Estimation and forecasting for binomial and negative binomial INAR(1) time series

Stima e previsione di modelli INAR(1) con innovazioni Binomiali e Binomiali Negative

Luisa Bisaglia, Margherita Gerolimetto and Paolo Gorgi

Abstract To model count time series Al-Osh & Alzaid (1987) and McKenzie (1988) introduced the INteger-valued AutoRegressive process (INAR). Usually the innovation term is assumed to follow a Poisson distribution. However, other distributional assumptions may be used instead. In this work we discuss the issue of estimating and forecasting in case of INAR(1) time series with over and under-dispersion, resorting respectively to the Binomial and Negative Binomial distributions. We calculate the maximum likelihood functions for the considered cases and via a Monte Carlo experiment we show that the resulting estimators have a good performance. Moreover, we also concentrate on the problem of producing coherent predictions based on estimates of the p-step ahead predictive mass functions assuming Binomial and Negative Binomial distributions of the error term. Finally, we compare the forecast accuracy of Binomial and Negative Binomial INAR with that of Poisson INAR and ARMA models with a Monte Carlo experiment.

Abstract *Per modellare serie storiche per dati di conteggio Al-Osh & Alzaid (1987) and McKenzie (1988) hanno introdotto i processi AR a valori interi (INAR). In generale, la distribuzione del termine d'errore si assume essere Poisson, tuttavia, allo scopo di poter trattare adeguatamente anche casi di sotto e sovra-dispersione, in questo lavoro verranno considerate le distribuzioni Binomiale e Binomiale Negativa. Più specificamente, sotto queste ipotesi distribuzionali, verranno dapprima sviluppate le procedure di stima e previsione, quindi, attraverso un esperimento di Monte Carlo, ne verranno valutate le relative performance.*

Key words: COUNT TIME SERIES, INAR(1) PROCESS, FORECASTING

L. Bisaglia
Dept. of Statistical Science, University of Padua, e-mail: bisaglia@stat.unip.it

M. Gerolimetto
Dept. of Economics, University of Ca' Foscari, Venice, e-mail: margherita.gerolimetto@unive.it

P. Gorgi
Dept. of Statistical Science, University of Padua, e-mail: paolo.gorgi@unipd.it

1 Introduction

Recently, there has been a growing interest in studying nonnegative integer-valued time series and, in particular, time series of counts. The most common approach to build an integer-valued autoregressive processes is based on a probabilistic operation called (binomial) thinning, as reported in Al-Osh & Alzaid (1987) and McKenzie (1988) who first introduced integer-valued autoregressive processes (INAR). While theoretical properties of INAR models have been extensively studied in the literature, relatively few contributions discuss the development of forecasting methods that are coherent, in the sense of producing only integer forecasts of the count variable. Freeland & McCabe (2004), in the context of INAR(1) process with Poisson innovations suggest the use of the median of the $k-$step-ahead conditional distribution to emphasize the intention of preserving the integer structure of the data in generating the forecasts. The Poisson distribution, however, has the disadvantage to take into account only equi-dispersion. Thus, unlike the usual applications where the error terms are usually Poisson, in oder to investigate also over and under-dispersion cases, we assume the error terms to follow the Binomial or Negative Binomial distributions.

The purpose of this work is, firstly, to calculate the maximum likelihood (ML) functions and consequently the estimators of the parameters of the Binomial and Negative Binomial INAR(1) (respectively BINAR(1) and BNINAR(1)) models, secondly, to obtain the predictive probability mass function (pmf) in order to derive coherent predictions for the considered models.

The remainder of the paper is divided as follows. Section 2 recalls INAR(1) model. Section 3 provides the theoretical results in order to obtain the ML function and the predictive pmf. Section 4 describes our simulation study to show the performance of the estimators and to compare the forecast accuracy of BINAR and BNINAR models with that of Poisson INAR (PoINAR) and ARMA ones.

2 The INAR(1) model

To introduce the class of INAR model we first recall the thinning operator, '$\circ$', defined as follows.

Definition *Let Y be a non negative integer-valued random variable, then for any $\alpha \in [0,1]$*

$$\alpha \circ Y = \sum_{i=1}^{Y} X_i$$

where X_i is a sequence of iid count random variables, independent of Y, with common mean α.

The INAR(1) process $\{Y_t; t \in \mathbf{Z}\}$ is defined by the recursion

$$Y_t = \alpha \circ Y_{t-1} + \varepsilon_t \tag{1}$$

where $\alpha \in [0,1]$, and ε_t is a sequence of iid discrete random variables with finite first and second moment. The components of the process $\{Y_t\}$ are the surviving elements of the process Y_{t-1} during the period $(t-1,t]$, and the number of elements which entered the system in the same interval, ε_t. Each element of Y_{t-1} survives with probability α and its survival has no effect on the survival of the other elements, nor on ε_t which is not observed and cannot be derived from the Y process in the INAR(1) model. If the error terms are distributed as a $Po(\lambda)$, the marginal distribution of the observed counts is a Poisson distribution with (unconditional) mean and variance equal to $\lambda/(1-\alpha)$ (Jung et al., 2005). Clearly, this model does not allow for under or over-dispersion in data and for this reason in the next Section we will consider Binomial and Negative Binomial distribution instead.

3 Theoretical results

Theoretical properties of PoINAR models have been discussed extensively in the literature. In this Section we extend the results of the PoINAR(1) model to the BINAR(1) and NBINAR(1) models.

3.1 Maximum Likelihood function for BINAR(1) and NBINAR(1)

We begin assuming that the distribution of the innovation term is $Bi(m,p)$. Differently from the PoINAR model, now the marginal distribution of Y_t is not known and the distribution of Y_0 could not be obtained explicitly. Thus we calculate the conditional likelihood function $L(\alpha,m,p|y_0) = \prod_{t=1}^{n} p(Y_t|Y_{t-1})$ and the log-likelihood that is:

$$\begin{aligned} l(\alpha,m,p|y_0) &= \log(L(\alpha,m,p|y_0)) = \sum_{t=1}^{n} \log(p(Y_t|Y_{t-1})) = \\ &= \left(\sum_{t=1}^{n} y_{t-1}\right)\log(1-\alpha) + mn\log(1-p) + \left(\sum_{t=1}^{n} y_t\right)\log\left(\frac{p}{(1-p)}\right) + \\ &+ \sum_{t=1}^{n}\log\left(\sum_{i=Max(0,y_t-m)}^{Min(y_t,y_{t-1})}\binom{y_{t-1}}{i}\binom{m}{y_t-i}\left(\frac{\alpha(1-p)}{p(1-\alpha)}\right)^i\right) \end{aligned} \tag{2}$$

The ML estimate are the values of α, m and p that maximizes the (2).

Now, we assume that the distribution of the innovation term is $NB(r,p)$. In this case it is possible to show that the conditional log-likelihood is:

$$
\begin{aligned}
l(\alpha,r,p|y_0) &= \log(L(\alpha,r,p|y_0)) = \sum_{t=1}^{n}\log(p(Y_t|Y_{t-1})) = \\
&= \left(\sum_{t=1}^{n} y_{t-1}\right)\log(1-\alpha) + nr\log(p) + \left(\sum_{t=1}^{n} y_t\right)\log(1-p) + \\
&+ \sum_{t=1}^{n}\log\left(\sum_{i=0}^{\min(y_t,y_{t-1})}\binom{y_{t-1}}{i}\frac{\Gamma(r+y_t-i)}{\Gamma(r)(y_t-i)!}\left(\frac{\alpha}{(1-p)(1-\alpha)}\right)^{i}\right)
\end{aligned}
\tag{3}
$$

The ML estimates are the values of α, r and p that maximizes the (3).

3.2 k-step ahead predictive probability mass function

Freeland & McCabe (2004) define the $p-$step ahead predictive probability mass function (pmf) for the PoINAR(1) model as:

$$
\begin{aligned}
P_k(Y_{T+k}=y \mid Y_T) = \; & \textstyle\sum_{s=0}^{\min(y,Y_T)}\binom{Y_T}{s}(\alpha^k)^s(1-\alpha^k)^{Y_T-s}\times \\
& \tfrac{1}{(y-s)!}\exp\left\{-\lambda\tfrac{1-\alpha^k}{1-\alpha}\right\}\times\left(\lambda\tfrac{1-\alpha^k}{1-\alpha}\right)^{y-s}
\end{aligned}
\tag{4}
$$

Consequently, $P_k(Y_{T+k}=y \mid Y_T)$ is the probability of the value y occurring, according to the $k-$step-ahead conditional distribution. In order to obtain coherent predictions for Y_{T+k}, Freeland & McCabe (2004) suggest using the median of the $k-$step-ahead pmf. In practice we compute $P_k(Y_{T+k}=y \mid Y_T,\hat{\alpha},\hat{\lambda})$, so typically (α,λ) are estimated via ML.

Unlike the PoINAR model, for the BINAR and BNINAR models it is not possible to derive a general form for the $p-$step ahead predictive probability mass function. However, it is possible to calculate it for some values of p and in particular, we obtain the expression of the conditional probabilities for BINAR and BNINAR models in case of p=1, 2. Here, for space reasons, we report only the expression relatively to the case of p=1 for the BNINAR:

$$
\begin{aligned}
p(Y_{t+1}|Y_t) = (1-\alpha)^{y_t}p^r(1-p)^{y_{t+1}} & \sum_{i=0}^{\min(y_{t+1},y_t)}\binom{y_t}{i} \\
& \times\left(\frac{\Gamma(r+y_{t+1}-i)}{\Gamma(r)(y_{t+1}-i)!}\right)\left(\frac{\alpha}{(1-p)(1-\alpha)}\right)^{i}
\end{aligned}
\tag{5}
$$

4 Monte Carlo experiments

In this Section we provide the details of the Monte Carlo experiment we carried out. We generate data from INAR(1) DGP's with Poisson errors, Binomial errors and Negative Binomial errors for different values of parameters and sample sizes. For each model we generated $s = 1000$ independent realizations. To compare the forecasting performance of the estimated models, we computed the Forecast Mean Square Error (FMSE) and Forecast Mean Absolute Error (FMAE) statistics of $k-$step-ahead forecasts, where $k = 1, 2$, and the Total variation and Bhattacharya distance (BC) between the true distribution $h-$ steps ahead and the predicted one.

For lack of space we report only some results for the BNINAR(1) case. As we can notice from Table 1 and Figure 1, the Negative Binomial hypothesis outperforms the Poisson hypothesis and the same holds for the Binomial hypothesis.

Table 1 *Total variation distance* and *Bhattacharyya distance* between the true distribution $h-$ steps ahead and the predicted one, calculated for N=1000 simulations for data generated from INAR(1) process with Negative Binomial innovations. The best results are in bold.

Total variation distance							
Parameters	Model	n=50		n=100		n=300	
		$h=1$	$h=2$	$h=1$	$h=2$	$h=1$	$h=2$
$\alpha=0.2$, $r=1.3$	Negative Bin	**0.085**	**0.080**	**0.060**	**0.056**	**0.034**	**0.031**
$p=0.3$	Poisson	0.262	0.261	0.261	0.259	0.257	0.254
$\alpha=0.5$, $r=1.3$	Negative Bin	**0.087**	**0.091**	**0.060**	**0.064**	**0.033**	**0.036**
$p=0.3$	Poisson	0.217	0.231	0.209	0.223	0.203	0.218
$\alpha=0.8$, $r=1.3$	Negative Bin	**0.086**	**0.096**	**0.061**	**0.067**	**0.034**	**0.037**
$p=0.3$	Poisson	0.171	0.185	0.165	0.177	0.161	0.170
$\alpha=0.2$, $r=12$	Negative Bin	**0.076**	**0.071**	**0.053**	**0.050**	**0.033**	**0.031**
$p=0.8$	Poisson	0.080	0.077	0.065	0.063	0.053	0.052
$\alpha=0.5$, $r=12$	Negative Bin	**0.073**	**0.078**	**0.053**	**0.058**	**0.031**	**0.034**
$p=0.8$	Poisson	**0.073**	0.079	0.057	0.064	0.042	0.047
$\alpha=0.8$, $r=12$	Negative Bin	0.074	0.084	0.052	0.060	**0.031**	**0.035**
$p=0.8$	Poisson	**0.070**	**0.081**	**0.051**	**0.059**	0.034	0.039

Bhattacharyya distance							
Parameters	Model	n=50		n=100		n=300	
		$h=1$	$h=2$	$h=1$	$h=2$	$h=1$	$h=2$
$\alpha=0.2$, $r=1.3$	Negative Bin	**0.0077**	**0.0071**	**0.0037**	**0.0032**	**0.0012**	**0.0011**
$p=0.3$	Poisson	0.0660	0.0658	0.0635	0.0630	0.0617	0.0615
$\alpha=0.5$, $r=1.3$	Negative Bin	**0.0089**	**0.0101**	**0.0043**	**0.0050**	**0.0013**	**0.0015**
$p=0.3$	Poisson	0.0491	0.0545	0.0471	0.0523	0.0427	0.0480
$\alpha=0.8$, $r=1.3$	Negative Bin	**0.0086**	**0.0111**	**0.0040**	**0.0050**	**0.0013**	**0.0017**
$p=0.3$	Poisson	0.0314	0.0377	0.0300	0.0357	0.0276	0.0322
$\alpha=0.2$, $r=12$	Negative Bin	**0.0064**	**0.0055**	**0.0037**	**0.0030**	**0.0012**	**0.0010**
$p=0.8$	Poisson	0.0073	0.0064	0.0050	0.0043	0.0029	0.0029
$\alpha=0.5$, $r=12$	Negative Bin	0.0073	0.0084	**0.0036**	**0.0042**	**0.0012**	**0.0014**
$p=0.8$	Poisson	**0.0070**	**0.0083**	0.0040	0.0047	0.0020	0.0024
$\alpha=0.8$, $r=12$	Negative Bin	0.0064	0.0085	**0.0030**	**0.0039**	**0.0011**	**0.0015**
$p=0.8$	Poisson	**0.0060**	**0.0083**	0.0031	0.0042	0.0015	0.0018

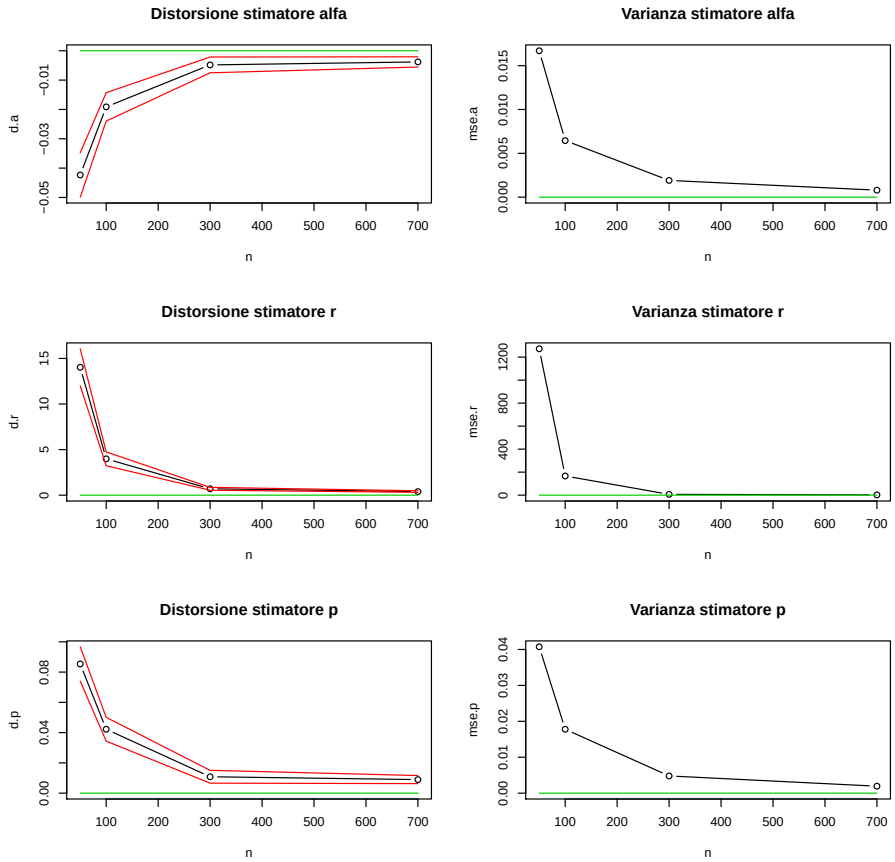

Fig. 1 DGP: NBINAR(1). On the left side of the panel the bias of $\hat{\alpha}$ and m (with the increase of the sample size) is reported, on the right side the variance (confidence level 95%, 1000 simulations with $\alpha = 0.5, r = 6, p = 0.5$).

References

Al-Osh, M. A. & Alzaid, A. A. (1987). First order integer-valued autoregressive INAR(1) process. *Journal of Time Series Analysis*, **8**(3), 261–275.

Freeland, R. K. & McCabe, B. P. M. (2004). Forecasting discrete valued low count time series. *International Journal of Forecasting*, **20**, 427–434.

Jung, R. C., G., R., & Tremayne, A. R. (2005). Estimation in conditional first order autoregression with discrete support. *Statistical Papers*, **46**, 195–224.

McKenzie, E. (1988). Some ARMA models for dependent sequences of poisson counts. *Advances in Applied Probability*, **20**, 822–835.