Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 3127–3139
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Occupancy numbers for dynamic heterogeneous populations: Estimate of particles lifetimes

Paola Tardelli*

Dipartimento di Ingegneria Elettrica, Facoltá di Ingegneria, Universitá dell'Aquila, I-67040, Poggio di Roio, L'Aquila, Italy

Received 11 October 2005; accepted 16 May 2006
Available online 16 March 2007

Abstract

A dynamic model of a heterogeneous population is studied. Particles belonging to a population are divided, at every time t, into a finite number of classes according to their types and the partition changes over time. The role of the occupancy numbers, namely the cardinality of each class, is highlighted. The relationship between the stochastic process of occupancy numbers and the process of particle types is analyzed. The main goal of this paper is the estimation of the lifetime of each particle at a given time t, when the observed data are the history of the process of the number of dead particles up to t. Furthermore, a discrete time approximation of the filter is given.

MSC: 60G09; 62N05; 60H30; 60G35; 60J75

Keywords: Heterogeneous populations; Exchangeability; Stochastic filtering; Jump processes

1. Introduction

In several fields of Bayesian statistics and applied probability there is a growing interest towards populations of particles partitioned into classes. In particular, for some stochastic epidemic models, potentially infectious contacts occur both within classes, and between classes.

In survival analysis an issue is to study the conditional distribution of residual lifetimes for surviving particles, given some kind of past history of failures for other particles. In particular, considering a heterogeneous population, the evolution of the conditional distribution of residual lifetimes is of interest, given the observation of the number of particles already dead.

More precisely, let H be a finite positive integer value. In Gerardi et al. (2000a, b) H identical particles are simultaneously in the same environment and are subject to some source of stress. A family of exchangeable random variables, $Z_1, \ldots, Z_H$, is introduced in order to define the partition of the population, setting $Z_i = k$ if and only if the particle labelled i, $i = 0, 1, \ldots, H$, belongs to the class characterized by the type k, $k = 0, 1, \ldots, d$. The law of the lifetimes given the variables $Z_1, \ldots, Z_H$ is given and the lifetimes are assumed to be independent given the partition of the population. In their set up, the authors prove the exchangeability of the sequence of lifetimes and study the conditional law of lifetimes given the number of particles dead up to time t.

* Tel.: +39 0862 434431; fax: +39 06 23318800.
E-mail address: tardelli@ing.univaq.it.

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2006.05.020

In Gerardi and Tardelli (2005), a dynamic setting is considered, generalizing the previous model in the sense that the type of any single particle can change over time as well as the partition of the population. So, at every time $t \in \mathbb{R}^+$, the type of any given particle is represented as a stochastic process $Z(t) = \{Z_i(t)\}_{1 \leqslant i \leqslant H}$ and each component $Z_i(t)$ takes integer values between 0 and d. This means that, at every time t, the population is divided into $d + 1$ classes.

The type 0 is absorbing, in the sense that if a particle is of type 0, it stays that way indefinitely. Furthermore, a suitable assumption of exchangeability is made for the process $Z(t)$, generalizing the property of exchangeability, for fixed t, assumed in Gerardi et al. (2000b).

In this paper, the process $Z(t)$ is assumed to be Markovian, as in Gerardi and Tardelli (2005). The exchangeability property, just introduced in Gerardi and Tardelli (2005), is investigated and the relation between the exchangeability property and the Markov property is clarified. Such properties are essential tools to prove that the family of the lifetimes is an exchangeable sequence of random variables, already proved in Gerardi and Tardelli (2005).

The occupancy numbers, which are defined as the cardinality of each class, are introduced and their role is highlighted, generalizing some concepts presented in Gerardi et al. (2000b) in a dynamic framework.

To this end, let $X(t) = (X^1(t), \ldots, X^d(t))$ be a vector such that $X^i(t)$ is the number of particles belonging to the class labelled i and $Y(t)$ is the number of particles belonging to the class labelled 0. The exchangeability assumption, given in Gerardi and Tardelli (2005), is generalized in order to obtain a one-to-one relation between the law of $Z(t)$ and the law of the occupancy numbers. Moreover, the dynamics of $Z(t)$ is assumed to depend just on the number of particles belonging to each class. This assumption provides a necessary and sufficient condition for the Markov property of the occupancy number process, $(X(t), Y(t))$.

As in Gerardi et al. (2000b) and Gerardi and Tardelli (2005), the partition of the population is supposed to be non-observable. The observation process is just the number of particles dead up to time t, namely the history of the process $Y(t)$. Again, the aim is to find a prediction of the residual lifetimes given the history of the observations. Then, a stochastic filtering problem arises.

More precisely, first, the conditional law of the lifetimes given the state at time t of the process $(X(t), Y(t))$ is written down. Then, a filtering problem is solved, providing the conditional law of $X(t)$ given the history of $Y(t)$, namely the filter.

While in discrete time the filter can be deduced by a Bayes formula, in continuous time, the filter satisfies a stochastic differential equation known as the Kushner–Stratonovich equation which is, in some sense, a dynamic version of the Bayes formula. By a straightforward procedure, given in Brémaud (1980), this equation is written down. In general, the Kushner–Stratonovich equation does not have a unique solution. Therefore, to deduce the properties of the filter from it, some kind of uniqueness is needed. Weak uniqueness could be obtained, by using the filtering Martingale problem approach, as discussed in Kurtz (1998) and Kurtz and Ocone (1988). But, taking into account the peculiarity of the model presented here, a stronger kind of uniqueness, namely pathwise uniqueness, can be reached in a simpler way, with a method which is a modification of that presented in Arjas et al. (1992).

Some general results, given in Gerardi and Tardelli (2005), are recalled in Section 2 to investigate and clarify the meaning and the consistency of the assumptions given there. In Section 3, after describing the model, its relationship with the model presented in Gerardi and Tardelli (2005) is discussed. Section 4 is devoted to the computation of the conditional law of lifetimes given the observations. Finally, in Section 5, a discrete time approximation is described and its convergence to the original filter is discussed.

As a final comment, let us note that our model is related to the closed migration process introduced, for instance, in Kelly (1994).

2. General setting

Let $\mathscr{P} = \{U_j\}_{j \geqslant 1}$ be a finite or countable population, where U_j are given identical particles. For H positive integer, let $\mathscr{P}_H = \{U_j\}_{j=1,\ldots,H}$ be a finite subpopulation of $\mathscr{P}$. Let $\mathscr{P}_H$ be heterogeneous in that its elements can be of d different types, labelled the natural numbers $1, \ldots, d$.

Then, given $t \in \mathbb{R}^+$, let $C_k(t)$ be the subset of all individuals of type k, $k = 1, \ldots, d$, at time t. Since in every class each particle can die, let $C_0(t)$ denote the class of the particles dead up to time t. In this way, $\mathscr{P}_H(t) = \bigcup_{k=0,1,\ldots,d} C_k(t)$.

In order to define the type of each particle at every time $t \in \mathbb{R}^+$, let $Z(t) = \{Z_i(t)\}_{1 \leqslant i \leqslant H}$ be defined as $Z_i(t) = k$ if and only if $U_i \in C_k(t)$, for $k = 0, 1, \ldots, d$ and for $i = 1, \ldots, H$. Events of the form $\{U_i \in C_k(t)\} = \{Z_i(t) = k\}$, for