some } n \ge m\right)\\ &\le P\left(|X_n - x_n| > \epsilon/2, |Z_n| \in I_m \text{ for some } n \ge m\right)\\ &\quad + P\left(|Z_n| \in I_m^c \text{ for some } n \ge m\right).\end{aligned}$$

Both probabilities above tend to zero as $m \to \infty$. Next, by Chebyshev inequality,

$$\begin{aligned}(G_{m,2}) &= P(|x_nZ_n - x_nZ_h| > 2^m\epsilon/2 \text{ for some } n \ge m)\\ &\le \frac{x_n^2}{4^{m-1}\epsilon^2}E\left[|Z_n - Z_h|^2\right] = O(p/4^m)\end{aligned}$$

which converges to zero as $m \to \infty$. Hence $\lim_{m\to\infty} P(D_m) = 0$ and the desired result in (13) is obtained.

For (b), note that by [14], who discussed the strong consistency of the stationary bootstrap of weak dependent sequences including associated processes, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(h_i^* - \bar{h}) \xrightarrow{\mathrm{d}} Z_h,\ P^*\text{-almost surely}.$$

It suffices to show that $\sqrt{n}\left[\frac{1}{\kappa}\sum_{k=1}^{\kappa}\bar{\Delta}_k\right]$ has the same limiting normality, letting $\Delta_k := \sum_{i=I_k}^{I_k+L_k-1}(h_{ni} - h_0)$ and $\bar{\Delta}_k := \frac{1}{L_k}\Delta_k$, which are sum and average, respectively, of functionals of observations in block $B(I_k, L_k)$ for $k = 1, \ldots, \kappa$. By Lemma 2.2(a) we have $\bar{\Delta}_k := \frac{1}{L_k}\sum_{i=I_k}^{I_k+L_k-1}(h_{ni} - h_0) = \sqrt{p}Z_{h,k} + O(p)$ a.s, where $\{Z_{h,k},\ k = 1, 2, \ldots, \kappa\}$ are normal random variables with mean zero and variance σ_h^2. Note that $\{Z_{h,k},\ k = 1, 2, \ldots, \kappa\}$ are (conditionally) independent because $\{\bar{\Delta}_k : k = 1, 2, \ldots, \kappa\}$ are (conditionally) independent, (given $\{X_1, \ldots, X_n\}$), by the independence of $\{(I_k, L_k) : k = 1, 2, \ldots\}$. Thus

$$\sqrt{n}\left[\frac{1}{\kappa}\sum_{k=1}^{\kappa}\bar{\Delta}_k\right] = \frac{\sqrt{n}}{\kappa}\sum_{k=1}^{\kappa}\left(\sqrt{p}Z_{h,k} + O(p)\right) = \frac{1}{\sqrt{\kappa}}\sum_{k=1}^{\kappa}Z_{h,k} + O(\sqrt{n}p),\ \text{a.s.}$$

since $\kappa = np + O(\sqrt{np})$. Under condition $\sqrt{n}p \to 0$ in (A2), the desired asymptotic normality is obtained. □

Proof of Lemma 2.3. As mentioned above we assume that X_0 has the uniform distribution on $[0,1]$ and rewrite

$$H_{I,L} \equiv H_{I,L}(s,t) = \sum_{j=I}^{I+L-1} \left\{ \mathbb{I}(s < X_{nj} \le t) - (t-s) \right\}.$$

Asymptotic normality of $\frac{1}{L} H_{I,L}$ is straightforwardly given by Lemma 2.2(a). By (1) and by the stationarity,

$$\begin{aligned} \sigma_H^2(s,t) &= \sum_{i=-\infty}^{\infty} \operatorname{Cov}\left(\mathbb{I}(s < X_0 \le t), \mathbb{I}(s < X_{|i|} \le t)\right) \\ &= \lim_{n\to\infty} \frac{1}{n} Var\left(\sum_{i=1}^{n} \mathbb{I}(s < X_i \le t) \right) \end{aligned}$$

exists for every s and t, and thus (a) is obtained.

For (b), let $m = m_n = [p^{-\epsilon}]$ be the integer part of $p^{-\epsilon}$ for some $\epsilon > 1/2$. For fixed s and t, let u and v be two integers such that

$$\frac{u-1}{m} < s \le \frac{u}{m} \le \frac{v-1}{m} < t \le \frac{v}{m}.$$

Define $D_{I,L} \equiv D_{I,L}(s,t)$ by

$$D_{I,L} := H_{I,L} - H_{I,L,m},$$

where

$$H_{I,L,m} \equiv H_{I,L}\left(\frac{u}{m}, \frac{v}{m}\right) := \sum_{j=I}^{I+L-1} \left\{ \mathbb{I}\left(\frac{u}{m} < X_{nj} \le \frac{v}{m}\right) - \frac{(v-u)}{m} \right\}.$$

Using inequality $2|xy| \le \frac{1}{a}x^2 + ay^2$ with $0 < a < 1$, we have the following inequalities a.s.

$$H_{I,L}^2 \le D_{I,L}^2 + H_{I,L,m}^2 + 2|D_{I,L}H_{I,L,m}| \le (1+a)H_{I,L,m}^2 + \left(1 + \frac{1}{a}\right) D_{I,L}^2 \tag{14}$$

and

$$H_{I,L}^2 \ge D_{I,L}^2 + H_{I,L,m}^2 - 2|D_{I,L}H_{I,L,m}| \ge (1-a)H_{I,L,m}^2 + \left(1 - \frac{1}{a}\right) D_{I,L}^2. \tag{15}$$

By subtracting the expectation of (15) from (14), then we have

$$\begin{aligned} & H_{I,L}^2 - EH_{I,L}^2 \\ &\le (1+a)H_{I,L,m}^2 + \left(1 + \frac{1}{a}\right) D_{I,L}^2 - (1-a)EH_{I,L,m}^2 - \left(1 - \frac{1}{a}\right) ED_{I,L}^2 \\ &\le (1+a)(H_{I,L,m}^2 - EH_{I,L,m}^2) + 2aEH_{I,L,m}^2 + \frac{2}{a}(D_{I,L}^2 + ED_{I,L}^2). \end{aligned}$$

By subtracting the expectation of (14) from (15), then we have

$$H_{I,L}^2 - EH_{I,L}^2$$

$$\geq (1-a)H_{I,L,m}^2 + \left(1-\frac{1}{a}\right)D_{I,L}^2 - (1+a)EH_{I,L,m}^2 - \left(1+\frac{1}{a}\right)ED_{I,L}^2$$
$$\geq (1-a)(H_{I,L,m}^2 - EH_{I,L,m}^2) - 2aEH_{I,L,m}^2 - \frac{2}{a}(D_{I,L}^2 + ED_{I,L}^2).$$

Hence we have, a.s.,

$$\begin{aligned}
&\frac{1}{L}|H_{I,L}^2 - EH_{I,L}^2| \\
&\leq 2\left(\frac{1}{L}|H_{I,L,m}^2 - EH_{I,L,m}^2| + \frac{a}{L}EH_{I,L,m}^2 + \frac{1}{aL}(D_{I,L}^2 + ED_{I,L}^2)\right) \\
&=: 2(\xi_1+\xi_2+\xi_3).
\end{aligned}$$

In order to prove (7) of Lemma 2.3(b), we show $\xi_i = O(n^{-g_i})$ a.s. for some $g_i > 0$, for each $i = 1, 2, 3$.

First we observe $\xi_2 := \frac{a}{L}EH_{I,L,m}^2 = \frac{a}{L}Var[H_{I,L,m}]$. Note that, with $q = 1-p$,

$$\begin{aligned}
Var[H_{I,L,m}] &= Var\left[\sum_{j=I}^{I+L-1}\left\{\mathbb{I}\left(\frac{u}{m} < X_{nj} \leq \frac{v}{m}\right) - \frac{(v-u)}{m}\right\}\right] \\
&= \sum_{i=1}^{n}\sum_{\ell=1}^{\infty}\frac{pq^{\ell-1}}{n}Var\left[\sum_{j=i}^{i+\ell-1}\left\{\mathbb{I}\left(\frac{u}{m} < X_{nj} \leq \frac{v}{m}\right) - \frac{(v-u)}{m}\right\}\right].
\end{aligned}$$

By Lemma 5.2, it follows that

$$Var\left[\sum_{j=i}^{i+\ell-1}\left\{\mathbb{I}\left(\frac{u}{m} < X_{nj} \leq \frac{v}{m}\right) - \frac{(v-u)}{m}\right\}\right] \leq K\left(\ell\left|\frac{v-u}{m}\right|^b + \ell^{1-c}\right)$$

for some $K > 0$, $0 < b < 1$ and $c > 0$. Thus,

$$\begin{aligned}
Var[H_{I,L,m}] &\leq K\sum_{i=1}^{n}\sum_{\ell=1}^{\infty}\frac{pq^{\ell-1}}{n}\left(\ell\left|\frac{v-u}{m}\right|^b + \ell^{1-c}\right) \\
&= K\sum_{\ell=1}^{\infty}pq^{\ell-1}\left(\ell\left|\frac{v-u}{m}\right|^b + \ell^{1-c}\right) \qquad (16)\\
&\leq K\left(\frac{1}{m^b}\sum_{\ell=1}^{\infty}pq^{\ell-1}\ell + \sum_{\ell=1}^{\infty}pq^{\ell-1}\ell^{1-c}\right) \\
&= K\left(p^{b\epsilon-1} + p^{c-1}\right)
\end{aligned}$$

since $m = [p^{-\epsilon}]$ and using $\sum_{\ell=1}^{\infty} q^{\ell-1}\ell^a = O(1/p^{a+1})$ for $a > 0$. Hence, $\xi_2 = \frac{a}{L}Var[H_{I,L,m}] \leq K(p^{b\epsilon-1} + p^{c-1})(p + o(p)) = O(p^{\min(b\epsilon,c)}) = O(n^{-g_2}) \xrightarrow{\text{p}} 0$ a.s. for some $g_2 > 0$ with condition of $p \sim n^{-\varrho}$ in (A2).

Secondly, we consider $\xi_1 := \frac{1}{L}|H^2_{I,L,m} - EH^2_{I,L,m}|$. In order to show $\xi_1 = o_p(n^{-g_1})$ for some $g_1 > 0$, we see that by Markov Inequality, for $\delta_1 > 0$,

$$P\left(\frac{n^{g_1}}{L}|H^2_{I,L,m} - EH^2_{I,L,m}| > \delta_1\right) \le \frac{n^{g_1}}{\delta_1}E\left[\frac{1}{L}|H^2_{I,L,m} - EH^2_{I,L,m}|\right].$$

We apply the Taylor expansion of function $f(x,y) = y/x$ to obtain $Y/X = EY/EX-(X-EX)EY/(EX)^2+(Y-EY)/EX+O_p((X-EX)+(Y-EY))^2$ for two random variables X and Y. Note that higher order terms of $Y-EY$ in the O-term become zero a.s. since $f_{yy} = 0$. Taking expectation yields that $E[Y/X]$ is approximated by $EY/EX + O(E([X-EX]^2)EY/(EX)^3)$. Set $X = L$ and $Y = |H^2_{I,L,m} - EH^2_{I,L,m}|$, the last term is

$$\frac{n^{g_1}}{\delta_1}\left(\frac{1}{EL}E\left[|H^2_{I,L,m} - EH^2_{I,L,m}|\right] + O\left(\frac{E[L-p^{-1}]^2}{(EL)^3}E\left[|H^2_{I,L,m} - EH^2_{I,L,m}|\right]\right)\right)$$

and it is equal to $O(n^{g_1}pVar[H_{I,L,m}])$. By (16), we obtain that

$$P\left(\frac{n^{g_1}}{L}|H^2_{I,L,m} - EH^2_{I,L,m}| > \delta_1\right) \le O(n^{g_1}p[p^{b\epsilon-1} + p^{c-1}]) = O(n^{g_1-b\epsilon\varrho} + n^{g_1-c\varrho})$$

for $0 < b < 1$ and $c > 0$, using condition of $p \sim n^{-\varrho}$. The last O-term tends to zero if we choose g_1 so that $g_1 < \min\{b\epsilon\varrho, c\varrho\}$, for which value of g_1, we thus have $\xi_1 = o(n^{-g_1})$ a.s..

Finally, we observe $\xi_3 := \frac{1}{aL}(D^2_{I,L} + ED^2_{I,L})$. We show that $\frac{1}{L}D^2_{I,L} = O(n^{-g_3})$ a.s. for some $g_3 > 0$. Recall

$$\begin{aligned} D_{I,L} &= H_{I,L} - H_{I,L,m} \\ &= \sum_{j=I}^{I+L-1}\left\{\mathbb{I}(s < X_{nj} \le t) - \mathbb{I}\left(\frac{u}{m} < X_{nj} \le \frac{v}{m}\right)\right\} - \left(t - s - \frac{v-u}{m}\right)L. \end{aligned}$$

By the choices of u and v, we have

$$\mathbb{I}\left(\frac{u}{m} < X_{nj} \le \frac{v-1}{m}\right) \le \mathbb{I}\left(s < X_{nj} \le t\right) \le \mathbb{I}\left(\frac{u-1}{m} < X_{nj} \le \frac{v}{m}\right)$$

and then

$$\begin{aligned} -\mathbb{I}\left(\frac{v-1}{m} < X_{nj} \le \frac{v}{m}\right) &\le \mathbb{I}\left(s < X_{nj} \le t\right) - \mathbb{I}\left(\frac{u}{m} < X_{nj} \le \frac{v}{m}\right) \\ &\le \mathbb{I}\left(\frac{u-1}{m} < X_{nj} \le \frac{u}{m}\right). \end{aligned}$$

Hence

$$\mathbb{I}\left(s < X_{nj} \le t\right) - \mathbb{I}\left(\frac{u}{m} < X_{nj} \le \frac{v}{m}\right) \le \max_u\left|\mathbb{I}\left(\frac{u-1}{m} < X_{nj} \le \frac{u}{m}\right)\right|$$

and thus

$$D_{I,L} \le \max_u\left|\sum_{j=I}^{I+L-1}\mathbb{I}\left(\frac{u-1}{m} < X_{nj} \le \frac{u}{m}\right)\right| + \frac{1}{m}L.$$

Since

$$H_{I,L}\left(\frac{u-1}{m},\frac{u}{m}\right)=\sum_{j=I}^{I+L-1}\mathbb{I}\left(\frac{u-1}{m}<X_{nj}\le\frac{u}{m}\right)-\frac{1}{m}L$$

we have

$$D_{I,L}\le\max_u\left|H_{I,L}\left(\frac{u-1}{m},\frac{u}{m}\right)\right|+\frac{2}{m}L.$$

Using $(a+b)^2\le 2(a^2+b^2)$, it follows that

$$\frac{1}{L}D_{I,L}^2\le\frac{2}{L}\max_u H_{I,L}^2\left(\frac{u-1}{m},\frac{u}{m}\right)+\frac{8}{m^2}L$$

which is written as

$$\frac{2}{L}\max_u\left[H_{I,L}^2\left(\frac{u-1}{m},\frac{u}{m}\right)-EH_{I,L}^2\left(\frac{u-1}{m},\frac{u}{m}\right)\right]+\frac{2}{L}\max_u EH_{I,L}^2\left(\frac{u-1}{m},\frac{u}{m}\right)+\frac{8}{m^2}L.$$

Its first and second terms can be shown to be $O(n^{-g_3})$ similarly to the arguments as in ξ_1 and ξ_2. To see the third term, recalling $L=1/p+o(1/p)$ a.s. and $m\sim p^{-\epsilon}$ for some $\epsilon>1/2$ and $p\sim n^{-\varrho}$, $L/m^2=O(n^{-\varrho(2\epsilon-1)})\to 0$ a.s.. Therefore $\xi_3=O(n^{-g_3})$ for some $g_3>0$, and we complete the proof of (7) in the case of the uniform distribution. In the case that X_0 has an arbitrary continuous distribution $F(x)$, the same discussion as in [2], p. 197, [22], p. 993, [25], pp. 895–896, can be done. □

Proof of Lemma 2.4. For fixed $s<t$,

$$\begin{aligned}v_n^2(s,t)&=Var^*(G_n^*(t)-G_n^*(s))\\&=Var^*\left[\sqrt{n}(F_n^*(t)-F_n(t))-\sqrt{n}(F_n^*(s)-F_n(s))\right]\\&=Var^*\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n\mathbb{I}(s<X_i^*\le t)-\frac{1}{\sqrt{n}}\sum_{i=1}^n\mathbb{I}(s<X_i\le t)\right]\\&=:Var^*[S_n^*-S_n],\end{aligned}$$

where

$$S_n\equiv S_n(s,t):=\frac{1}{\sqrt{n}}\sum_{i=1}^n\{\mathbb{I}(s<X_i\le t)-(t-s)\},\tag{17}$$

$$S_n^*\equiv S_n^*(s,t):=\frac{1}{\sqrt{n}}\sum_{i=1}^n\{\mathbb{I}(s<X_i^*\le t)-(t-s)\}.\tag{18}$$

Note that

$$\begin{aligned}E^*S_n^*&=\sqrt{n}E^*\{\mathbb{I}(s<X_1^*\le t)-(t-s)\}\\&=\sqrt{n}\sum_{i=1}^n\sum_{\ell=1}^\infty\frac{1}{n}pq^{\ell-1}E^*[\mathbb{I}(s<X_1^*\le t)-(t-s)|L_1=\ell,I_1=i]\end{aligned}$$

$$= \sqrt{n} \sum_{i=1}^{n} \sum_{\ell=1}^{\infty} \frac{1}{n} p q^{\ell-1} \{\mathbb{I}(s < X_i \le t) - (t-s)\} = S_n.$$

We observe $E^*[S_n^* - S_n]^2$. Let

$$\tilde{\Delta}_k := \frac{1}{L_k} H_{I_k, L_k} = \frac{1}{L_k} \sum_{j=I_k}^{I_k+L_k-1} \{\mathbb{I}(s < X_{nj} \le t) - (t-s)\}. \tag{19}$$

Note that $\{\tilde{\Delta}_k : k = 1, 2, \ldots, \kappa\}$ are independent, conditionally on $\{X_1, \ldots, X_n\}$, i.e., under P^*. By Lemma 2.2(b), conditionally on $X_1, \ldots, X_n$, we have

$$S_n^* - S_n = \frac{\sqrt{n}}{\kappa} \sum_{k=1}^{\kappa} \tilde{\Delta}_k + o(1) \text{ a.s..} \tag{20}$$

Thus, using the (conditional) independence of $\{\tilde{\Delta}_k\colon k = 1, \ldots, \kappa\}$, we have

$$E^*[S_n^* - S_n]^2 = \frac{n}{\kappa^2} \sum_{k=1}^{\kappa} E^* \tilde{\Delta}_k^2 + o(1) = \frac{n}{\kappa} E^* \tilde{\Delta}_1^2 + o(1) \text{ a.s.},$$

where we can take (conditional) expectation to squares of both sides in (20) by the dominated convergence theorem with some dominated function, which is an integrable square function of the normal random variable by Lemma 2.2(a). We write

$$E^* \tilde{\Delta}_1^2 = E^* \left[\frac{1}{L^2} H_{I,L}^2 \right] = E^* \left[\frac{1}{L^2} (H_{I,L}^2 - E H_{I,L}^2) \right] + E \left[\frac{1}{L^2} E H_{I,L}^2 \right]$$

and we will show that

$$\frac{n}{\kappa} \frac{1}{L^2} [H_{I,L}^2 - E H_{I,L}^2] \xrightarrow{\text{p}} 0, \tag{21}$$

$$\lim_{n\to\infty} \frac{n}{\kappa} E \left[\frac{1}{L^2} E H_{I,L}^2 \right] \text{ exists.} \tag{22}$$

To verify (21), using Lemma 2.3(b), for some $g > 0$, and using $1/L = p + o(p)$ a.s., we have

$$\frac{n}{\kappa} \frac{1}{L^2} [H_{I,L}^2 - E H_{I,L}^2] = \frac{n}{\kappa} \frac{1}{L} O(n^{-g}) = \frac{n}{\kappa} (p + o(p)) O(n^{-g}) \text{ a.s.}$$

which is $O(n^{-g})(1 + o(1)) \to 0$ a.s. since $\kappa = np + O(\sqrt{np})$.

To verify (22), we show that

$$\frac{1}{L} H_{I,L}^2 - \frac{1}{pL^2} E H_{I,L}^2 = o(1) \quad \text{a.s..} \tag{23}$$

The left-hand side of (23) can be written as, using $pL = 1 + o(1)$ a.s.,

$$\begin{aligned} \frac{pL H_{I,L}^2 - E H_{I,L}^2}{pL^2} &= \frac{(1 + o(1)) H_{I,L}^2 - E H_{I,L}^2}{L(1 + o(1))} \\ &= \frac{(H_{I,L}^2 - E H_{I,L}^2)}{L} + o(H_{I,L}^2 / L) + o(1) \text{ a.s.} \end{aligned}$$

which tends to zero by Lemma 2.3(b), and thus (23) is shown.

Note that by Lemma 2.3(a),

$$\sigma_H^2(s,t) = \lim_{n\to\infty} Var\left(\frac{1}{\sqrt{L}} H_{I,L}\right) = \lim_{n\to\infty} E\left(\frac{1}{L_k} H_{I_k,L_k}^2\right).$$

Therefore, by (23) we have

$$\sigma_H^2(s,t) = \lim_{n\to\infty} \frac{1}{p} E\left[\frac{1}{L^2} E H_{I,L}^2\right]$$

and in (22),

$$\lim_{n\to\infty} \frac{n}{\kappa} E\left[\frac{1}{L^2} E H_{I,L}^2\right] = \lim_{n\to\infty} \frac{n}{\kappa} p[\sigma_H^2(s,t) + o(1)] = \sigma_H^2(s,t).$$

Hence (22) holds, and therefore

$$\lim_{n\to\infty} v_n^2(s,t) = \sigma_H^2(s,t). \tag{24}$$

Also, by Lemma 5.2, (8) holds. □

Proof of Lemma 2.5. In the proof of Lemma 2.4, (see (6), (19) and (20)), we note that

$$G_n^*(t) - G_n^*(s) = \frac{\sqrt{n}}{\kappa} \sum_{k=1}^{\kappa} \tilde{\Delta}_k + o(1) \text{ a.s.},$$

and $\frac{\sqrt{n}}{\kappa}\tilde{\Delta}_k = O((np)^{-1/2})$ a.s.. Applying Bennett's inequality (see [29], p. 192), for every $\eta > 0$ we have

$$\begin{aligned} P^*(|G_n^*(t) - G_n^*(s)| > \eta) &\le P^*\left(\left|\frac{\sqrt{n}}{\kappa}\sum_{k=1}^{\kappa}\tilde{\Delta}_k\right| > \eta\right) \\ &\le 2\exp\left\{-\frac{1}{2}\frac{\eta^2}{C(|t-s|^b + n^{-c})}\Lambda\left(\frac{\eta(np)^{-1/2}}{C(|t-s|^b+n^{-c})}\right)\right\}, \end{aligned}$$

where $\Lambda(\cdot)$ is a continuous and decreasing function with $\Lambda(0+) = 1$ as in Bennett's inequality of [29], p. 192, and C is a generic constant. If $|t-s| < \delta$, then the last expression cannot exceed

$$2\exp\left\{-\frac{1}{2}\frac{\eta^2}{C(\delta^b + n^{-c})}\Lambda\left(\frac{\eta n^{-\frac{1}{2}(1-\varrho)}}{C(\delta^b + n^{-c})}\right)\right\}$$

for $p \sim n^{-\varrho}$ with $1/2 < \varrho < 1$, which in turn is bounded above by

$$2\exp\left\{-\frac{1}{2}\frac{\eta^2}{C(\delta^b + n^{-c})}\lambda\right\} \le 2\exp\left\{-\frac{1}{4}\frac{\eta^2}{C\delta^b}\lambda\right\}$$

if $\lambda \le \Lambda\left(\frac{\eta n^{-\frac{1}{2}(1-\varrho)}}{C(\delta^b+n^{-c})}\right)$ and if $\delta^b > n^{-c}$. That is, since $\Lambda(\cdot)$ is continuous and decreasing,

$$\Lambda^{-1}(\lambda) \ge \frac{\eta n^{-\frac{1}{2}(1-\varrho)}}{C(\delta^b + n^{-c})} \ge \frac{\eta n^{-\frac{1}{2}(1-\varrho)}}{2C\delta^b}.$$

Hence, we choose δ satisfying $\delta^b > \max\{\eta n^{-\frac{1}{2}(1-\varrho)}/(2C\Lambda^{-1}(\lambda)), n^{-c}\}$ so that

$$P^*(|G_n^*(t) - G_n^*(s)| > \eta) \leq 2\exp\left\{-\frac{1}{4}\frac{\eta^2}{C\delta^b}\lambda\right\}.$$

This is a similar result to that of Lemma 4.1 of [22], p. 990, and the key step in applying a restricted chaining argument given in Theorem VII.26 of [29] applied with the semimetric $d_0(s,t) = C|t-s|^{b/2}$ and the covering number $N(\delta, d_0, T_0) = 1 + [(\delta/C)^{-2/b}]$ for $0 < b < 1$ and $T_0 = [0,1]$. Following the same arguments as in [22], pp. 990–992, and [25], p. 886, (in the proof of Proposition 3.2), it can be shown that $\limsup_{n\to\infty} P^*\left[\sup_{t\in[0,1]} |G_n^*(t) - G_n^*(s)| > \eta\right] = 0$ a.s., which is the same form as in Lemma 4.3 of [22]. The desired result is given by applying Theorem VII.26 of [29], p. 160. Its detailed proof is omitted because the discussion follows in the same way. □

Proof of Theorem 2.1. By Lemma 2.5, the tightness holds under the conditions. The proof is completed by verifying convergence of the finite-dimensional distributions. Note that

$$\sigma_H^2(s,t) = \lim_{n\to\infty} Var\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbb{I}(s < X_i \leq t)\right) = \lim_{n\to\infty} Var(G_n(t) - G_n(s))$$

and $\sigma_H^2(s,t) = \sigma_H^2(0,t) + \sigma_H^2(s,0) - 2\gamma(s,t)$ where $\gamma(s,t) = \mathrm{Cov}(G(s), G(t))$. Also, $v_n^2(s,t) = v_n^2(0,t) + v_n^2(s,0) - 2\Gamma_n(s,t)$ where $\Gamma_n(s,t) = \mathrm{Cov}^*(G_n^*(t), G_n^*(s))$. By (24) in the proof of Lemma 2.4, the limits of $v_n^2(s,t)$ and the two variance terms as $n \to \infty$ exist a.s., and thus the limit of the covariance term above exists a.s. and $\lim_{n\to\infty}\Gamma_n(s,t) = \gamma(s,t)$. For a positive integer J, let $0 \leq t_1 < t_2 < \cdots < t_J \leq 1$ and let $a_1, a_2, \ldots, a_J$ be real numbers. It will be shown that as $n \to \infty$,

$$\sum_{j=1}^J a_j G_n^*(t_j) \xrightarrow{\mathrm{d}} \sum_{j=1}^J a_j G(t_j).$$

We write

$$\begin{aligned}\sum_{j=1}^J a_j G_n^*(t_j) &= \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^J a_j \mathbb{I}(X_i^* \leq t_j) - \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^J a_j \mathbb{I}(X_i \leq t_j)\right] \\ &=: \sqrt{n}[\bar{Y}_n^* - E^*Y_1^*],\end{aligned}$$

letting $Y_i^* = \sum_{j=1}^J a_j \mathbb{I}(X_i^* \leq t_j)$ and $\bar{Y}_n^* = \sum_{i=1}^n Y_i^*/n$. By the central limit theorem of the stationary sequence, the asymptotic normality of $\sum_{j=1}^J a_j G_n^*(t_j)$ holds with asymptotic variance

$$\sigma_J^2 = \lim_{n\to\infty} Var^*\left[\sum_{j=1}^J a_j G_n^*(t_j)\right]$$

$$= \lim_{n\to\infty} \left[\sum_{j=1}^{J} a_j^2 Var^*(G_n^*(t_j)) + 2\sum_{j=1}^{J-1} \sum_{j'=j+1}^{J} \mathrm{Cov}^*(G_n^*(t_j), G_n^*(t_{j'})) \right]$$

which is equal to, by the above argument,

$$\sum_{j=1}^{J} a_j^2 Var(G(t_j)) + 2\sum_{j=1}^{J-1} \sum_{j'=j+1}^{J} \mathrm{Cov}(G(t_j), G(t_{j'})) = Var\left[\sum_{j=1}^{J} a_j G(t_j)\right].$$

Thus the desired convergence of the finite-dimensional distributions is obtained. □

Proof of Theorem 3.2. To prove Theorem 3.2(a), we first show that

$$t_{n,q}^* - t_q = \frac{q - F_n^*(t_q)}{F'(t_q)} + o_p(n^{-1/2}). \tag{25}$$

To show (25), we define $\zeta_n^* = \sqrt{n}[t_{n,q}^* - t_q]$, and

$$\eta_n^* = \frac{\sqrt{n}}{F'(t_q)}[q - F_n^*(t_q)] = \frac{\sqrt{n}}{F'(t_q)}[F(t_q) - F_n^*(t_q)],$$
$$\psi_{t,n}^* = \frac{\sqrt{n}}{F'(t_q)}[F(t_q + t/\sqrt{n}) - F_n^*(t_q + t/\sqrt{n})].$$

By Theorem 2.1, $\sqrt{n}[F_n(t_q) - F_n^*(t_q)]$ converges to a normal limit, and thus $\{\sqrt{n}[(q - F_n(t_q)) + (F_n(t_q) - F_n^*(t_q))] : n \in \mathbb{N}\}$ is tight and hence the sequence $\{\eta_n^* : n \in \mathbb{N}\}$ is also tight. We show that

$$\psi_{t,n}^* - \eta_n^* \xrightarrow{\mathrm{p}^*} 0. \tag{26}$$

It suffices to show that

$$E\left[E^*\left(\sqrt{n}[F(t_q + t/\sqrt{n}) - F_n^*(t_q + t/\sqrt{n}) - F(t_q) + F_n^*(t_q)]\right)^2\right] \to 0.$$

Its left-hand side is written as

$$\begin{aligned} &nE[F(t_q + t/\sqrt{n}) - F_n(t_q + t/\sqrt{n}) - F(t_q) + F_n(t_q)]^2 \\ &\quad + nE\left[E^*\left\{F_n(t_q + t/\sqrt{n}) - F_n^*(t_q + t/\sqrt{n}) - F_n(t_q) + F_n^*(t_q)\right\}^2\right] \\ &=: \Pi_n^{(1)} + \Pi_n^{(2)}. \end{aligned}$$

In order to show $\Pi_n^{(2)} \to 0$, we observe

$$\Pi_n^{(2)} = E\left[E^*\left\{S_n^*(t_q, t_q + t/\sqrt{n}) - S_n(t_q, t_q + t/\sqrt{n})\right\}^2\right],$$

where $S_n(\cdot,\cdot)$ and $S_n^*(\cdot,\cdot)$ are as in (17) and (18), and by (8)

$$\begin{aligned} &E^*\left\{S_n^*(t_q, t_q + t/\sqrt{n}) - S_n(t_q, t_q + t/\sqrt{n})\right\}^2 \\ &= v_n(t_q, t_q + t/\sqrt{n}) \le K_2(|t/\sqrt{n}|^b + n^{-c}) \end{aligned}$$

which tends to zero as $n \to \infty$. The convergence of $\Pi_n^{(1)}$ to zero can be done similarly due to the weak convergence of the empirical processes of the associated sequence. Therefore, the convergence in (26) holds.

It is easily shown that $\{\zeta_n^* \le t\} = \{\psi_{t,n}^* \le t_n\}$ where $t_n = \frac{\sqrt{n}}{F'(t_q)}[F(t_q + t/\sqrt{n}) - q]$ and that $t_n \to t$ as $n \to \infty$. By these facts and by (26), we have

$$\begin{aligned}\lim_{n\to\infty} P(\zeta_n^* \le t,\ \eta_n^* \ge t+\epsilon) &= \lim_{n\to\infty} P(\psi_{t,n}^* \le t_n,\ \eta_n^* \ge t+\epsilon)\\ &\le \lim_{n\to\infty} P(\psi_{t,n}^* \le t+\epsilon/2,\ \eta_n^* \ge t+\epsilon) = 0.\end{aligned}$$

Thus $\zeta_n^* - \eta_n^* \xrightarrow{\mathrm{p}^*} 0$ by Lemma 5.3, and (25) is obtained.

Now we write, by Lemma 3.1 and by (25),

$$\begin{aligned}t_{n,q}^* - \hat{t}_{n,q} &= (t_{n,q}^* - t_q) - (\hat{t}_{n,q} - t_q)\\ &= \frac{q - F_n^*(t_q)}{F'(t_q)} + \frac{q - F_n(t_q)}{F'(t_q)} + o_p(n^{-1/2})\\ &= \frac{F_n(t_q) - F_n^*(t_q)}{F'(t_q)} + o_p(n^{-1/2})\end{aligned}$$

and thus the desired result in (a) is obtained. Theorem 3.2(b) is straightforwardly given by Theorem 2.1 and Theorem 3.2(a). □

Proof of Theorem 3.3. Noticing that

$$\begin{aligned}&M_n^*(x) - M_n(x)\\ &= \frac{1}{1-F_n^*(x)}\left[-\int_x^\infty (F_n^*(t) - F_n(t))dt + M_n(x)(F_n^*(x) - F_n(x))\right]\end{aligned} \tag{27}$$

we first show

$$\sup_{0<t<\infty} |F_n^*(t) - F_n(t)| \xrightarrow{\text{a.s.}} 0 \tag{28}$$

and

$$\int_0^\infty |F_n^*(t) - F_n(t)|\, dt \xrightarrow{\text{a.s.}} 0 \tag{29}$$

in (conditional) probability P^*.

For (28), we use

$$\begin{aligned}|F_n^*(t) - F_n(t)| &\le \left|\frac{1}{n}\sum_{i=1}^n \mathbb{I}(X_i^* \le t) - \frac{1}{n}\sum_{k=1}^{\kappa}\sum_{i=I_k}^{I_k+L_k-1} \mathbb{I}(X_{ni} \le t)\right|\\ &\quad + \left|\frac{1}{n}\sum_{k=1}^{\kappa}\sum_{i=I_k}^{I_k+L_k-1}\{\mathbb{I}(X_{ni} \le t) - F(t)\}\right| + \left|F_n(t) - \frac{s_\kappa}{n}F(t)\right|\\ &=: J_n^{(1)} + J_n^{(2)} + J_n^{(3)},\end{aligned} \tag{30}$$

where $s_\kappa := L_1 + \cdots + L_\kappa$.

For $J_n^{(1)}$, by the same argument of Step 1 in the proof of Theorem 3.2 of [14], pp. 493–494, we write

$$J_n^{(1)} = \frac{1}{n}\left[\sum_{i=n+1}^{s_\kappa}(\mathbb{I}(X_{ni} \le t) - F_n(t))\right] + \frac{1}{n}(s_\kappa - n)F_n(t)$$

we may show

$$\frac{1}{n}\sum_{j=I}^{I+R-1}(\mathbb{I}(X_{ni} \le t) - F_n(t)) \xrightarrow{\text{a.s.}} 0 \text{ and } \frac{1}{n}R \cdot F_n(t) \xrightarrow{\text{a.s.}} 0,$$

where $R = L_\kappa - R_1$ with $R_1 = n - s_{\kappa-1}$, noting that R, conditionally on $(R_1, s_{\kappa-1})$ has a geometric distribution with mean $1/p$ by the memoryless property. For $\epsilon > 0$,

$$\begin{aligned} &P\left(\frac{1}{n}\left|\sum_{j=I}^{I+R-1}(\mathbb{I}(X_{ni} \le t) - F_n(t))\right| > \epsilon\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{r=1}^{\infty} pq^{r-1} P\left(\frac{1}{n}\left|\sum_{j=i}^{i+r-1}(\mathbb{I}(X_{ni} \le t) - F_n(t))\right| > \epsilon\right) \end{aligned}$$

and

$$\begin{aligned} &P\left(\frac{1}{n}\left|\sum_{j=i}^{i+r-1}(\mathbb{I}(X_{ni} \le t) - F_n(t))\right| > \epsilon\right) \\ &\le \frac{1}{\epsilon^2 n^2} E\left|\sum_{j=i}^{i+r-1}\mathbb{I}(X_{ni} \le t) - rF_n(t)\right|^2 = O(r/n^2). \end{aligned}$$

Thus, using $\sum_{r=1}^{\infty} q^{r-1} r^a = O(1/p^{a+1})$, $a \ge 1$,

$$P\left(\frac{1}{n}\left|\sum_{j=I}^{I+R-1}(\mathbb{I}(X_{ni} \le t) - F(t))\right| > \epsilon\right) = O(1/(n^2 p)) = O(n^{-2+\varrho})$$

for $1 < 2 - \varrho < 3/2$, by (A2): $p \sim n^{-\varrho}$ for $\varrho \in (\frac{1}{2}, 1)$.

Also, similarly we have

$$\begin{aligned} P\left(\frac{1}{n}R \cdot F_n(t) > \epsilon\right) &= \sum_{r=1}^{\infty} pq^{r-1} P\left(\frac{1}{n} rF_n(t) > \epsilon\right) \\ &\le \sum_{r=1}^{\infty} pq^{r-1}\frac{r^2}{\epsilon^2 n^2} Var(F_n(t)) \\ &= O(1/(n^3 p^2)) = O(n^{-3+2\varrho}),\ 1 < 3 - 2\varrho < 3. \end{aligned}$$

Therefore, we have

$$\sum_{n=1}^{\infty} P^*(J_n^{(1)} > \epsilon) \le \sum_{n=1}^{\infty} O(n^{-2+\varrho}) + \sum_{n=1}^{\infty} O(n^{-3+2\varrho}) < \infty.$$

By Borel-Cantelli Lemma, the desired almost sure convergence of $J_n^{(1)}$ holds.

For $J_n^{(2)}$, we consider a sequence $m = m_n$ with $m \to \infty$ and $m/(np) \to 1$ as $n \to \infty$. For $k = 1, 2, \ldots, m$, let $H_k := \sum_{i=I_k}^{I_k+L_k-1} \{\mathbb{I}(X_{ni} \le t) - F(t)\}$. Note that $\{H_k,\, k = 1, \ldots, m\}$ are independent and $H_k = H_{I_k,L_k}(-\infty, t)$ where $H_{I_k,L_k}(\cdot,\cdot)$ is given as in (5). For $\epsilon > 0$ and $\delta > 0$

$$\begin{aligned} & P\left(\left|\frac{1}{n}\sum_{k=1}^{m}\sum_{i=I_k}^{I_k+L_k-1}\{\mathbb{I}(X_{ni} \le t) - F(t)\}\right| > \epsilon\right) \\ & = P\left(\left|\frac{1}{n}\sum_{k=1}^{m} H_k\right| > \epsilon\right) \\ & \le \frac{1}{\epsilon^{2+\delta} n^{2+\delta}} E\left|\sum_{k=1}^{m} H_k\right|^{2+\delta} \le \frac{m^{1+\delta/2}}{\epsilon^{2+\delta} n^{2+\delta}} E\left|H_{I,L}(-\infty, t)\right|^{2+\delta}. \end{aligned} \tag{31}$$

Using $1 = 1/(pL) + o(1)$ a.s. and $\tilde{H} := H_{I,L}(-\infty, t) = O(1/\sqrt{p})$ a.s.,

$$\tilde{H} = \frac{1}{pL}\tilde{H} + o(\tilde{H}) = \frac{1}{pL}\tilde{H} + o(1/\sqrt{p}) = \frac{1}{p}\left[\frac{1}{L}\tilde{H} + o(\sqrt{p})\right] \text{ a.s.,}$$

$$E|\tilde{H}|^{2+\delta} = \frac{1}{p^{2+\delta}} E\left|\frac{1}{L}\tilde{H} + o(\sqrt{p})\right|^{2+\delta} = \frac{1}{p^{1+\delta/2}} E\left|Z_0 + O(\sqrt{p}) + o(1)\right|^{2+\delta}$$

by Lemma 2.3(a), where Z_0 is a normal random variable with mean zero and variance $\sigma_0^2(t) = \lim_{n\to\infty} \frac{1}{n} Var\left(\sum_{i=1}^{n} \mathbb{I}(X_i \le t)\right)$. Thus the right-hand side of (31) is $O((m/n^2p)^{1+\delta/2}) = O(n^{-1-\delta/2})$. Hence $\sum_{n=1}^{\infty} P(J_n^{(2)} > \epsilon) = \sum_{n=1}^{\infty} O(n^{-1-\delta/2}) < \infty$. By Borel-Cantelli Lemma, the desired almost sure convergence of $J_n^{(2)}$ holds.

For $J_n^{(3)} = \left|F_n(t) - \frac{s_\kappa}{n}F(t)\right|$ which is less than or equal to $|F_n(t) - F(t)| + \left|F(t) - \frac{s_\kappa}{n}F(t)\right|$. By Theorem 2.1 of [37], we have $\sup_x |F_n(t) - F(t)| \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$ under condition (A4). Also, we may write $\left|F(t) - \frac{s_\kappa}{n}F(t)\right| = (s_\kappa - n)/nF(t) = (R/n)F(t)$ which converges to zero a.s.. We complete the proof of (28).

For (29), let $I_n^* := \int_0^\infty |F_n^*(t) - F_n(t)|\, dt$. Since X_i^* is nonnegative and $E^*X_1^* = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} EX < \infty$ by the SLLN, and noting that $P^*(X_1^* > t) = E[E^*(\mathbb{I}(X_1^* > t)|I] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(X_i > t) = 1 - F_n(t)$, we may have $E^*X_1^* = \int_0^\infty P^*(X_1^* > t)dt = \int_0^\infty (1 - F_n(t))\, dt < \infty$. Hence, for arbitrary small $\epsilon > 0$ we can choose $\beta_n > 0$, which is depending on $F_n(\cdot)$, so large that $\int_{\beta_n}^\infty (1 - F_n(t))\, dt < \epsilon/2$. Let $\beta = \limsup_{n\to\infty} \beta_n$. then we have $\int_\beta^\infty (1 - F_n(t))\, dt < \epsilon/2$ for sufficiently large n.

Let $A_n^{(1)} := \int_0^\beta |F_n^*(t) - F_n(t)|\,dt$, $A_n^{(2)} := \int_\beta^\infty |1 - F_n^*(t)|\,dt$ and $A_n^{(3)} := \int_\beta^\infty |1 - F_n(t)|\,dt$. Then we have $I_n^* \le A_n^{(1)} + A_n^{(2)} + A_n^{(3)}$ and $A_n^{(3)} < \epsilon/2$. Observe that, by (28)

$$A_n^{(1)} = \int_0^\beta |F_n^*(t) - F_n(t)|\,dt \le \beta \sup_{0<t<\infty} |F_n^*(t) - F_n(t)| \xrightarrow{\text{a.s.}} 0.$$

For $A_n^{(2)} = \int_\beta^\infty (1 - F_n^*(t))\,dt = \frac{1}{n}\sum_{i=1}^n \int_\beta^\infty \mathbb{I}(X_i^* > t)\,dt$, let $h_i^*(\beta) = \int_\beta^\infty \mathbb{I}(X_i^* > t)\,dt$, $h_{ni}(\beta) = \int_\beta^\infty \mathbb{I}(X_{ni} > t)\,dt$ and $\bar{h}(\beta) = \frac{1}{n}\sum_{i=1}^n h_{ni}(\beta)$. We have

$$A_n^{(2)} = \frac{1}{n}\sum_{i=1}^n h_i^*(\beta) = \frac{1}{n}\sum_{i=1}^n [h_i^*(\beta) - \bar{h}(\beta)] + \bar{h}(\beta).$$

Note that $\bar{h}(\beta) = \frac{1}{n}\sum_{i=1}^n h_{ni}(\beta) = \frac{1}{n}\sum_{i=1}^n \int_\beta^\infty \mathbb{I}(X_{ni} > t)\,dt = A_n^{(3)}$, and thus $\bar{h}(\beta) < \epsilon/2$. Thus, finally it suffices to show that $\frac{1}{n}\sum_{i=1}^n [h_i^*(\beta) - \bar{h}(\beta)] \xrightarrow{\text{a.s.}} 0$. Its proof follows the same arguments as that of (28). In the proof of (28) with $J_n^{(1)}, J_n^{(2)}$ and $J_n^{(3)}$ above in (30), $\mathbb{I}(X_i^* \le t)$ and $\mathbb{I}(X_{ni} \le t)$ are replaced by $h_i^*(\beta)$ and $h_{ni}(\beta)$, respectively. In other words, $\frac{1}{n}\sum_{i=1}^n [h_i^*(\beta) - \bar{h}(\beta)]$ is decomposed by three terms as follows and it is bounded above by

$$\left|\frac{1}{n}\sum_{i=1}^n h_i^*(\beta) - \frac{1}{n}\sum_{k=1}^{\kappa}\sum_{i=I_k}^{I_k+L_k-1} h_{ni}(\beta)\right| + \left|\frac{1}{n}\sum_{k=1}^{\kappa}\sum_{i=I_k}^{I_k+L_k-1}\{h_{ni}(\beta) - E[h_{n1}(\beta)]\}\right|$$
$$+ \left|\bar{h}(\beta) - \frac{s_\kappa}{n}E[h_{n1}(\beta)]\right| =: J_{h,n}^{(1)} + J_{h,n}^{(2)} + J_{h,n}^{(3)}$$

and thus in the same way as above, the almost sure convergences of $J_{h,n}^{(1)}$, $J_{h,n}^{(2)}$ and $J_{h,n}^{(3)}$ can be obtained. Therefore, $\limsup_{n\to\infty} I_n^* < \epsilon$ a.s. for all small $\epsilon > 0$ and we complete the proof of (29).

If $x < T = \min\{T_F, T_{F_n}\}$, then $F_n(x) \ne 1, F_n(x) < 1$, that is, for some $i_0 \in \{1, 2, \ldots, n\}$, $\mathbb{I}(X_{i_0} \le x) = 0$, $X_{i_0} > x$. Also, $P^*(X_1^* = X_{i_0}) = \frac{1}{n}$ and thus $F_n^*(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(X_i^* \le x) \ne 1$ a.s. since $\mathbb{I}(X_i^* \le x) = \sum_{j=1}^n \mathbb{I}(X_i^* = X_j)\mathbb{I}(X_j \le x)$, which is zero with probability $\frac{1}{n}$. By (28) and (29) along with $1 - F_n^*(x)$ for $x < T$, the result of (a) holds.

For (b), first note that $\alpha_n^*(F(t)) = \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^n \mathbb{I}(U_i^* \le F(t)) - F(t)\right] = G_n^*(t)$ and we write, from (27) and by change of variables,

$$\begin{aligned}
W_n^*(x) &= \frac{1}{1 - F_n^*(x)}\left[-\int_x^\infty G_n^*(t)dt + M_n(x)G_n^*(x)\right] \\
&= \frac{1}{1 - F_n^*(x)}\left[-\int_x^\infty \alpha_n^*(F(t))dt + M_n(x)\alpha_n^*(F(x))\right] \\
&= \frac{1}{1 - F_n^*(x)}\left[-\int_{F(x)}^1 \alpha_n^*(s)dQ(s) + M_n(x)\alpha_n^*(F(x))\right].
\end{aligned}$$

By the result in (a), by Theorem 3.1 of [35], by Theorem 2.1 of [37], and by (10) and (11), the last expression converges to $W(x)$ in (12) in distribution, and the proof of (b) is completed. $\square$

5. Appendix

This appendix presents some existing lemmas required in proving our results.

Lemma 5.1 ([15])**.** *Let $\{L = L(n) : n = 1, 2, \dots\}$ be a sequence of geometrically distributed random variables with parameter $p = p_n$, satisfying condition* (A2) *on the rate of p, and let $\{\varepsilon_t : t = 1, 2, \dots\}$ be a strictly stationary and associated sequence of random variables satisfying condition* (A1)*. Then, as $n \to \infty$, we have*

$$P\left(|p_n L(n) - 1| > \delta_n\right) \to 0$$

for any nonincreasing sequence δ_n of positive numbers tending to 0*, and*

$$\frac{1}{\sqrt{L(n)}} \sum_{t=1}^{L(n)} \varepsilon_t \xrightarrow{\mathrm{d}} N(0, \sigma_\varepsilon^2),$$

where $\sigma_\varepsilon^2 = \lim_{n\to\infty} Var(\sum_{t=1}^{n} \varepsilon_t)/n < \infty$.

As mentioned above, Lemma 5.1 is a modification of the random central limit theorem of [15] as a special case with a geometrically distributed random sample size. Its proof can be given straightforwardly from Assumption 2(b) and Theorem 3.1 of [15].

Lemma 5.2 ([25])**.** *Assume that $\{X_i\}$ is a stationary sequence of random variables uniformly distributed on* $[0, 1]$*, with the condition* (4)*. Then for every $0 \le s < t \le 1$ and every $n \ge 1$, there exist a positive constant $K > 0$, $0 < b < 1$ and $c > 0$ such that*

$$\frac{1}{n} Var\left(\sum_{i=1}^{n} \mathbb{I}(s < X_i \le t)\right) \le K(|t - s|^b + n^{-c}).$$

See Eq. (4.15) in Lemma 4.3 and its proof of [25], p. 893.

Lemma 5.3 ([12])**.** *Let $\{\zeta_n, n \in \mathbb{N}\}$ and $\{\eta_n, n \in \mathbb{N}\}$ be two sequences of random variables such that* (i) *the sequence $\{\eta_n, n \in \mathbb{N}\}$ is tight, and* (ii) *for a real number t, and $\epsilon > 0$*

$$\lim_{n\to\infty} P(\zeta_n \le t,\ \eta_n \ge t + \epsilon) = 0 \quad \textit{and} \quad \lim_{n\to\infty} P(\zeta_n \ge t + \epsilon,\ \eta_n \le t) = 0.$$

Then we have $\zeta_n - \eta_n \xrightarrow{\mathrm{p}} 0$ as $n \to \infty$.

References

[1] M. Asadi and A. Berred, *Properties and estimation of the mean past lifetime*, Statistics **46** (2012), no. 3, 405–417. https://doi.org/10.1080/02331888.2010.540666

[2] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York, 1968.

[3] P. Bühlmann, *Blockwise bootstrapped empirical process for stationary sequences*, Ann. Statist. **22** (1994), no. 2, 995–1012. https://doi.org/10.1214/aos/1176325508

[4] ———, *The blockwise bootstrap for general empirical processes of stationary sequences*, Stochastic Process. Appl. **58** (1995), no. 2, 247–265. https://doi.org/10.1016/0304-4149(95)00019-4

[5] A. Bulinski and A. Shashkin, *Limit theorems for associated random fields and related systems*, Advanced Series on Statistical Science & Applied Probability, 10, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007. https://doi.org/10.1142/9789812709417

[6] P. Doukhan, *Stochastic models for time series*, Mathématiques & Applications (Berlin), 80, Springer, Cham, 2018. https://doi.org/10.1007/978-3-319-76938-7

[7] P. Doukhan, G. Lang, A. Leucht, and M. H. Neumann, *Dependent wild bootstrap for the empirical process*, J. Time Series Anal. **36** (2015), no. 3, 290–314. https://doi.org/10.1111/jtsa.12106

[8] P. Doukhan and S. Louhichi, *A new weak dependence condition and applications to moment inequalities*, Stochastic Process. Appl. **84** (1999), no. 2, 313–342. https://doi.org/10.1016/S0304-4149(99)00055-1

[9] P. Doukhan and D. Surgailis, *Functional central limit theorem for the empirical process of short memory linear processes*, C. R. Acad. Sci. Paris Sér. I Math. **326** (1998), no. 1, 87–92. https://doi.org/10.1016/S0764-4442(97)82718-8

[10] F. El Ktaibi, B. G. Ivanoff, and N. C. Weber, *Bootstrapping the empirical distribution of a linear process*, Statist. Probab. Lett. **93** (2014), 134–142. https://doi.org/10.1016/j.spl.2014.06.019

[11] F. El Ktaibi and B. G. Ivanoff, *Bootstrapping the empirical distribution of a stationary process with change-point*, Electron. J. Stat. **13** (2019), no. 2, 3572–3612. https://doi.org/10.1214/19-ejs1613

[12] J. K. Ghosh, *A new proof of the Bahadur representation of quantiles and an application*, Ann. Math. Statist. **42** (1971), 1957–1961. https://doi.org/10.1214/aoms/1177693063

[13] E. Hwang and D. W. Shin, *Stationary bootstrapping for non-parametric estimator of nonlinear autoregressive model*, J. Time Series Anal. **32** (2011), no. 3, 292–303. https://doi.org/10.1111/j.1467-9892.2010.00699.x

[14] ———, *Strong consistency of the stationary bootstrap under ψ-weak dependence*, Stat. Probab. Lett. **82** (2012), 488–495.

[15] ———, *Random central limit theorems for linear processes with weakly dependent innovations*, J. Korean Stat. Soc. **41** (2012), 313–322.

[16] ———, *Stationary bootstrapping realized volatility under market microstructure noise*, Electron. J. Stat. **7** (2013), 2032–2053.

[17] ———, *Stationary bootstrapping realized volatility*, Stat. Probab. Lett. **83** (2013), 2045–2051.

[18] ———, *Stationary bootstrapping for semiparametric panel unit root tests*, Comput. Statist. Data Anal. **83** (2015), 14–25. https://doi.org/10.1016/j.csda.2014.09.004

[19] H. R. Künsch, *The jackknife and the bootstrap for general stationary observations*, Ann. Statist. **17** (1989), no. 3, 1217–1241. https://doi.org/10.1214/aos/1176347265

[20] S. Lee, *Random central limit theorem for the linear process generated by a strong mixing process*, Statist. Probab. Lett. **35** (1997), no. 2, 189–196. https://doi.org/10.1016/S0167-7152(97)00013-8

[21] S. Louhichi, *Weak convergence for empirical processes of associated sequences*, Ann. Inst. H. Poincaré Probab. Statist. **36** (2000), no. 5, 547–567. https://doi.org/10.1016/S0246-0203(00)00140-0

[22] U. V. Naik-Nimbalkar and M. B. Rajarshi, *Validity of blockwise bootstrap for empirical processes with stationary observations*, Ann. Statist. **22** (1994), no. 2, 980–994. https://doi.org/10.1214/aos/1176325507

[23] M. S. Noughabi and M. Kayid, *Bivariate quantile residual life: a characterization theorem and statistical properties*, Statist. Papers **60** (2019), no. 6, 2001–2012. https://doi.org/10.1007/s00362-017-0905-9

[24] A. Parvardeh, *A note on the asymptotic distribution of the estimation of the mean past lifetime*, Statist. Papers **56** (2015), no. 1, 205–215. https://doi.org/10.1007/s00362-013-0575-1

[25] M. Peligrad, *On the blockwise bootstrap for empirical processes for stationary sequences*, Ann. Probab. **26** (1998), no. 2, 877–901. https://doi.org/10.1214/aop/1022855654

[26] H. Pishro-Nik, *Introduction to Probability Statistics and Random Processes*, Kappa Research. LLC. www.probabilitycourse.com, 2016.

[27] D. N. Politis and J. P. Romano, *The stationary bootstrap*, J. Amer. Statist. Assoc. **89** (1994), no. 428, 1303–1313.

[28] ______, *Limit theorems for weakly dependent Hilbert space valued random variables with applications to the stationary bootstrap*, Stat. Sin. **4** (1994), 461–476.

[29] D. Pollard, *Convergence of Stochastic Processes*, Springer Series in Statistics, Springer-Verlag, New York, 1984. https://doi.org/10.1007/978-1-4612-5254-2

[30] D. Radulović, *The bootstrap for empirical processes based on stationary observations*, Stochastic Process. Appl. **65** (1996), no. 2, 259–279. https://doi.org/10.1016/S0304-4149(96)00102-0

[31] ______, *The bootstrap of empirical processes for α-mixing sequences*, in High Dimensional Probability (E. Eberlein, M. Hahn and M. Talagrand, eds.), 315–330, Birkhäuser, Basel, 1998.

[32] G. G. Roussas, *Kernel estimates under association: strong uniform consistency*, Statist. Probab. Lett. **12** (1991), no. 5, 393–403. https://doi.org/10.1016/0167-7152(91)90028-P

[33] X. Shao, *The dependent wild bootstrap*, J. Amer. Statist. Assoc. **105** (2010), no. 489, 218–235. https://doi.org/10.1198/jasa.2009.tm08744

[34] Q. M. Shao and H. Yu, *Bootstrapping the sample means for stationary mixing sequences*, Stochastic Process. Appl. **48** (1993), no. 1, 175–190. https://doi.org/10.1016/0304-4149(93)90113-I

[35] ______, *Weak convergence for weighted empirical processes of dependent sequences*, Ann. Probab. **24** (1996), no. 4, 2098–2127. https://doi.org/10.1214/aop/1041903220

[36] B. Wieczorek, *Blockwise bootstrap of the estimated empirical process based on ψ-weakly dependent observations*, Stat. Inference Stoch. Process. **19** (2016), no. 1, 111–129. https://doi.org/10.1007/s11203-015-9120-2

[37] H. Yu, *A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated sequences*, Probab. Theory Related Fields **95** (1993), no. 3, 357–370. https://doi.org/10.1007/BF01192169

Eunju Hwang
Department of Applied Statistics
Gachon University
Seongnam 13120, Korea
Email address: ehwang@gachon.ac.kr