JOURNAL OF APPLIED STATISTICS
2022, VOL. 49, NO. 2, 336–356
https://doi.org/10.1080/02664763.2020.1812544

Bayesian generalizations of the integer-valued autoregressive model

Paulo C. Marques F. [a], Helton Graziadei [b] and Hedibert F. Lopes [a]

[a]Insper Institute of Education and Research, São Paulo, Brazil; [b]Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil

ABSTRACT
We develop two Bayesian generalizations of the Poisson integer-valued autoregressive model. The AdINAR(1) model accounts for overdispersed data by means of an innovation process whose marginal distributions are finite mixtures, while the DP-INAR(1) model is a hierarchical extension involving a Dirichlet process, which is capable of modeling a latent pattern of heterogeneity in the distribution of the innovations rates. The probabilistic forecasting capabilities of both models are put to test in the analysis of crime data in Pittsburgh, with favorable results.

ARTICLE HISTORY
Received 7 February 2020
Accepted 13 August 2020

1. Introduction

Integer-valued time series models are essential inferential tools in areas such as epidemiology, econometrics, environmental studies, and public policy [6]. An important pioneering development was the proposal of the integer-valued autoregressive model (INAR(1) model hereafter) by McKenzie [14] and Al-Osh and Alzaid [1]. In a nutshell, the INAR(1) model defines a time-homogeneous Markovian process for which the counts at one epoch are a random fraction of the counts at the previous epoch plus the counts associated with an independent innovation process with marginal Poisson distributions. In recent decades, a large body of research has been dedicated to modifications and generalizations of this original INAR(1) model [11,17,18,21,22].

In this paper, we develop two Bayesian generalizations of the INAR(1) model. In our first proposal, the adaptive integer-valued autoregressive model (AdINAR(1) model hereafter), we make use of a finite mixture to define the marginal distributions of the innovation process. The motivation for this AdINAR(1) model is the possibility to account for overdispersed time series. Our second contribution is a hierarchical extension of the INAR(1) model, implemented with the help of a Dirichlet process [10] placed at the top of the model hierarchy. In this DP-INAR(1) model, the innovation rates may vary through time, and in their modeling, we benefit from the clustering properties of the Dirichlet process. Our main goal in the paper is to assess the forecasting capabilities of the two new models regarding out-of-sample predictions.

CONTACT Paulo C. Marques F. paulo.c.marques.f@gmail.com