Nonparametric Estimation for Semi-Markov Processes Based on its Hazard Rate Functions

BRAHIM OUHBI* and NIKOLAOS LIMNIOS**
Ecole Nationale Superieure d'Arts et Metiers, Universite de Technologie de Compiége

Abstract. The problem of estimating the Markov renewal matrix and the semi-Markov transition matrix based on a history of a finite semi-Markov process censored at time $T$ (fixed) is addressed for the first time. Their asymptotic properties are studied. We begin by the definition of the transition rate of this process and propose a maximum likelihood estimator for the hazard rate functions and then we show that this estimator is uniformly strongly consistent and converges weakly to a normal random variable. We construct a new estimator for an absolute continuous semi-Markov kernel and give detailed derivation of uniform strong consistency and weak convergence of this estimator as the censored time tends to infinity.

Key words: semi-Markov process, maximum likelihood estimator, transition rate, semi-Markov kernel, Markov renewal matrix, semi-Markov transition matrix, asymptotic properties.

1. Introduction

Semi-Markov process (SMP) or Markov renewal process (MRP) is a class of stochastic processes which generalize Markov jump process as well as the renewal process, (cf. Pyke [20]). They can be applied in different situations (Weibull, Erlang etc. distributions), and allow the computation of a large number of indicators. A fundamental question in semi-Markov problems is to know the semi-Markov transition matrix. To our knowledge, there is no paper which gives estimator of this matrix. The aims of the present paper is to introduce nonparametric estimators of the Markov renewal matrix and the semi-Markov matrix which are seen to have the desirable asymptotic properties and are simple to compute in practice. There are many papers which have given estimators of the semi-Markov kernel. In [14], Lagakos et al., proposed certain estimators and derived approximate variance and covariance of these estimators. Gill [7], Moore and Pyke [16], and Ouhbi and Limnios [17] have given estimators with several attractive asymptotic properties but these estimators present inconvenience when the censored time or the sample size is small.

* Postal address: Ecole Nationale Superieure d'Arts et Metiers, Marjane II, Meknes Ismailia, B.P. 4024 Beni M'Hamed, Meknes, Maroc
** Postal address: Université de Technologie de Compiége, Division Mathématiques Appliquées, B.P. 20529, 60205, Compiége Cedex, France
In this paper a maximum likelihood estimator of SMP kernels which have desirable asymptotic properties and do not present inconvenience when the censored time or the sample size is small is given. In our procedure, the central role is played by the SM kernel hazard rate.

This paper is organized as follows. In Section 2, the basic definitions are sketched. In Sections 3 and 4, hazard rate associated to a SM kernel is estimated and its asymptotic properties are also given. Section 5 is dedicated to estimate and to give asymptotic properties of the SM kernel. Section 6 introduces the Markov renewal matrix and its estimator as well as its asymptotic properties. In Section 7, we introduce the SM transition matrix as the solution of the well-known Markov renewal equation and we derive its estimator as well as its asymptotic properties.

2. Preliminaries

Let us consider a Markov renewal process, \((J, S) = (J_n, S_n)_{n \geq 0}\), defined on a probability complete space, where \((J_n)_{n \geq 0}\) is a Markov chain with values in \(E = \{1, \ldots, s\}\), \((s < \infty)\), which is the state space of the process, and \((S_n)_{n \geq 0}\) are the jump times which will be in \(\mathbb{R}_+ = [0, \infty]\). \(J_0, J_1, \ldots, J_n, \ldots\) are the consecutive states to be visited by the MRP and \(X_0 = 0, X_1, X_2, \ldots\) defined by \(X_n = S_n - S_{n-1}\), for \(n \geq 1\), are the sojourn times in these states taking values in \(\mathbb{R}_+\).

A MRP can be completely determined if we know its initial law and its semi-Markov kernel (which we shall estimate) defined respectively by \(P(J_0 = k) = p(k)\) and

\[
P(J_{n+1} = k, X_{n+1} \leq x | J_0, J_1, \ldots, J_n, X_1, X_2, \ldots, X_n) = Q_{i,k}(x) \quad (a.s.) \tag{1}
\]

for all \(x \in \mathbb{R}_+\) and \(1 \leq k \leq s\). The probabilities \(p_{ij} = Q_{ij}(\infty) = \lim_{t \to \infty} Q_{ij}(t)\) are the transition probabilities of the Markov chain \((J_n)_{n \geq 0}\) (cf. Pyke [20]).

In the sequel, we will assume that, \(Q_{ii}(t) \equiv 0\), for all \(i \in E\). Define the distribution function of the sojourn time in state \(i\) by,

\[
H_i(t) = \sum_{j=1}^{s} Q_{ij}(t), \quad \forall t \in \mathbb{R}.
\]

Let us also consider the distribution function associated to sojourn time in state \(i\), before going to state \(j\), defined by

\[
F_{ij}(\cdot) = \begin{cases} p_{ij}^{-1} \times Q_{ij}(\cdot) & \text{if } p_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}
\]

and \(\tilde{F}_{ij}(\cdot) = 1 - F_{ij}(\cdot)\) is the corresponding survival function to \(F_{ij}(\cdot)\).