Теорія Ймовір. та Матем. Статист.
Вип. 96, 2017

Theor. Probability and Math. Statist.
No. 96, 2018, Pages 15–26
https://doi.org/10.1090/tpms/1031
Article electronically published on October 5, 2018

THE UNIFORM CLT FOR EMPIRICAL ESTIMATOR OF A GENERAL STATE SPACE SEMI-MARKOV KERNEL INDEXED BY FUNCTIONS

UDC 519.21

S. BOUZEBDA AND N. LIMNIOS

Abstract. In this paper we mainly deal with the uniform CLT for empirical estimator of a general state space semi-Markov process indexed by functions under the uniformly integrable entropy condition. A way to describe the uniform CLT is to translate the problem into martingale difference sequences to obtain the desired results.

1. Introduction and notation

Semi-Markov processes are an extension of jump Markov processes and renewal processes. More specifically, they allow the use of any distribution for the sojourn times instead of the exponential (geometric) distributions in the Markov processes (chains) case. This feature has led to successful applications in survival analysis [1], reliability [21], queuing theory, finance and insurance [28]. The basic theory of the semi-Markov processes is given by [25, 26] and in the particular case of renewal process we refer to [9]. For recent references in this area along with statistical applications see, e.g., [21]. Nonparametric estimation of semi-Markov kernel and conditional transition probability has been the subject of intense investigation for many years leading to the development of a large variety of methods; see, e.g., [7, 8, 20] and the references therein.

At the same time, we have very important advances on empirical processes theory and their applications in several directions in statistical theory and practice, e.g., kernel-type estimation, copula, M-estimation, rate of statistical estimators, etc.; we may refer to [11, 18, 19, 23, 24, 30, 33]. In the present paper we would like to connect some of these theoretical advances to semi-Markov estimation problems, first of all to the semi-Markov kernel estimation, which is the basis for further estimation problems. In particular we extend previous results ([20] and [8]) to a more abstract setting and we consider a general state space semi-Markov process in the present work. In order to obtain our results we will transfer the problem to the martingale differences setting that permits us to circumvent the use of the delicate chaining and symmetrization arguments. This simplifies our proof by using the uniform CLT of [2] and [3].

We start by giving some notation and definitions that are needed for the forthcoming sections. All the following random processes will be supposed to be defined on a complete probability space $(\Omega, \mathcal{F}, \mathsf{P})$. This is a technical requirement which allows for the construction of the Gaussian processes in our theorems and is not restrictive since one can expand a probability space to make it rich enough (see, e.g., Appendix 2 in [10]

2010 *Mathematics Subject Classification.* Primary 60J28, 60K15, 60G09, 62F40.
Key words and phrases. Semi-Markov process, semi-Markov kernel, empirical estimator, uniform central limit theorem, invariance principle.

and [5, Lemma A1]). To define semi-Markov processes or equivalently Markov renewal processes, it is natural, first, to define semi-Markov kernels (see, for example, [21] for a description in detail). For convergence of stochastic processes in a functional setting, see, for example, [15, 31]. We are using the empirical processes terminology concerning the uniform CLT, which of course is a functional CLT.

Let $\mathbb{R}_+ = [0, \infty)$ be a set of nonnegative real numbers and let $\mathbb{N}$ be a set of natural numbers $\{0, 1, 2, \dots\}$.

Definition 1.1. Let $(E, \mathscr{E})$ be a Borel measurable space. Let $\mathscr{B}_+$ denote the Borel σ-algebra of subsets of $\mathbb{R}_+$. Let $\{P((x, s), A \times \Gamma) \colon (x, s) \in E \times \mathbb{R}_+, A \in \mathscr{E}, \Gamma \in \mathscr{B}_+\}$ be a Markov transition function on $(E \times \mathbb{R}_+, \mathscr{E} \times \mathbb{R}_+)$. The function Q defined by

$$Q\big(x, A \times (\Gamma - s)\big) = P\big((x, s), A \times \Gamma\big), \qquad (x, s) \in E \times \mathbb{R}_+,\ A \in \mathscr{E},\ \Gamma \in \mathbb{B}_+,$$

where $\Gamma - s = \{t \in \mathbb{R}_+ \colon t + s \in \Gamma\}$, is a semi-Markov kernel.

For any $(x, s) \in E \times \mathbb{R}_+$, there exists a probability measure $\mathbb{P}_{(x,s)}$ on $(\Omega, \mathcal{F})$ and a sequence of random variables $\{J_n, S_n \colon n \in \mathbb{N}\}$, such that

$$\mathbb{P}_{(x,s)}(J_0 = x, S_0 = s) = 1$$

and, for $n > 0$, $A \in \mathscr{E}$, and $\Gamma \in \mathscr{B}_+$,

$$\begin{aligned}
&\mathbb{P}_{(x,s)}\big((J_{n+1}, S_{n+1}) \in A \times \Gamma \mid \sigma(J_k, S_k; k \le n)\big) \\
&\quad = \mathbb{P}_{(x,s)}\big((J_{n+1}, S_{n+1}) \in A \times \Gamma \mid \sigma(J_n, S_n)\big) \\
&\quad = Q\big(J_n, A \times (\Gamma - S_n)\big).
\end{aligned}$$

Let $\{N(t) \colon t \in \mathbb{R}_+\}$ be defined by

$$N(t) = \sup\{n \ge 0 \colon S_n \le t\}.$$

The probability measure $\mathbb{P}_{(x,s)}$, for any $(x, s) \in E \times \mathbb{R}_+$, is uniquely defined in the usual way for the Markov case on cylinder sets on the canonical measurable space $(\Omega, \mathcal{F}) = \big(E^{\mathbb{N}}, \mathscr{E}^{\otimes \mathbb{N}}\big)$; see, e.g., [21].

Definition 1.2. The stochastic process $\{J_n, S_n \colon n \in \mathbb{N}\}$ is called a Markov renewal process. The stochastic process $\mathbf{Z} = \{Z(t) \colon t \in \mathbb{R}_+\}$, defined by $Z(t) = J_{N(t)}$ for $t \ge 0$ (or $J_n = Z(S_n)$ for $n \ge 0$), is a semi-Markov process associated with (J_n, S_n).

To be exact, $\mathbf{Z}$ is an $(E, \mathscr{E})$-valued càdlàg homogeneous semi-Markov process. The process $J = \{J_n \colon n \in \mathbb{N}\}$ (called the embedded Markov chain of $\mathbf{Z}$) is a Markov chain with state space $(E, \mathscr{E})$ and transition probability kernel

$$P(x, dy) = Q(x, dy \times \mathbb{R}_+).$$

The process (S_n) is the sequence of jump times of $\mathbf{Z}$, with

$$0 \le S_0 < S_1 < \cdots < S_n < S_{n+1} < \dots,$$

and inter-jump times $X_n = S_n - S_{n-1}$, for $n \ge 1$. The process $\{J_n, X_n \colon n \in \mathbb{N}\}$ is a Markov chain with state space $(E \times \mathbb{R}_+, \mathscr{E} \times \mathscr{B}_+)$ and transition probability kernel $Q(x, dy \times dt)$. The point process $\{N(t) \colon t \in \mathbb{R}_+\}$ counts the jumps in the time interval $(0, t]$. Let H denote the distribution function of the sojourn times, that is,

$$H(x, t) = Q(x, E \times [0, t]) \quad \text{for } (x, t) \in E \times \mathbb{R}_+,$$

and set $\overline{H} = 1 - H$. Let us define the transition operator Q, for $f \colon E \times E \times \mathbb{R}_+ \to \mathbb{R}_+$ measurable, as follows:

$$Qf(x) = \iint_{E \times \mathbb{R}_+} f(x, y, s)\, Q(x, dy, ds).$$

Also let $m_k(x)$ denote the k-th moment of the sojourn time in state $x \in E$, that is,

$$\begin{aligned} m_k(x) &= \mathsf{E}\left(X_{n+1}^k \mid J_n = x\right) \\ &= \int_{\mathbb{R}_+} s^k\, Q(x, E \times ds) = \int_{\mathbb{R}_+} s^k\, H(x, ds), \qquad k \in \mathbb{N}^*,\ x \in E; \end{aligned}$$

we set $m(x) = m_1(x)$ for the mean sojourn time in state $x \in E$. We will also need the following definitions. Let $\mathcal{P}(E)$ be the set of all probability distributions α on $(E, \mathscr{E})$; note that α will denote both the distribution and its density when it exists.

Definition 1.3. Let α be a probability measure in $\mathcal{P}(E)$. The probability measure $\mathbb{P}_\alpha$ is defined on $(\Omega, \mathcal{F})$ by

$$\mathbb{P}_\alpha(C) = \int_E \alpha(dx)\, \mathbb{P}_x(C), \qquad C \in \mathcal{F},$$

where

$$\mathbb{P}_x(C) = \mathbb{P}(C \mid J_0 = x).$$

Let β be a probability measure on $(E \times \mathbb{R}_+, \mathscr{E} \times \mathscr{B}_+)$. The probability measure $\mathbb{P}_\beta$ is defined on $(\Omega, \mathcal{F})$ by

$$\mathbb{P}_\beta(C) = \iint_{E \times \mathbb{R}_+} \beta(dx \times ds)\, \mathbb{P}_{(x,s)}(C), \qquad C \in \mathcal{F},$$

where

$$\mathbb{P}_{(x,s)}(C) = \mathbb{P}(C \mid J_0 = x, S_0 = s).$$

Let $L^1(\beta Q)$ be the space of all real βQ-measurable integrable functions g defined on $E \times E \times \mathbb{R}_+$. The functional μQ is defined on $L^1(\mu Q)$ by

$$\mu Q g = \iiint_{E \times E \times \mathbb{R}_+} \mu(dx)\, Q(x, dy \times ds) g(x, y, s), \qquad g \in L^1(\mu Q).$$

We set $\widehat{g} = \nu Q g$, and if g is a function of only one variable $x \in E$, we put

$$\mu g = \mu Q g = \int \mu(dx)\, g(x).$$

We consider here regular semi-Markov processes, that is, $\mathbb{P}_x(N(t) < +\infty) = 1$, for any $t \geq 0$ and $x \in E$. Moreover, the distribution functions of the sojourn times are not degenerate, i.e., δ_0, Dirac distribution concentrated at 0. It is worth noticing that Dirac distributions concentrated at the point $a > 0$ are allowed; these distributions are useful in order to introduce fixed durations.

All the asymptotic results in this paper require the following assumptions.

A.1. The Markov chain $\{J_k \colon k \geq 0\}$ possesses a unique stationary distribution ν;
A.2. $\overline{m} = \int_E m(x)\, \nu(dx) < \infty$.

Conditions A.1–A.2 imply that the semi-Markov process $\{Z_t \colon t \geq 0\}$ has a unique stationary distribution, π, given by

$$\pi(B) = \frac{1}{\overline{m}} \int_B m(x)\, \nu(dx), \quad \text{for } B \in \mathscr{E}.$$

The structure of the present paper is as follows. The uniform central limit theorem for a function-indexed semi-Markov process, under the uniformly integrable entropy condition, is obtained in Section 2. Finally, detailed mathematical developments are given in Section 3.

2. The uniform CLT for empirical estimator of a semi-Markov kernel

2.1. **Preliminaries and notation.** Given a class $\mathscr{F}$ of measurable functions defined on a measurable space $(\mathbf{X}, \mathcal{X}) := (E \times E \times \mathbb{R}_+, \mathscr{E} \times \mathscr{E} \times \mathscr{B}_+)$, the covering number $N(\varepsilon, \mathscr{F}, \|\cdot\|)$, simply denoted $N(\varepsilon)$ when there is no risk of ambiguity, is the minimum number of balls $\{g\colon \|g-h\| < \varepsilon\}$ of radius ε needed to cover $\mathscr{F}$. Let $\mathbf{F}$ be an envelope of $\mathscr{F}$. That is, $\mathbf{F}$ is a measurable function from $\mathbf{X}$ to $[0,\infty)$ such that

$$\sup_{f\in\mathscr{F}} |f(x)| \le \mathbf{F}(x) \quad \text{for all } x \in \mathbf{X}.$$

Let $M(\mathbf{X}, \mathbf{F})$ be the set of all measures γ on $(\mathbf{X}, \mathcal{X})$ with

$$\gamma\left(\mathbf{F}^2\right) := \int_{\mathbf{X}} \mathbf{F}^2 \, d\gamma < \infty.$$

Given a measure γ on $(\mathbf{X}, \mathcal{X})$, we define

$$d_\gamma^{(2)}(f,g) := \left[\gamma(f-g)^2\right]^{1/2} = \left[\int_{\mathbf{X}} (f-g)^2 \, d\gamma\right]^{1/2}.$$

Say that $\mathscr{F}$ has a uniformly integrable entropy with respect to L_2-norm if

$$\int_0^\infty \sup_{\gamma\in M(\mathbf{X},\mathbf{F})} \left[\log N\left(\varepsilon\left[\gamma\left(\mathbf{F}^2\right)\right]^{1/2}, \mathscr{F}, d_\gamma^{(2)}\right)\right]^{1/2} d\varepsilon < \infty.$$

When the class $\mathscr{F}$ has uniformly integrable entropy, $\left(\mathscr{F}, d_\gamma^{(2)}\right)$ is totally bounded for any measure γ. Many important classes of functions, such as VC graph classes, have uniformly integrable entropy. See Section 2.6 of [33], and we may refer also to [18].

Example 2.1. The set $\mathscr{F}$ of all indicator functions $\mathbb{1}_{(-\infty,t]}$ of cells in $\mathbb{R}$ satisfies

$$N\left(\varepsilon, \mathscr{F}, d_{\mathbb{P}}^{(2)}\right) \le \frac{2}{\varepsilon^2},$$

for any probability measure $\mathbb{P}$ and $\varepsilon \le 1$. Notice that

$$\int_0^1 \sqrt{\log\left(\frac{1}{\varepsilon}\right)}\, d\varepsilon \le \int_0^\infty u^{1/2}\exp(-u)\, du \le 1.$$

For more details and discussion on this example refer to Example 2.5.4 of [33] and [18, p. 157]. The covering numbers of the class of cells $(-\infty, t]$ in higher dimension satisfy a similar bound, but with higher power of $(1/\varepsilon)$; see Theorem 9.19 of [18].

Example 2.2 (Classes of functions that are Lipschitz in a parameter, Section 2.7.4 in [33])**.** Let $\mathscr{F}$ be the class of functions $x \mapsto f(t,x)$ that are Lipschitz in the index parameter $t \in T$. Suppose that

$$|f(t_1,x) - f(t_2,x)| \le d(t_1,t_2)\kappa(x)$$

for some metric d on the index set T, the function $\kappa(\cdot)$ defined on the sample space $\mathcal{X}$, and all x. According to Theorem 2.7.11 of [33] and Lemma 9.18 of [18], it follows, for any norm $\|\cdot\|_{\mathcal{F}}$ on $\mathcal{F}$, that

$$N\big(\varepsilon\|F\|_{\mathscr{F}}, \mathscr{F}, \|\cdot\|_{\mathscr{F}}\big) \le N(\varepsilon/2, T, d).$$

Hence if (T,d) satisfies $J(\infty, T, d) < \infty$, then the conclusion holds for $\mathscr{F}$.

Example 2.3. Let us consider as an example the classes of functions that are smooth up to order α; see Section 2.7.1 of [33] and Section 2 of [32]. For $0 < \alpha < \infty$ let $\lfloor\alpha\rfloor$ be the greatest integer strictly smaller than α. For any vector $k = (k_1, \ldots, k_d)$ of d integers define the differential operator

$$D^{k.} = \frac{\partial^{k.}}{\partial^{k_1} \ldots \partial^{k_d}},$$

where

$$k. = \sum_{i=1}^{d} k_i \, .$$

Then, for a function $f\colon \mathcal{X} \to \mathbb{R}$, let

$$\|f\|_\alpha = \max_{k. \le \lfloor \alpha \rfloor} \sup_x \left| D^k f(x) \right| + \max_{k. \le \lfloor \alpha \rfloor} \sup_x \frac{D^k f(x) - D^k f(y)}{\|x-y\|^{\alpha - \lfloor \alpha \rfloor}},$$

where the suprema are taken over all x, y in the interior of $\mathcal{X}$ with $x \neq y$. Let $C_M^\alpha(\mathcal{X})$ be the set of all continuous functions $f\colon \mathcal{X} \to \mathbb{R}$ with

$$\|f\|_\alpha \le M.$$

Note that for $\alpha \le 1$ this class consists of bounded functions f that satisfy the Lipschitz condition. [16] computed the entropy of the classes of $C_M^\alpha(\mathcal{X})$ for the uniform norm. As a consequence of their results [32] shows that there exists a constant K depending only on α, d, and the diameter of $\mathcal{X}$ such that for every measure γ and every $\varepsilon > 0$,

$$\log N_{[\,]}\big(\varepsilon M \gamma(\mathcal{X}), C_M^\alpha(\mathcal{X}), L_2(\gamma)\big) \le K \left(\frac{1}{\varepsilon}\right)^{d/\alpha};$$

$N_{[\,]}$ is the bracketing number. Refer to Definition 2.4, and we refer to Theorem 2.7.1 of [33] for a variant of the last inequality. By Lemma 9.18 of [18], we have

$$\log N\big(\varepsilon M \gamma(\mathcal{X}), C_M^\alpha(\mathcal{X}), L_2(\gamma)\big) \le K \left(\frac{1}{2\varepsilon}\right)^{d/\alpha}.$$

Definition 2.4. Let ρ be a pseudometric on $\mathscr{F}$. Given two functions l and u, the bracket $[l, u]$ is the set of all functions f with $l \le f \le u$. An ε-bracket is a bracket $[l, u]$ with $\rho(l, u) < \varepsilon$. The bracketing number $N_{[\,]}(\varepsilon, \mathscr{F}, \rho)$ is the minimum number of ε-brackets needed to cover $\mathscr{F}$.

2.2. **Semi-Markov empirical process.** Let us consider the class $\mathscr{F}$, with envelope $\mathbf{F}$, of real measurable functions $f\colon E \times E \times \mathbb{R}_+ \to \mathbb{R}$, such that

A.3. $\iiint_{E \times E \times \mathbb{R}_+} \nu(dx)\, \mathbf{F}^2(x, y, s)\, Q(x, dy, ds) < \infty$.

This condition implies that $\nu(dx) Q(x, dy, ds)$ pertains to the class $M(\mathbf{X}, \mathbf{F})$.

Let us define the random sequence $\{Y_\ell(f)\colon \ell \ge 1, f \in \mathscr{F}\}$ by

$$Y_\ell(f) := f(J_{\ell-1}, J_\ell, X_\ell) - Qf(J_{\ell-1}),$$

and for $k \ge 1$, define the following sums:

$$S_k(f) := \sum_{\ell=1}^{k} Y_\ell(f), \quad \text{for } f \in \mathscr{F}.$$

Let us define the filtration $\mathcal{F}_\ell := \sigma(J_0, X_1, J_1, \ldots, X_\ell, J_\ell)$, for $\ell \ge 1$ and $\mathcal{F}_0 := \sigma(J_0)$. Clearly it is straightforward to see that, almost surely,

$$\mathbb{E}(Y_\ell(f) \mid \mathcal{F}_{\ell-1}) = 0, \tag{1}$$

together with A.1, and that the sequence $(Y_\ell(f))$ is an $(\mathcal{F}_\ell)$-martingale difference.

For each $n \geq 1$, consider a process $\{\mathbb{G}_n(f)\colon f \in \mathscr{F}\}$ defined by

$$\left\{\mathbb{G}_n(f) = \frac{1}{n^{1/2}} S_n(f)\colon f \in \mathscr{F}\right\}.$$

This is the empirical process of the semi-Markov process. Define

$$\sigma_n^2(f,g) = \frac{1}{n}\sum_{\ell=1}^{n} \mathsf{E}\left((Y_\ell(f) - Y_\ell(g))^2 \mid \mathcal{F}_{\ell-1}\right), \quad \text{for } f,g \in \mathscr{F}.$$

2.3. **The uniform CLT.** Establishing a uniform CLT, for the process $\{\mathbb{G}_n(f)\colon f \in \mathscr{F}\}$, essentially means showing, in the sense of Definition 2.5, that as $n \to \infty$,

$$\{\mathcal{L}(\mathbb{G}_n(f))\colon f \in \mathscr{F}\} \to \{\mathcal{L}(Z(f))\colon f \in \mathscr{F}\},$$

where the processes are indexed by $\mathscr{F}$ and are considered as random elements of the Banach space

$$\ell^\infty(\mathscr{F}) := \left\{z\colon \mathscr{F} \to \mathbb{R}\colon \|z\|_{\mathscr{F}} := \sup_{f\in\mathscr{F}} |z(f)| < \infty\right\},$$

the space of the bounded real-valued functions on $\mathscr{F}$, taken with the sup norm. The limiting process $Z = \{Z(f)\colon f \in \mathscr{F}\}$ is a Gaussian process whose sample paths are contained in the class of uniformly continuous bounded function

$$U_B(\mathscr{F},\rho) := \{z \in \ell^\infty(\mathscr{F})\colon z \text{ is uniformly continuous with respect to } \rho\},$$

where ρ is a metric on $\mathscr{F}$. Notice that $(\ell^\infty(\mathscr{F}), \|\cdot\|_{\mathscr{F}})$ is a Banach space and $U_B(\mathscr{F},\rho)$ is a closed subspace of $(\ell^\infty(\mathscr{F}), \|\cdot\|_{\mathscr{F}})$ and hence is a Banach space too. In particular $U_B(\mathscr{F},\rho)$ is separable if and only if $(\mathscr{F},\rho)$ is totally bounded. We equip the space $\mathscr{F}$ with the pseudometric $d_Q^{(2)}$ so that $\left(\mathscr{F}, d_Q^{(2)}\right)$ is totally bounded.

In the sequel, we use the following definition of weak convergence, which is originally due to [14]. Throughout this paper events are identified with their indicator functions, and $\mathbb{E}^*$ denotes the upper expectation with respect to the outer probability $\mathbb{P}^*$; refer to [33] for a definition.

Definition 2.5. A sequence of $\ell^\infty(\mathscr{F})$-valued random functions $\{T_n\colon n \geq 1\}$ converges in law to an $\ell^\infty(\mathscr{F})$-valued Borel measurable random function T whose law concentrates on a separable subset of $\ell^\infty(\mathscr{F})$, denoted $T_n \rightsquigarrow T$, if

$$\mathbb{E}g(T) = \lim_{n\to\infty} \mathsf{E}^*\, g(T_n), \quad \text{for all } g \in C\left(\ell^\infty(\mathscr{F}), \|\cdot\|_{\mathscr{F}}\right),$$

where $C\left(\ell^\infty(\mathscr{F}), \|\cdot\|_{\mathscr{F}}\right)$ is the set of all bounded, continuous functions from the space $(\ell^\infty(\mathscr{F}), \|\cdot\|_{\mathscr{F}})$ into $\mathbb{R}$.

Theorem 2.6. *Suppose that $\mathscr{F}$ has uniformly integrable entropy with envelope function* **F** *fulfilling assumption* A.3, *and assumptions* A.1 *and* A.2 *are also fulfilled. Then the following weak convergence holds, as $n \to \infty$:*

$$\{\mathbb{G}_n(f)\colon f \in \mathscr{F}\} \rightsquigarrow \{\mathbb{Z}(f)\colon f \in \mathscr{F}\} \quad \textit{as random elements of } \ell^\infty(\mathscr{F}).$$

The limiting process $\{\mathbb{Z}(f)\colon f \in \mathscr{F}\}$ is a zero mean Gaussian process with covariance structure

$$\mathsf{E}\,\mathbb{Z}(f)\mathbb{Z}(g) = \int_E \nu(dx)\left(\iint_{E\times\mathbb{R}_+} f(x,y,s)g(x,y,s)\,Q(x,dy,ds) - \iint_{E\times\mathbb{R}_+} f(x,y,s)\,Q(x,dy,ds)\iint_{E\times\mathbb{R}_+} g(x,y,s)\,Q(x,dy,ds)\right),$$

and the sample paths of $\{\mathbb{Z}(f)\colon f \in \mathscr{F}\}$ *are bounded and uniformly continuous with respect to the metric* $d_Q^{(2)}$, *i.e., belong to* $U_B(\mathscr{F}, d_Q^{(2)})$.

Consider now a process $\{\mathbf{G}_n(s,f)\colon (s,f) \in [0,1] \times \mathscr{F}\}$ defined by

$$\left\{\mathbf{G}_n(s,f) = \frac{1}{n^{1/2}} S_{\lfloor ns \rfloor}(f)\colon (s,f) \in [0,1] \times \mathscr{F}\right\}.$$

Suppose that $\mathscr{F}$ has uniformly integrable entropy with envelope function $\mathbf{F}$ fulfilling A.3 and the same conditions of Theorem 2.6. Then by the same arguments as those used in Theorem 4 of [2] (refer also to [22, Theorem 2.1]), one can show the following theorem.

Theorem 2.7. *The following weak convergence holds:*

$$\{\mathbf{G}_n(s,f)\colon (s,f) \in [0,1] \times \mathscr{F}\} \rightsquigarrow \{\mathcal{Z}(s,f)\colon (s,f) \in [0,1] \times \mathscr{F}\}$$

as random elements of $\ell^\infty([0,1] \times \mathscr{F})$. *The limiting process* $\{\mathcal{Z}(s,f)\colon (s,f) \in [0,1] \times \mathscr{F}\}$ *is a centered Gaussian process with covariance structure*

$$\begin{aligned} &\mathsf{E}\,\mathcal{Z}(s,f)\mathcal{Z}(t,g) \\ &\quad = (s \wedge t) \int_E \nu(dx) \left(\iint_{E \times \mathbb{R}_+} f(x,y,s) g(x,y,s)\, Q(x,dy,ds) \right. \\ &\qquad\qquad \left. - \iint_{E \times \mathbb{R}_+} f(x,y,s)\, Q(x,dy,ds) \iint_{E \times \mathbb{R}_+} g(x,y,s)\, Q(x,dy,ds) \right), \end{aligned}$$

and the sample paths of $\{\mathcal{Z}(s,f)\colon (s,f) \in [0,1] \times \mathscr{F}\}$ *are bounded and uniformly continuous with respect to the metric* $|\cdot| + d_Q^{(2)}$, *i.e., belong to* $U_B([0,1] \times \mathscr{F}, |\cdot| + d_Q^{(2)})$.

Notice that we have

$$\mathsf{E}\big(\mathbf{G}_n(s,f)\mathbf{G}_n(t,f)\big) = \frac{\lfloor ns \rfloor \wedge \lfloor nt \rfloor}{n}\, \mathsf{E}\,\mathbb{Z}(f)\mathbb{Z}(g).$$

Then

$$\mathsf{E}\,\mathcal{Z}(s,f)\mathcal{Z}(t,g) = (s \wedge t)\, \mathsf{E}\,\mathbb{Z}(f)\mathbb{Z}(g).$$

By the multivariate CLT for stationary and ergodic martingale differences, see, for example, [15], the finite dimensional distributions of the process $\{\mathbf{G}_n(s,f)\colon (s,f) \in [0,1] \times \mathscr{F}\}$ converge to those of $\{\mathcal{Z}(s,f)\colon (s,f) \in [0,1] \times \mathscr{F}\}$. We conclude by applying Theorem 2.12.1 of [33], which says: "A class of measurable functions is functionally Donsker if and only if it is Donsker."

2.4. **Weighted bootstrap.** In a variety of statistical problems, the bootstrap provides a simple method for circumventing technical difficulties due to the intractable distribution theory and has become a powerful tool for setting confidence intervals and critical values of tests for composite hypotheses. Note that, in general, the bootstrap, according to Efron's original formulation (see [12]), presents some drawbacks. More specifically, some observations may be used more than once while others are not sampled at all. To overcome that problem, a more general formulation of the bootstrap has been introduced, the *weighted* bootstrap, which has also been shown to be computationally more efficient in several applications. For a survey of further results on weighted bootstrap the reader is referred to [4, 29] and the references therein. In this section, we provide a multiplier central limit theorem. A more general bootstrap will be considered elsewhere.

Let $\{z_i\}_{i \geq 1}$ be a sequence of random variables satisfying the following assumption.

B. The $\{z_i\}_{i \geq 1}$ are independent and identically distributed, on the probability space $(\Omega_z, \mathcal{A}_z, P_z)$ with mean zero and variance 1.

Throughout the section, we assume that the bootstrap weights z_i are independent from the data $\{\mathbb{G}_n(f)\colon n \in \mathbb{N}, f \in \mathcal{F}\}$. Define, for a measurable function $f \in \mathcal{F}$,

$$S_k^*(f) := \sum_{\ell=1}^{k} z_\ell Y_\ell(f), \qquad f \in \mathscr{F}.$$

For each $n \geq 1$, consider a bootstrapped process $\{\mathbb{G}_n^*(f)\colon f \in \mathscr{F}\}$ defined by

$$\left\{\mathbb{G}_n^*(f) = \frac{1}{n^{1/2}} S_n^*(f)\colon f \in \mathscr{F}\right\}.$$

This is the bootstrapped (or wild bootstrap of) empirical process of the semi-Markov process. Notice that we have from conditions (A.3) and (B) that the second moments of the martingale difference are finite, i.e.,

$$\begin{aligned}
&\iiiint_{E\times E\times\mathbb{R}_+\times\mathbb{R}_+} z^2\nu(dx)\mathbf{F}^2(x,y,s)Q(x,dy,ds)P_z(dz)\\
&\quad= \iiint_{E\times E\times\mathbb{R}_+} \nu(dx)\mathbf{F}^2(x,y,s)\,Q(x,dy,ds)\int_{\mathbb{R}} z^2 P_z(dz)\\
&\quad= \iiint_{E\times E\times\mathbb{R}_+} \nu(dx)\mathbf{F}^2(x,y,s)\,Q(x,dy,ds) < \infty.
\end{aligned}$$

It is worth noticing that the random variables z_i given in the definition of

$$\left\{\mathbb{G}_n^*(f) = \frac{1}{n^{1/2}} S_n^*(f)\colon f \in \mathscr{F}\right\}$$

do not effect the geometry of $\mathscr{F}$. Thus we can state our result with envelopes $|z_i|\mathbf{F}$. The following result gives application of Theorem 2.6 to the bootstrap.

Corollary 2.8. *Suppose that $\mathscr{F}$ has uniformly integrable entropy with envelope function* $\mathbf{F}$ *fulfilling assumption* A.3*, and assumptions* A.1*,* A.2*, and* B *are also fulfilled. Then the following unconditional weak convergence holds:*

$$\{\mathbb{G}_n^*(f)\colon f \in \mathscr{F}\} \rightsquigarrow \left\{\widetilde{\mathbb{Z}}(f)\colon f \in \mathscr{F}\right\} \quad \textit{as random elements of } \ell^\infty(\mathscr{F}).$$

The limiting process $\{\widetilde{\mathbb{Z}}(f)\colon f \in \mathscr{F}\}$ *is an independent copy of the process defined in Theorem* 2.6*.*

It is well known that Corollary 2.8 can be used easily, through routine bootstrap sampling, to evaluate the limiting distributions.

3. Proof of Theorem 2.6

This section is devoted to the detailed proofs of our results. The previously displayed notation continue to be used in the sequel. The proof of this theorem consists of two parts. The first one concerns the convergence of finite dimensional distributions, and the second one the tightness of the family of the initial processes. Finally, we derive the limit covariance.

The finite dimensional distributions convergence, by the Cramér–Wold device and the linearity of $\mathbb{G}_n(\cdot)$, suffices to show that, as $n \to \infty$, $\mathbb{G}_n(f) \to Z(f)$, for any $f \in \mathscr{F}$. This comes from the CLT for semi-Markov processes; see, e.g., [21]. We also use the fact that, under our conditions, for fixed $f \in \mathscr{F}$, as $n \to \infty$,

$$\mathbb{G}_n(f) \to N\left(0, \sigma^2(f)\right),$$

where

$$\sigma^2(f) = \int_E \nu(dx) \left(\iint_{E\times\mathbb{R}_+} f^2(x,y,s)\, Q(x,dy,ds) - \left(\iint_{E\times\mathbb{R}_+} f(x,y,s)\, Q(x,dy,ds) \right)^2 \right).$$

For each $f \in \mathscr{F}$, we have

$$\begin{aligned}
\frac{1}{n}\sum_{\ell=1}^n \mathsf{E}\left(Y_\ell(f)^2 \mid \mathcal{F}_{\ell-1}\right) &= \frac{1}{n}\sum_{\ell=1}^n \mathsf{E}\left(\left(f(J_{\ell-1},J_\ell,X_\ell) - Qf(J_{\ell-1})\right)^2 \,\middle|\, \mathcal{F}_{\ell-1}\right) \\
&= \frac{1}{n}\sum_{\ell=1}^n \left(\iint_{E\times\mathbb{R}_+} f^2(J_{\ell-1},y,s)\, Q(J_{\ell-1},dy,ds) \right. \\
&\qquad \left. - \left(\iint_{E\times\mathbb{R}_+} f(J_{\ell-1},y,s)\, Q(J_{\ell-1},dy,ds) \right)^2 \right) \\
&\xrightarrow{\mathbb{P}\text{-a.s.}} \int_E \nu(dx) \left(\iint_{E\times\mathbb{R}_+} f^2(x,y,s)\, Q(x,dy,ds) \right. \\
&\qquad \left. - \left(\iint_{E\times\mathbb{R}_+} f(x,y,s)\, Q(x,dy,ds) \right)^2 \right) \\
&< \infty.
\end{aligned} \tag{2}$$

Now, we show that the following convergence holds as n tends to infinity, in probability,

$$\frac{1}{n}\sum_{\ell=1}^n \mathsf{E}\left(Y_\ell^2(f)\mathbb{1}_{\{Y_\ell(f)>\eta\sqrt{n}\}} \mid \mathcal{F}_{\ell-1}\right) \to 0.$$

Recall that, from the condition A.3, we have

$$\int_E \nu(dx) \left(\iint_{E\times\mathbb{R}_+} \mathbf{F}^2(x,y,s)\, Q(x,dy,ds) - \left(\iint_{E\times\mathbb{R}_+} \mathbf{F}(x,y,s)\, Q(x,dy,ds) \right)^2 \right) < \infty.$$

Using the inequality $(a+b)^2 \le 2(a^2+b^2)$ in combination with a straightforward application of the Lebesgue dominated convergence theorem, it is easy to see that

$$\begin{aligned}
&\frac{1}{n}\sum_{\ell=1}^n \mathsf{E}\left(Y_\ell^2(f)\mathbb{1}_{\{Y_\ell(f)>\eta\sqrt{n}\}} \mid \mathcal{F}_{\ell-1}\right) \\
&\quad \le \frac{1}{n}\sum_{\ell=1}^n \mathsf{E}\left(\left\{ 2f^2(J_{\ell-1},J_\ell,X_\ell) \right.\right. \\
&\qquad\qquad \left. +2\left(\int_E \nu(dx) \iint_{E\times\mathbb{R}_+} f^2(x,y,s) Q(x,dy,ds) \right)^2 \right\} \\
&\qquad\qquad\qquad \left. \times\, \mathbb{1}_{\{Y_\ell(f)>\eta\sqrt{n}\}} \,\middle|\, \mathcal{F}_{\ell-1} \right)
\end{aligned}$$

$$\leq \frac{1}{n}\sum_{\ell=1}^{n} \mathsf{E}\Bigg(\Bigg\{2\mathbf{F}^2(J_{\ell-1},J_\ell,X_\ell) + 2\left(\int_E \nu(dx)\iint_{E\times\mathbb{R}_+}\mathbf{F}(x,y,s)Q(x,dy,ds)\right)^2\Bigg\} \times \mathbb{1}_{\{\mathbf{F}(J_{\ell-1},J_\ell,X_\ell)>\eta\sqrt{n}\}}\ \Bigg|\ \mathcal{F}_{\ell-1}\Bigg)$$

$$= \frac{2}{n}\sum_{\ell=1}^{n} \mathsf{E}\left(\mathbf{F}^2(J_{\ell-1},J_\ell,X_\ell)\mathbb{1}_{\{\mathbf{F}(J_{\ell-1},J_\ell,X_\ell)>\eta\sqrt{n}\}}\ \Big|\ \mathcal{F}_{\ell-1}\right) + \frac{2}{n}\sum_{\ell=1}^{n} \mathsf{E}\Bigg(\left(\int_E \nu(dx)\iint_{E\times\mathbb{R}_+}\mathbf{F}(x,y,s)Q(x,dy,ds)\right)^2 \times \mathbb{1}_{\{\mathbf{F}(J_{\ell-1},J_\ell,X_\ell)>\eta\sqrt{n}\}}\ \Bigg|\ \mathcal{F}_{\ell-1}\Bigg)$$

$$= \frac{2}{n}\sum_{\ell=1}^{n} \mathsf{E}\Bigg(\mathbf{F}^2(J_{\ell-1},J_\ell,X_\ell) \times \mathbb{1}_{\{\mathbf{F}(J_{\ell-1},J_\ell,X_\ell)>\eta\sqrt{n}\}}\ \Bigg|\ \mathcal{F}_{\ell-1}\Bigg) + \left(\int_E \nu(dx)\iint_{E\times\mathbb{R}_+}\mathbf{F}(x,y,s)Q(x,dy,ds)\right)^2 \times\left\{\frac{2}{n}\sum_{\ell=1}^{n}\mathsf{E}\left(\mathbb{1}_{\{\mathbf{F}(J_{\ell-1},J_\ell,X_\ell)>\eta\sqrt{n}\}}\ \Big|\ \mathcal{F}_{\ell-1}\right)\right\} \to 0.$$

This gives the desired assertion. Now, we prove that there exists a positive constant D such that, as $n\to\infty$,

$$\mathbb{P}^*\left(\sup_{f,g\in\mathscr{F}}\frac{\sigma_n^2(f,g)}{d^2(f,g)}\geq D\right)\to 0. \tag{3}$$

But this follows, as in [3], from the fact, as $n\to\infty$, that

$$\mathsf{E}^*\sup_{f,g\in\mathscr{F}}\sum_{\ell=1}^{n}\frac{\mathsf{E}\left((Y_\ell(f)-Y_\ell(g))^2\mid\mathcal{F}_{\ell-1}\right)}{nd_Q^2(f,g)}\to 1.$$

Now, from relations (1), (2), (3), and by virtue of the uniform central limit theorem for martingale differences (see, e.g., [3, Theorem 2]), we get the desired result. Finally, we give here the calculus of limit covariance function, for $f,g\in\mathscr{F}$,

$$\frac{1}{n}\sum_{\ell=1}^{n}\mathsf{E}\left(Y_\ell(f)Y_\ell(g)\mid\mathcal{F}_{\ell-1}\right)$$
$$= \frac{1}{n}\sum_{\ell=1}^{n}\mathsf{E}\big(\left(f(J_{\ell-1},J_\ell,X_\ell)-Qf(J_{\ell-1})\right)\left(g(J_{\ell-1},J_\ell,X_\ell)-Qg(J_{\ell-1})\right)\mid\mathcal{F}_{\ell-1}\big)$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{\ell=1}^{n} \mathsf{E}\big(f(J_{\ell-1},J_\ell,X_\ell)g(J_{\ell-1},J_\ell,X_\ell)\mid \mathcal{F}_{\ell-1}\big)\\
&\quad - \mathsf{E}\big(f(J_{\ell-1},J_\ell,X_\ell)Qg(J_{\ell-1})\mid \mathcal{F}_{\ell-1}\big) - \mathsf{E}\big(g(J_{\ell-1},J_\ell,X_\ell)Qf(J_{\ell-1})\mid \mathcal{F}_{\ell-1}\big)\\
&\quad + \mathsf{E}\big(Qf(J_{\ell-1})Qg(J_{\ell-1})\mid \mathcal{F}_{\ell-1}\big)\\
&= \frac{1}{n}\sum_{\ell=1}^{n}\Bigg(\iint_{E\times\mathbb{R}_+} f(J_{\ell-1},y,s)g(J_{\ell-1},y,s)Q(J_{\ell-1},dy,ds)\\
&\qquad - \iint_{E\times\mathbb{R}_+} f(J_{\ell-1},y,s)Q(J_{\ell-1},dy,ds)\\
&\qquad\qquad \times \iint_{E\times\mathbb{R}_+} g(J_{\ell-1},y,s)Q(J_{\ell-1},dy,ds)\Bigg)\\
&\xrightarrow{\mathbb{P}\text{-a.s.}} \int_E \nu(dx)\Bigg(\iint_{E\times\mathbb{R}_+} f(x,y,s)g(x,y,s)Q(x,dy,ds)\\
&\qquad - \iint_{E\times\mathbb{R}_+} f(x,y,s)Q(x,dy,ds)\iint_{E\times\mathbb{R}_+} g(x,y,s)Q(x,dy,ds)\Bigg).
\end{aligned}
$$

Hence the proof of Theorem 2.6 is complete. □

Acknowledgments

The authors would like to thank the editor in chief and the referee for their very helpful comments.

Bibliography

[1] P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding, *Statistical Models Based on Counting Processes*, Springer Ser. Statist., Springer-Verlag, New York, 1993. MR1198884

[2] J. Bae and M. J. Choi, *The uniform CLT for martingale difference of function-indexed process under uniformly integrable entropy*, Commun. Korean Math. Soc. **14** (1999), no. 3, 581–595. MR1791682

[3] J. Bae, D. Jun, and S. Levental, *The uniform CLT for martingale difference arrays under the uniformly integrable entropy*, Bull. Korean Math. Soc. **47** (2010), no. 1, 39–51. MR2604230

[4] P. Barbe and P. Bertail, *The Weighted Bootstrap*, Lecture Notes in Statist., vol. 98, Springer-Verlag, New York, 1995. MR2195545

[5] I. Berkes and W. Philipp, *Approximation theorems for independent and weakly dependent random vectors*, Ann. Probab. **7** (1979), no 1, 29–54. MR515811

[6] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons Inc., New York, 1968. MR0233396

[7] S. Bouzebda, Chr. Chrysanthi Papamichail, and N. Limnios, *On a multidimensional general bootstrap for empirical estimator of continuous-time semi-Markov kernels with applications*, J. Nonparametr. Stat. (2017). MR3756233

[8] S. Bouzebda and N. Limnios, *The uniform CLT for empirical estimator of a semi-Markov kernel indexed by functions with applications* (2017) (under revision).

[9] V. V. Buldygin, O. I. Klesov, and J. G. Steinebach, *Asymptotic properties of absolutely continuous functions and strong laws of large numbers for renewal processes*, Theory Probab. Math. Statist. **87** (2013), 1–12. MR3241442

[10] M. Csörgő and L. Horváth, *Weighted Approximations in Probability and Statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Ltd., Chichester, 1993. MR1215046

[11] R. M. Dudley, *Uniform Central Limit Theorems*, Cambridge Stud. Adv. Math., vol. 63, Cambridge University Press, Cambridge, 1999. MR1720712

[12] B. Efron, *Bootstrap methods: another look at the jackknife*, Ann. Statist. **7** (1979), no. 1, 1–26. MR515681

[13] S. Georgiadis and N. Limnios, *A multidimensional functional central limit theorem for an empirical estimator of a continuous-time semi-Markov kernel*, J. Nonparametr. Stat. **24** (2012), no. 4, 1007–1017. MR2995489

[14] J. Hoffmann-Jørgensen, *Stochastic Processes on Polish Spaces*, Various Publications Series (Aarhus), vol. 39, Aarhus Universitet Matematisk Institut, Aarhus, 1991. MR1217966
[15] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, second edition, Grundlehren Math. Wiss., vol. 288, Springer-Verlag, Berlin, 2003. MR1943877
[16] A. N. Kolmogorov and V. M. Tihomirov, *ε-entropy and ε-capacity of sets in functional space*, Amer. Math. Soc. Transl. **17** (1961), no. 2, 277–364. MR0124720
[17] V. S. Koroliuk and N. Limnios, *Stochastic Systems in Merging Phase Space*, World Scientific, New York, 2005. MR2205562
[18] M. R. Kosorok, *Introduction to Empirical Processes and Semiparametric Inference*, Springer Ser. Statist., Springer, New York, 2008. MR2724368
[19] M. Ledoux and M. Talagrand, *Probability in Banach Spaces. Isoperimetry and Processes*, Ergeb. Math. Grenzgeb., vol. 23, Springer-Verlag, Berlin, 1991. MR1102015
[20] N. Limnios, *A functional central limit theorem for the empirical estimator of a semi-Markov kernel*, J. Nonparametr. Stat. **16** (2004), no. 1–2, 13–18. MR2053060
[21] N. Limnios and G. Oprişan, *Semi-Markov Processes and Reliability*, Stat. Ind. Technol., Birkhäuser Boston Inc., Boston, MA, 2001. MR1843923
[22] Y. Nishiyama, *Some central limit theorems for l^∞-valued semimartingales and their applications*, Probab. Theory Related Fields **108** (1997), no. 4, 459–494. MR1465638
[23] D. Pollard, *Convergence of Stochastic Processes*, Springer Ser. Statist., Springer-Verlag, New York, 1984. MR762984
[24] D. Pollard, *Empirical Processes: Theory and Applications*, NSF-CBMS Regional Conference Series in Probability and Statistics, 2, Institute of Mathematical Statistics, Hayward, CA; American Statistical Association, Alexandria, VA, 1990. MR1089429
[25] R. Pyke, *Markov renewal processes: definitions and preliminary properties*, Ann. Math. Statist. **32** (1961), 1231–1242. MR0133888
[26] R. Pyke, *Markov renewal processes with finitely many states*, Ann. Math. Statist. **32** (1961), 1243–1259. MR0154324
[27] M. Rosenblatt, *Remarks on some nonparametric estimates of a density function*, Ann. Math. Statist. **27** (1956), 832–837. MR0079873
[28] J. Janssen and N. Limnios (eds.), *Semi-Markov Models and Applications*. Selected papers from the 2nd International Symposium on Semi-Markov Models: Theory and Applications held in Compiègne, December 1998. Kluwer Academic Publishers, Dordrecht, 1999. MR1772933
[29] J. Shao and D. S. Tu, *The Jackknife and Bootstrap*, Springer Ser. Statist., Springer-Verlag, New York, 1995. MR1351010
[30] G. R. Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986. MR838963
[31] D. Silvestrov, *Limit Theorems for Randomly Stopped Stochastic Processes*, Springer-Verlag, London, 2004. MR2030998
[32] A. van der Vaart, *New Donsker classes*, Ann. Probab. **24** (1996), no. 4, 2128–2140. MR1415244
[33] A. W. van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes*, Springer Ser. Statist., Springer-Verlag, New York, 1996. MR1385671

Sorbonne Universités, Université de Technologie de Compiègne, Laboratoire de Mathématiques Appliquées de Compiègne, Rue du Dr Schweitzer, CS 60319, 60205 Compiegne Cedex, France

Email address: salim.bouzebda@utc.fr

Sorbonne Universités, Université de Technologie de Compiègne, Laboratoire de Mathématiques Appliquées de Compiègne, Rue du Dr Schweitzer, CS 60319, 60205 Compiegne Cedex, France

Email address: nikolaos.limnios@utc.fr

Received 7/FEB/2017
Originally published in English