Statistics
Exchangeably weighted bootstraps of empirical estimators of a semi-Markov kernel

Le bootstrap échangeable pondéré de l'estimateur empirique du noyau semi-markovien

Salim Bouzebda, Nikolaos Limnios

Laboratoire de mathématiques appliquées de Compiègne, université de technologie de Compiègne, CS 60319, 60205 Compiègne cedex, France

A rticle info

Article history:
Received 9 February 2012
Accepted after revision 11 July 2013
Available online 8 August 2013
Presented by Paul Deheuvels

A bstract

A general notion of bootstrapped empirical estimators, of the semi-Markov kernels and of the conditional transition probabilities for semi-Markov processes with countable state space, constructed by exchangeably weighting sample, is introduced. Asymptotic properties of these generalized bootstrapped empirical distributions are obtained by means of the martingale approach.

R ésum é

Nous introduisons la notion du bootstrap échangeable des estimateurs empiriques des noyaux semi-markoviens et des probabilités de transition conditionnelles pour les processus semi-markoviens à espace d'état dénombrable. Nous obtenons nos résultats asymptotiques en utilisant les approches martingales.

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 [16], queueing theory, finance and insurance [14]. The theory of the semi-Markov processes is given by [20,21]. For recent references in this area along with statistical applications, see, e.g., [17] and [3]. 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., [16] and [3] and references therein. Note that the limiting distributions of these estimators, or their functionals, are rather complicated, which does not permit explicit computation in practice. To overcome that difficulty, we will propose a general bootstrap of empirical semi-Markov kernels and of the conditional transition probabilities and study some of its asymptotic properties by means of martingale techniques. The interest in considering general bootstrap instead of particular cases lies in the fact that we need, in general, a more flexible modeling to handle the problems in practice. In a variety of statistical problems, bootstrap provides a simple method for circumventing technical difficulties due to intractable distribution theory and has become a powerful tool for setting confidence intervals and critical values.

E-mail addresses: salim.bouzebda@utc.fr (S. Bouzebda), nikolaos.limnios@utc.fr (N. Limnios).
of tests for composite hypotheses. A substantial body of literature, reviewed in [4], gives conditions for the bootstrap to be satisfied in order to provide desirable distributional approximations. In [7], the performance of different kinds of bootstrap procedures is investigated through asymptotic results and small sample simulation studies. Note that bootstrapping, according to Efron’s original formulation (see [13]), presents some drawbacks. Namely, some observations may be used more than once, while others are not sampled at all. To overcome that problem, a more general formulation of bootstrap has been introduced, weighted (or smooth) 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 [2]. Another resampling scheme was proposed in [22] and was extensively studied by [6], who suggested the name “weighted bootstrap”: e.g., Bayesian bootstrap when the vector of weights \( (W_{m1}, \ldots, W_{m}) = (D_{m1}, \ldots, D_{m}) \), is equal in distribution to the vector of \( n \) spacings of \( n - 1 \) ordered uniform \((0, 1)\) random variables, that is, \((D_{m1}, \ldots, D_{m}) \sim \text{Dirichlet}(n, 1, \ldots, 1)\). These resampling plans lead to the interest of a unified approach, generally designated as general weighted resampling, was first proposed by [18] and amongst others extended by [19]. To the best of our knowledge, general weighted resampling in the semi-Markov setting was open, giving the main motivation to our paper. The results obtained in the present paper are useful in many statistical problems, in the semi-Markov framework, as it is illustrated in [9] and [8]. We start by giving some notation and definitions that are needed for the forthcoming section. All random processes are defined on a complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\). This is a technical requirement which allows the construction of the Gaussian processes in our theorems, and is not restrictive, since one can expand the probability space to make it rich enough (see, e.g., Appendix 2 in [11], [12] 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, [17] for a description in details). Consider an infinite countable set, say \( E \), and an \( E \)-valued càdlàg time-homogeneous semi-Markov process \( \{Z_t; t \geq 0\} \), with an embedded Markov renewal process \( \{(J_k, S_k); \ k \in \mathbb{N}\} \), where \( \{J_k; k \in \mathbb{N}\} \) is the \( E \)-valued embedded Markov chain (EMC) of the successive visited states, and \( 0 = S_0 \leq S_1 \leq \cdots \leq S_k \leq S_{k+1} \leq \cdots \) are the jump times of the semi-Markov process \( \{Z_t; t \geq 0\} \). Define also \( X_k := S_k - S_{k-1}, \) for \( k \geq 1 \), the sojourn times, and the process \( \{N(t); t \geq 0\} \), which counts the number of jumps of \( \{Z_t; t \geq 0\} \), in the time interval \([0, t]\), by \( N(t) := \sup\{k \geq 0; S_k \leq t\} \). Let us also define \( \{N_i(t); t \geq 0\} \), the number of visits of \( \{Z_t; t \geq 0\} \) to state \( i \in E \) up to time \( t \), and \( \{N_{ij}(t); t \geq 0\} \), the number of direct jumps of \( \{Z_t; t \geq 0\} \) from state \( i \) to state \( j \) up to time \( t \) (see, e.g., [17]). More precisely, 

\[
N_i(t) := \sum_{k=1}^{N(t)} \mathbb{1}_{\{J_{k-1} = i\}} \quad \text{and} \quad N_{ij}(t) := \sum_{k=1}^{N(t)} \mathbb{1}_{\{J_k = i, J_{k-1} = j\}},
\]

where \( \mathbb{1}_A \) stands for the indicator function of the event \( A \). Particularly, in the case where we consider the renewal process \( \{S_k^i; k \geq 0\}, \ i \in E \), (eventually delayed, see, e.g., [17]), \( \{N_i(t); t \geq 0\} \) is the counting process of renewals. Let \( \mu_{ii} \) be the mean recurrence time of \( \{S_k^i; k \geq 0\}, \ i \in E \), i.e., \( \mu_{ii} = E[S_k^i - S_{k-1}^i] \), and let us denote by:

\[
v = \{v_i; \ i \in E\}
\]

the stationary distribution of the embedded Markov chain \( \{J_k; k \geq 0\} \). Let \( Q(t) = (Q_{ij}(t), i, j \in E, t \geq 0) \), be the semi-Markov kernel, which is defined by:

\[
Q_{ij}(t) := \mathbb{P}(J_{k+1} = j, X_{k+1} \leq t | J_k = i) = P(i, j) F_{ij}(t), \quad \text{for } t \geq 0, \ i, j \in E,
\]

where \( P(i, j) := \mathbb{P}(J_{k+1} = j | J_k = i) \), is the transition kernel of the EMC \( \{J_k; k \geq 1\} \), and \( F_{ij}(t) := \mathbb{P}(X_{k+1} \leq t | J_k = i, J_{k+1} = j), \) for \( t \geq 0, \ i, j \in E \), is the conditional distribution function of the sojourn times. Let us define also the distribution function \( H_i(t) := \sum_{j \in E} Q_{ij}(t) \), and its mean value \( m_i \), which is the mean sojourn time of \( \{Z_t; t \geq 0\} \) in the state \( i \in E \), i.e., \( m_i := \int_0^\infty (1 - H_i(t)) \, dt \). The mean sojourn time of the semi-Markov process \( \{Z_t; t \geq 0\} \) is defined by:

\[
\bar{m} := \sum_{i \in E} v_im_i.
\]

In the sequel, we need to recall the following useful property:

\[
\mu_{ii} = \bar{m}/v_i.
\]

Let us define the following observation in the time interval \([0, t]\):

\[
H_t := [Z_u, 0 \leq u \leq t] = \begin{cases} \{J_0, X_1, \ldots, J_{N(t)}, U_t\}, & \text{if } N(t) > 0, \\ \{J_0, U_t = t\}, & \text{if } N(t) = 0, \end{cases}
\]

where \( U_t := t - S_{N(t)} \) denotes the backward recurrence time. According to [16], we define the empirical estimator of the semi-Markov kernel by:

\[
\hat{Q}_{ij}(x, t) := \begin{cases} \frac{1}{N(t)} \sum_{k=1}^{N(t)} \mathbb{1}_{\{J_{k-1} = i, J_k = j, X_k \leq x\}}, & \text{for } 0 \leq x \leq t, \ i, j \in E, \\ 0, & \text{whenever } N(t) = 0. \end{cases}
\]
and the empirical estimator of the conditional transition distribution functions by:

$$
\hat{F}_{ij}(x, t) := \begin{cases} 
\frac{1}{N_{ij}(t)} \sum_{k=1}^{N_{ij}(t)} 1_{\{j_{k-1}=i, j_k=j, X_k \leq x\}}, & \text{for } 0 \leq x \leq t, \ i, j \in E, \\
0, & \text{whenever } N_{ij}(t) = 0.
\end{cases}
$$

(1.5)

All the asymptotic results in this note require the following assumptions:

(i) The Markov chain \(\{J_k: k \geq 0\}\) is ergodic with stationary distribution \(\nu\);
(ii) The mean sojourn times \(\{m_i: i \in E\}\) are finite;
(iii) The distribution functions of the sojourn times are not degenerate, i.e., not equal to \(\delta_0\), that is, Dirac distributions concentrated at 0.

Remark 1. The first and second assumptions ((i)–(ii)) are used in order to obtain the asymptotic result, since we are working on a single trajectory of the semi-Markov process \(\{Z_t: t \geq 0\}\). In particular, the stationarity of the semi-Markov process \(\{Z_t: t \geq 0\}\) is guaranteed by the assumption (ii). Finally, the assumption (iii) is used to avoid the explosion of the semi-Markov process \(\{Z_t: t \geq 0\}\). However, the Dirac distributions concentrated at the point \(a > 0\) are useful in order to introduce fixed durations.

The functional central limit theorems for estimators given in (1.4) and (1.5) are established in [16]. Notice that, in the aforementioned empirical estimators, the backward recurrence times are neglected. In fact, for the asymptotic properties, i.e., as \(t\) tends to infinity, \(U_t\) adds no significant information. The bootstraps of \(\{Q_{ij}(x, t): 0 \leq x \leq t, \ i, j \in E\}\) and \(\{\hat{F}_{ij}(x, t): 0 \leq x \leq t, \ i, j \in E\}\) are introduced in detail and their asymptotic properties are given in Section 2.

2. Main results

In this section, we shall establish the asymptotic properties of bootstrapping under quite general conditions in the framework of semi-Markov processes. Let \(W = (W_{nj}: \ j = 1, \ldots, n, \ n = 1, 2, \ldots)\) be a triangular array of random variables. This array determines a weighted bootstrap of the empirical estimator of the semi-Markov kernel by:

$$
Q_{ij}^W(x, t) := \begin{cases} 
\frac{1}{n_{ij}(t)} \sum_{k=1}^{N_{ij}(t)} W_{nj}(t)k1_{\{j_{k-1}=i, j_k=j, X_k \leq x\}}, & \text{for } 0 \leq x \leq t, \ i, j \in E, \\
0, & \text{whenever } N_{ij}(t) = 0,
\end{cases}
$$

(2.1)

and the weighted bootstrap of the empirical estimator of the conditional transition distribution functions by:

$$
F_{ij}^W(x, t) := \begin{cases} 
\frac{1}{n_{ij}(t)} \sum_{k=1}^{N_{ij}(t)} W_{nj}(t)k1_{\{j_{k-1}=i, j_k=j, X_k \leq x\}}, & \text{for } 0 \leq x \leq t, \ i, j \in E, \\
0, & \text{whenever } N_{ij}(t) = 0.
\end{cases}
$$

(2.2)

The bootstrap weights \(W_{ni}\)'s are assumed to belong to the class of exchangeable bootstrap weights investigated in [19]. We shall assume the following conditions:

(W1) The vector \(W_n = (W_{n1}, \ldots, W_{nn})^T\) is exchangeable for any \(n = 1, 2, \ldots, \) i.e., for any permutation \(\pi = (\pi_1, \ldots, \pi_n)\) of \((1, \ldots, n)\), the joint distribution of \(\pi(W_n) = (W_{n\pi_1}, \ldots, W_{n\pi_n})^T\) is the same as that of \(W_n\);
(W2) \(W_{ni} \geq 0\) for all \(n, i\) and \(\sum_{i=1}^{n} W_{ni} = n\) for all \(n\);
(W3) \(\limsup_{n \to \infty} \|W_{ni}\|_{2,1} \leq C < \infty\), where \(\|W_{ni}\|_{2,1} := \int_0^\infty \sqrt{\mathbb{E}(W_{ni}^2 > u)} \, du\); 
(W4) \(\lim_{n \to \infty} \sup_{\pi} \sup_{t \geq \lambda} \mathbb{E}(W_{ni} > t) = 0\);
(W5) \(1/n \sum_{i=1}^{n} (W_{ni} - 1)^2 \overset{p}{\to} c^2 > 0\).

Efron’s nonparametric bootstrap corresponds to the choice \(W_n \sim \text{Mult}(n; n^{-1}, \ldots, n^{-1})\) for which conditions W1–W5 are satisfied. In general, conditions W3–W5 are satisfied under some moment conditions on \(W_{ni}\), see [19, Lemma 3.1]. In addition to Efron’s nonparametric bootstrap, the sampling schemes that satisfy conditions W1–W5, include Bayesian bootstrap, Multiplier bootstrap, Double bootstrap and Urn bootstrap. This list is sufficiently long to indicate that conditions W1–W5 are not unduly restrictive. Notice that the value of \(c\) in W5 is independent of the sample size and depends on the resampling method, e.g., \(c = 1\) for the nonparametric bootstrap and Bayesian bootstrap, whereas \(c = \sqrt{2}\) for the double bootstrap.

A more precise discussion of this general formulation of the bootstrap and further details can be found in [18], [19], [23, §3.6.2, p. 353], [15, §10, p. 179], [10] and references therein. Throughout the paper, we assume that the bootstrap weights \(W_{ni}\)'s are independent from the data \(\{Z_u: 0 \leq u \leq t\}\). For any fixed states \(i, j \in E\) and fixed \(x \geq 0\), let us define the random sequences \(\{Y_{\ell}: \ell \geq 1\}\) and \(\{Y^*_{\ell}: \ell \geq 1\}\), respectively, by:

$$
Y_{\ell} = Y_{\ell}(i, j, x) := 1_{\{j_{\ell-1}=i, j_{\ell}=j, X_{\ell} \leq x\}} - 1_{\{j_{\ell-1}=i\}} Q_{ij}(x),
$$

$$
Y^*_{\ell} = Y^*_{\ell}(i, j, x) := W[nt]1_{\{j_{\ell-1}=i, j_{\ell}=j, X_{\ell} \leq x\}} - 1_{\{j_{\ell-1}=i\}} \hat{Q}_{ij}(x, nt),
$$

respectively.
Theorem 2.3. We have the following weak convergence, as $n \to \infty$:

$$S_{k}^{ij}(x) := \sum_{\ell=1}^{k} Y_{\ell}(i, j, x) \quad \text{and} \quad S_{k}^{eij}(x) := \sum_{\ell=1}^{k} Y_{\ell}^{e} = \sum_{\ell=1}^{k} Y_{\ell}(i, j, x). \quad (2.3)$$

In the sequel, $\{W(t), t \geq 0\}$ denotes the standard Brownian motion, that is, a centered Gaussian process with continuous sample paths and covariance function $E(W(s)W(t)) = \min(s, t)$, for $s, t \geq 0$. The following weak convergence result is due to [16]. For any fixed states $i, j \in E$ and fixed $x \geq 0$, as $n$ tends to infinity, we have:

$$n^{-1/2} S_{[nt]}^{ij}(x) \xrightarrow{d} \sigma_{ij}(x) W(t), \quad (2.4)$$

with the arrow “$\xrightarrow{d}$” denoting the weak convergence of random element in the Skorohod space $D[0, \infty)$, provided that $\sigma_{ij}^{2}(x) > 0$, where $\sigma_{ij}^{2}(x) := v_{i}Q_{ij}(x)(1 - Q_{ij}(x))$. We are now in position to state our first result in the following theorem, which gives the bootstrap version of (2.4).

**Theorem 2.1.** Let $W$ be a triangular array of bootstrap weights satisfying assumptions W1–W5. For any fixed states $i, j \in E$ and fixed $x \geq 0$, the following weak convergence holds, as $n \to \infty$:

$$n^{-1/2} S_{[nt]}^{eij}(x) \xrightarrow{d} (1 + c^{2})^{1/2} \sigma_{ij}(x) W(t), \quad \text{for } t > 0. \quad (2.5)$$

**Theorem 2.2.** Let $W$ be a triangular array of bootstrap weights satisfying assumptions W1–W5. For any fixed $i, j \in E$, and fixed $x \geq 0$, we have the following weak convergence, as $n \to \infty$:

$$n^{1/2} (Q_{ij}^{W}(x, nt) - \hat{Q}_{ij}(x, nt)) \xrightarrow{d} (1 + c^{2})^{1/2} b_{ij}(x) W(t)/t, \quad \text{for } t > 0, \quad (2.6)$$

provided that $b_{ij}^{2}(x) > 0$, where:

$$b_{ij}(x) := \mu_{ii}Q_{ij}(x)(1 - Q_{ij}(x)).$$

For any fixed states $i, j \in E$ and fixed $x \geq 0$, let us define the random sequences $\{Y_{\ell}^{i} : \ell \geq 1\}$ and $\{Y_{\ell}^{e} : \ell \geq 1\}$, respectively, by:

$$Y_{\ell}^{i} = Y_{\ell}^{i}(i, j, x) := \mathbb{I}_{\{j_{\ell-1} = i, j_{\ell} = j, x_{\ell} \leq x\}} - \mathbb{I}_{\{j_{\ell-1} = i\}} F_{ij}(x),$$

$$Y_{\ell}^{e} = Y_{\ell}^{e}(i, j, x) := W_{[nt]}(\mathbb{I}_{\{j_{\ell-1} = i, j_{\ell} = j, x_{\ell} \leq x\}} - \mathbb{I}_{\{j_{\ell-1} = i\}} \tilde{F}_{ij}(x, nt)), \quad (2.3)$$

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

$$S_{k}^{ij}(x) := \sum_{\ell=1}^{k} Y_{\ell}^{i} = \sum_{\ell=1}^{k} Y_{\ell}^{i}(i, j, x) \quad \text{and} \quad S_{k}^{eij}(x) := \sum_{\ell=1}^{k} Y_{\ell}^{e} = \sum_{\ell=1}^{k} Y_{\ell}^{e}(i, j, x).$$

The following result has already been derived by [16]. For any fixed states $i, j \in E$ and fixed $x \geq 0$, the following weak convergence holds, as $n$ tends to infinity:

$$n^{-1/2} S_{[nt]}^{ij}(x) \xrightarrow{d} \sigma_{ij}(x) W(t), \quad (2.7)$$

provided that $\sigma_{ij}^{2}(x) > 0$, where $\sigma_{ij}^{2}(x) := v_{i}Q_{ij}(x)(1 - F_{ij}(x))$. We state our result in the following theorem which gives the bootstrap version of (2.7).

**Theorem 2.3.** Let $W$ be a triangular array of bootstrap weights satisfying assumptions W1–W5. For any fixed states $i, j \in E$ and fixed $x \geq 0$, the following weak convergence holds, as $n \to \infty$:

$$n^{-1/2} S_{[nt]}^{eij}(x) \xrightarrow{d} (1 + c^{2})^{1/2} \sigma_{ij}(x) W(t), \quad \text{for } t > 0. \quad (2.8)$$

**Theorem 2.4.** Let $W$ be a triangular array of bootstrap weights satisfying assumptions W1–W5. For any fixed $i, j \in E$, and fixed $x \geq 0$, we have the following weak convergence, as $n \to \infty$:

$$n^{1/2} (F_{ij}^{W}(x, nt) - \tilde{F}_{ij}(x, nt)) \xrightarrow{d} (1 + c^{2})^{1/2} c_{ij}(x) W(t)/t, \quad \text{for } t > 0, \quad (2.9)$$

provided that $c_{ij}^{2}(x) > 0$, where:

$$c_{ij}^{2}(x) := \frac{\mu_{ii}}{P(i, j)} F_{ij}(x)(1 - F_{ij}(x)).$$
Acknowledgement

The authors would like to thank the referee for careful reading and valuable comments on the manuscript, which improved the presentation of the paper.

References