Performance Evaluation 102 (2016) 1–20

Contents lists available at ScienceDirect

Performance Evaluation

journal homepage: www.elsevier.com/locate/peva

Analytical issues regarding the lack of identifiability of the non-stationary MAP_2

Joanna Rodríguez [a,*], Rosa E. Lillo [a], Pepa Ramírez-Cobo [b]

[a] Departamento de Estadística, Universidad Carlos III de Madrid, Spain

[b] Departamento de Estadística e Investigación Operativa, Universidad de Cádiz, Spain

ARTICLE INFO

Article history:
Received 5 October 2015
Received in revised form 13 April 2016
Accepted 30 June 2016
Available online 7 July 2016

Keywords:
Markovian arrival process (*MAP*)
Stationary vs. non-stationary versions
Identifiable process
Canonical representation
Moments characterization

ABSTRACT

This paper studies in detail different problems concerning the identifiability of the non-stationary version of the MAP_2. First, a matrix-based methodology to build equivalent processes is given. Second, a unique, canonical representation of the process, so that the infinite, equivalent versions of a process can be reduced to its canonical counterpart is provided. Finally, a characterization of the process in terms of five descriptors representing moments of the three first inter-event time distributions is given.

1. Introduction

The Markovian Arrival Process (*MAP*) was introduced by Lucantoni [1], providing a convenient and more tractable representation of the preceding model proposed by Neuts [2]. It is a large class of point processes that generalizes the Poisson arrival process allowing for non-exponential and dependent inter-event times. By event we are referring to any type of occurrence in any context, for example, engineering, reliability or teletraffic. The phase-type (*PH*) renewal processes and the Markov Modulated Poisson process are well-known special cases of the *MAP*. The *MAP*s are versatile point processes (the stationary *MAP* is dense in the class of point processes, see Asmussen and Koole [3]) and have been widely considered in a number of real-life contexts, as reliability, queueing, teletraffic or climatology. See for example, Montoro-Cazorla et al. [4,5], Okamura et al. [6], Kang et al. [7], Casale et al. [8], Wu et al. [9] and Ramírez-Cobo et al. [10,11].

The non-stationary *MAP* is a matrix generalization of the Poisson process, that allows for correlated event occurrences and non-identically distributed inter-event times, unlike its stationary version where the inter-event times are considered to be identically distributed. It is important to point out that most of the works found in the literature are related to the stationary version of the process. Recently, the non-stationary *MAP* has gained attention in the literature because of its high potential in modeling problems where the inter-event times are correlated but not identically distributed. There are data in reliability and medical contexts that support the use of this version of the process. In the reliability context, consider the work of Rodríguez et al. [12], where the available data are the inter-failure times of 926 electrical generators of an electrical utility company. These generators are assumed to be structurally equal, but the inter-failure times associated with each generator cannot be assumed to be independent nor identically distributed. A non-stationary two-state *MAP* (noted MAP_2) was fitted to the data set, since it overperformed when compared to its stationary version. The parameter estimation

* Correspondence to: Departamento de Estadística, Universidad Carlos III de Madrid, Av. de la Universidad 30, 28911 Leganés (Madrid), Spain.
E-mail addresses: joannavirginia.rodriguez@uc3m.es (J. Rodríguez), rosaelvira.lillo@uc3m.es (R.E. Lillo), pepa.ramirez@uca.es (P. Ramírez-Cobo).

http://dx.doi.org/10.1016/j.peva.2016.06.008
0166-5316/

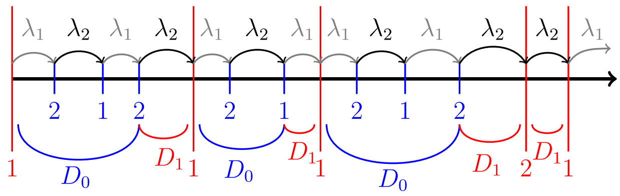

Fig. 1. Transition diagram for the MAP_2. The blue line corresponds to transitions without events, governed by D_0, and the red lines correspond to transitions with an event occurring, governed by D_1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

approach for fitting the data set is a direct consequence from the results found in this paper. In a medical context, in Mahé and Chevret [13] the events are the clinical relapses of cancer patients (usually a reduced number), which motivates the use of the non-stationary version of the process. Other works have studied the modeling capabilities of the non-stationary processes, for example, Zhu [14] studied the ruin probabilities for risk processes with dependent and non-stationary arrival processes that satisfy a large deviation principle. Ma and Whitt [15] created a way to efficiently simulate a non-Poisson non-stationary processes and studied the performance of the associated queueing models with those type of arrival processes. Kim and Whitt [16] analyzed the use of a non-homogeneous Poisson process to model the arrival data of call centers and hospitals.

It is well known that *MAP*-related processes suffer from identifiability problems. Identifiability issues occur when the parameters of a model do not uniquely determine the performance of the process, in the sense that the distribution functions related to the model are the same under at least two model representations. In order to develop an estimation method to fit real data sets to the model, a detailed examination of the identifiability property of the process is crucial. Since the objective is to estimate or identify the true parameters of the model, to do so it is important to find a proper characterization of the process with a reduced number of parameters. Motivated by this fact, the contributions of this work are threefold. (1) A matrix-based technique to build equivalent non-stationary MAP_2s to a given one is presented. The construction of equivalent representations for the stationary MAP_2 was undertaken in Ramírez-Cobo et al. [17], but unlike that work, here we adopt a matrix-based approach. (2) In a similar way as in Bodrog et al. [18], we derive a canonical, unique representation for the process. This assures the unimodality of the likelihood function and avoids the typical switching problems and dependence of the starting point of the likelihood-based statistical inference methods (see Congdon [19]), which reduces the space storage requirements in manipulating many distributions. (3) A characterization of the process in terms of five moments of the inter-event times distribution is given. The last two contributions help to design a fitting approach, as can be seen in [12].

This paper is organized as follows. The non-stationary m-state *MAP* is introduced in Section 2. In Section 3, a matrix-based methodology to construct equivalent non-stationary MAP_2s is derived. In Section 4 a canonical, unique representation of the process is presented. In Section 5 a moments-based characterization of the process, which is of practical applications when estimation of the process is carried out, is provided. Finally, conclusions and possible extensions to this work are given in Section 6.

2. The non-stationary *MAP*

The non-stationary m-state *MAP*, denoted MAP_m, is a matrix generalization of the Poisson process to allow for correlated event occurrences and non-exponential non-identically distributed inter-event times, unlike its stationary version where the inter-event times are considered to be identically distributed.

The MAP_m is a doubly stochastic process $\{J(t), N(t)\}$, where $J(t)$ represents an irreducible, continuous, Markov process with state space $\mathcal{S} = \{1, \ldots, m\}$ and generator matrix D. The process $\{N(t),\ t \geq 0\}$ counts the number of events in the interval $(0, t]$. For a thorough definition of the general m-state *MAP*, we refer the reader to Neuts [2] and Lucantoni et al. [1]. The MAP_m behaves as follows: the initial state $i_0 \in \mathcal{S}$ is generated according to an initial probability vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$ and at the end of an exponentially distributed sojourn time in state i, with mean $1/\lambda_i$, two types of transitions can occur. On one hand, with probability p_{ij0}, $J(t)$ may jump from one state to another state (necessarily different), and no event occur. On the other hand, with probability p_{ij1}, $J(t)$ may jump from one state to another state (possibly the same) and an event occurs. The transition probabilities satisfy

$$\sum_{j=1, j\neq i}^{m} p_{ij0} + \sum_{j=1}^{m} p_{ij1} = 1, \quad \text{for all } i \in \mathcal{S}.$$

Fig. 1 illustrates the evolution of the MAP_2 in terms of the event times. The blue line corresponds to transitions without events, governed by D_0, and the red lines correspond to transitions with an event occurring, governed by D_1. Note that the role of the states in this model is to provide inter-event times distributed as a sum of non-identical, random exponentials.

The non-stationary MAP_m can thus be expressed in terms of $\{\boldsymbol{\alpha}, \boldsymbol{\lambda}, P_0, P_1\}$ where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$, $P_0 = \{p_{ij0}\}_{i\neq j}$ and $P_1 = \{p_{ij1}\}$. Instead of transition probability matrices, any non-stationary MAP_m can be also characterized by $\mathcal{M} = \{\boldsymbol{\alpha}, D_0, D_1\}$