MODELLING TIME SERIES OF COUNTS:
AN INAR APPROACH

MARIA EDUARDA SILVA

ABSTRACT. Time series of counts arise when the interest lies on the number of certain events occurring during a specified time interval. Many of these data sets are characterized by low counts, asymmetric distributions, excess zeros, over dispersion, ruling out normal approximations. Several approaches and diversified models that explicitly account for the discreteness of the data have been considered in the literature, among which are the INteger-valued AutoRegressive, INAR, models. These models are based on random operations which, operating on discrete variables ensure an integer-valued result. The INAR models are attractive since they are linear-like models for discrete time series and exhibit recognizable correlation structures. This paper considers INAR models for analyzing time series of counts and discusses associated statistical inference, comprising estimation, diagnostics and model assessment.

1. INTRODUCTION

The problem of modelling time series of low counts has attracted many researchers over the last few decades. In fact time series of counts arise in many different contexts, usually as counts of certain events or objects in specified time intervals as for example: social science [31, 37], queueing systems [1], experimental biology [54], environmental processes [48, 11, 24], economics and finance [6, 39, 15, 28, 27], epidemiology [10, 53], international tourism demand [8, 17, 7], statistical control processes [50, 52], telecommunications [30], optimal alarm systems [34], and in the biopharmaceutical industry [3].

In many cases, the discrete variates are large numbers and it may make sense to approximate them by continuous variates. Often, however, this is not possible or even desirable and it is necessary to develop an appropriate modelling strategy for the statistical analysis of time series of counts. One of the approaches developed is based on a random operation called *thinning operation* capable of preserving the integer valued nature of the variables, giving rise to the class of INteger valued AutoRegressive, INAR, models. This paper considers the first order INAR, INAR(1), models for analyzing time series of counts and discusses associated statistical inference, comprising estimation, diagnostics and model assessment. The plan of the paper is as follows. Section 2 introduces the INAR(1) models with including discussion of the most relevant properties. Section 3 considers parameter estimation and presents a set of tools appropriate to check the adequacy of fitted INAR models, an important part of any iterative modelling exercise in applied time series analysis. Section 4 illustrates the fitting of the models to a time series of counts of the stock type. Section 5 provides some concluding remarks.

2. First order INteger valued AutoRegressive models

Definition 2.1. *The first order integer autoregressive, INAR(1), model is defined on the discrete support $\mathbb{N}_0$ by the recursive equation*

$$(2.1) \qquad\qquad\qquad X_t = \alpha \diamond X_{t-1} + \epsilon_t$$

where $\{\epsilon_t\}$ is a sequence of independent and identically distributed non-negative integer valued random variables, for each t independent of X_{t-1} and of $\alpha \diamond X_{t-1}$, with finite mean μ_ϵ and variance σ_ϵ^2 and, conditional on X_{t-1}, $\alpha \diamond X_{t-1}$ is an integer valued random variable whose probability distribution depends on the parameter α.[1]

Thus, $'\diamond'$[2] denotes a random operator, usually called thinning operator, which always produces integer values and introduces serial dependence via the conditioning on X_{t-1}. Consider now that in model (2.1) we require that the marginal distribution of $\{X_t\}_t$ is of the same family as $\{\epsilon_t\}$. [25] proposes an approach to solve this problem within the convolution-closed infinitely divisible class of (marginal) distributions. The random operator $'\diamond'$ is required not only to introduce serial dependence and preserve the integer-valued status of the random variable but also to be unconditionally of the same family as $\{\epsilon_t\}_t$.

---

[1]In fact, α may be a vector of parameters but at this point we prefer the simpler scalar notation

[2]The operator is in fact $\alpha \diamond$ as it depends on α but usually the simpler notation is used.

The intuition behind the operator $'\diamond'$, as described by that author is the following.

Let F_θ denote a convolution-closed infinitely divisible parametric family such that $F_{\theta_1} * F_{\theta_2} = F_{\theta_1 + \theta_2}$, with $'*'$ denoting the convolution operator. Let Y_1, Y_2 be independent random variables each with distribution $F_{(1-\alpha)\theta}$ (pmf $f_{(1-\alpha)\theta}(\cdot)$) and Y_{12} be a another random variable independent from Y_1, Y_2 with distribution $F_{\alpha\theta}$. The distribution of Y_{12} given $Y_{12} + Y_1 = y$ is denoted by $G_{\alpha\theta,(1-\alpha)\theta,y}$ and its pmf by $g(\cdot|y)$. [25] writes the joint distribution of (X_t, X_{t-1}) as being the same as that of $(Y_{12}+Y_2, Y_{12}+Y_1)$, in which case Y_{12} represents a common latent or unobserved component of the pair (X_t, X_{t-1}) that carries the dependence of the observations between two consecutive time periods and $Y_i, i = 1, 2$ represent the arrivals in the model.

[25] shows that the processes defined in (2.1) are Markov order 1, time reversible and stationary with non-negative serial dependence, $\rho_k = \alpha^k$. The transition probabilities are given by

$$\mathrm{P}(X_t = k | X_{t-1} = l) = \sum_{j=0}^{\min\{k,l\}} g(j|l)\mathrm{P}(\epsilon_t = k - j)$$

(2.2)
$$= \sum_{j=0}^{\min\{k,l\}} g(j|l)f_{(1-\alpha)\theta}(k - j).$$

Moreover, the INAR(1) model is a member of the class of conditional linear first order autoregressive, CLAR(1), models introduced by [20].

2.1. **The Poisson INAR(1).** Consider the case where F_θ is Poisson $\theta = \lambda/(1 - \alpha)$. Setting $G_{\alpha\theta,(1-\alpha)\theta,x}$ as Binomial(x, α), leads to the most common thinning operator which is the binomial thinning, denoted by $'\circ'$ and originally introduced by [46] to extend the notions of self-decomposability (DSD) and stability to integer-valued time series.

Definition 2.2. *Let X be a non-negative integer-valued random variable. Then, for any $\alpha \in [0, 1]$ define the binomial thinning operator as*

(2.3)
$$\alpha \circ X := \sum_{i=1}^{X} Y_i,$$

where $\{Y_i\}_i$ is a sequence of independent and identically distributed Bernoulli random variables with $\mathrm{P}(Y_i = 1) = \alpha$, called the counting series of $\alpha \circ X$, which is also independent of X.

For properties of the binomial thinning operation see [47, 51, 43, 44].

The binomial thinning based INAR(1) model was originally proposed by [2] and [32]. The conditional distribution of X_t given X_{t-1}, $f_{X_t|X_{t-1}}(x_t|x_{t-1}) = p(X_t|X_{t-1})$ is now the convolution of the two components, binomial and Poisson, as follows:

$$(2.4) \qquad p(X_t|X_{t-1}) = \sum_{i=0}^{M_t} \binom{X_{t-1}}{i} \alpha^i (1-\alpha)^{X_{t-1}-i} \frac{e^{-\lambda} \lambda^{X_t-i}}{(X_t-i)!},$$

where $M_t = \min(X_{t-1}, X_t)$. The bivariate pgf of the PoINAR(1) process, given by [4]

$$(2.5) \quad P_{X_t, X_{t-1}}(s_1, s_2) = P_{X_{t-1}}(s_1(1-\alpha-\alpha s_2)) P_{\epsilon_t}(s_2)$$
$$= \exp\left\{-\frac{\lambda}{1-\alpha}[2-\alpha-(1-\alpha)(s_1+s_2)-\alpha s_1 s_2]\right\}$$

is symmetric in its arguments s_1 and s_2 and therefore the process is time-reversible. The PoINAR(1) may be interpreted as an infinite server queue. The service time is geometric with parameter $1-\alpha$ and the arrival process is Poisson with mean λ. A fundamental result in queueing theory, Little Flow's equation, states that the expected length of the queue is equal to the arrival rate times the expected waiting time. It is thus possible to compute the expected number of time units a newly arrival stays in the system and which is given by: $1/(1-\alpha)$.

3. Parameter Estimation and Diagnostic Tools

3.1. **Parameter estimation.** This section considers the estimation of the parameters in the INAR(1) models discussed previously. Estimation can be carried out in several ways leading to the following broad categories of estimators: moment based estimators (MM), regression based or conditional least squares (CLS) estimators and likelihood based (ML) estimators. All these approaches have been considered in detail in the literature for the Poisson model. Additionally Bayesian methodology has been considered by [36] and [45]. However, the most common approach for the estimation of the INAR(1) model is maximum likelihood method.

Let $\mathbf{x} = (X_1, \ldots, X_n)$ represent the observed time series and $\boldsymbol{\theta}$ the $s \times 1$ vector of model parameters to be estimated. Since the INAR(1) model (2.1) is a first order stationary Markov chain, the likelihood function is written as

$$(3.1) \qquad L_n(\boldsymbol{\theta}|\mathbf{x}) = P(X_1 = x_1) \prod_{t=2}^{n} f_{X_t|X_{t-1}}(x_t|x_{t-1})$$

where $f_{X_t|X_{t-1}}(x_t|x_{t-1})$ is given by (2.2) and $P(X_1 = x_1)$ represents the stationary marginal distribution. For asymptotic analysis the standard approach is to consider the conditional log-likelihood function given by

$$(3.2) \qquad l_n(\boldsymbol{\theta}) = \sum_{t=2}^{n} \log\left(f_{X_t|X_{t-1}}(x_t|x_{t-1})\right).$$

Based on the results of [5] for Markov processes, [25, 26] prove the following result

Theorem 3.1. *The maximum likelihood estimators, MLE, of the parameters $\boldsymbol{\theta}$ in INAR(1) models are consistent, asymptotically normal and asymptotically efficient.*

The proof requires that some regularity conditions hold, ([26], pp. 318).

The limit distribution of $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is $N(0, \boldsymbol{\Sigma}_{MLE})$, $\boldsymbol{\Sigma}_{MLE} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = (\sigma_{ij}(\boldsymbol{\theta}))$ is a non-singular $s \times s$ matrix with elements

$$\sigma_{ij}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}\left(\dot{l}_i(\boldsymbol{\theta}; x_1, x_2)\dot{l}_j(\boldsymbol{\theta}; x_1, x_2)\right)$$

In general, for numerical maximum likelihood estimation, a quasi-Newton method can be used with an input of the negative log-likelihood function and output of the MLE and inverse Hessian matrix at the MLE. The initial estimates required by the optimization algorithm are based on the method of moments. The inverse Hessian evaluated at the maximum, $\hat{\boldsymbol{\theta}}$, can be used as the estimated variance-covariance matrix of the ML estimator $\hat{\boldsymbol{\theta}}$.

3.2. **Diagnostic tools.** A crucial step in any statistical investigation is the assessment of the adequacy of the models proposed and fitted to the data under analysis. Various methods for model validation and diagnostics in discrete-valued time series have been proposed in the literature. These methods can be broadly classified as: parametric resampling methods; residual based methods; methods based on the predictive distributions; model comparisons using scores and information criteria.

3.2.1. *Parametric resampling methods.* [49] proposes a procedure based on parametric bootstrap and special functionals designed to show the specific features of interest. Specifically, the fitted model is used to generate many samples, all with the same number of observations as the original data set. The samples generated are then used to construct an empirical distribution of the functional of interest. If the fitted model is adequate in describing the feature of interest, the functional quantity of the original data should be a reasonable point with respect to the empirical distribution. The functional of interest may be

a spectral density function or the autocorrelation function. Here, the autocorrelation properties are of interest. Thus M artificial data sets, with the same length of the original data set, are generated from the fitted model. Based on these, M sample autocorrelation functions (ACF) are obtained. For each fixed lag of the ACF, the $(1 - \alpha/2)$ and $\alpha/2$ quantiles of the empirical distribution, denominated acceptance bounds, are computed. Then, a probability interval is obtained for each lag and an envelope is obtained for the ACF. If the fitted model is adequate at each lag, the sample ACF of the original data should be largely within the envelope. Plotting the envelope and the sample ACF of the data jointly, gives rise to a graphical display that can be used to assess the overall goodness of fit of the fitted model with respect to the serial correlation properties. A model is considered to be adequately reproducing the correlation structure of the data, if the sample ACF of the observed data lies within these acceptance bounds. Note that since the sample ACF at different lags are correlated, the acceptance envelope considered is not a joint $100(1 - \alpha)\%$ confidence interval of the sample ACF.

3.2.2. *Residual based methods.* The dynamic structure in the mean and dispersion properties may be checked using tools based on the Pearson residuals defined by

$$(3.3) \qquad\qquad r_t = \frac{X_t - \mathrm{E}(X_t|X_{t-1})}{\mathrm{Var}(X_t|X_{t-1})^{1/2}},$$

where the population quantities are replaced by their estimated counterparts. If the model is correctly specified, these residuals should exhibit mean zero and variance one and no (significant) serial correlation.

However, the structure of the INAR(1) model suggests additional residuals checks. In fact, the INAR(1) model maybe seen as a structural model in the sense that it considers the data to be composed of a set of unobserved components each of which captures a feature of the data: the first component $\alpha \diamond X_{t-1}$ specifies the random number departures or its complement the random number of survivors from the past while ϵ_t represents the new arrivals at the system at time t. This interpretation leads to a residual decomposition that allows to check the adequacy of each component. For details on the residual decomposition and subsequent testing procedures see [16].

3.2.3. *Methods based on the predictive distributions.* A useful tool to check the adequacy of the distributional assumptions of the models is a suitably specified and modified version of the probability integral transform, PIT, originally proposed by [40]. This device has been used in the assessment of predictive distributions of a continuous type by [40], [14] and recently by [18] and [19]. The

PIT has a uniform distribution when the underlying model is continuous. For discrete valued variables the associate distribution functions are step functions, therefore adjustments are necessary. Several authors propose a *randomized PIT* obtained by perturbing the step function nature of the distribution function of the discrete random variables. Additionally, [12] introduce a nonrandomized version of the PIT suitable to count data. For details and references see [12].

Further evaluation of the model based on its predictive performance may be carried out using scoring rules suggested by [12] and [28].

3.2.4. *Information criteria.* Akaike information criterion, AIC and its many variants has been one of the most popular tools for model selection in time series analysis. In the context of time series of counts, [41] studies an automatic criterion for selecting the order of an INAR(p) model based on the corrected version of Akaike Information Criterion, AICC of [22]. Some authors have used AIC as means of choosing between non nested models for time series of counts, regardless of the lack of studies concerning the performance of the criterion in this framework. Moreover, [38] examine the ability of widely used information criteria such as AIC, BIC and the Hannan-Quinn criterion (HQ) [21] to distinguish between some nonlinear times series models that have been popular with practitioners. After performing an extensive simulation study they argue that all three criteria have a useful role to play in a time series model selection exercise.

4. Illustration

This section illustrates the modelling procedure with a data set consisting of the number of different IP addresses accessing the server of the pages of the Department of Statistics of the University of Würzburg in two-minute periods from 10 am to 6 pm on the 29th November 2005, in a total of 241 observations. This data set was originally studied by [50] and exhibits small but significant autocorrelation as indicated by figure 1. The sample mean and variance $\overline{x} = 1.31$ and $\hat{\sigma}^2 = 1.39$ do not indicate overdispersion.

Fitting a PoINAR(1) model to the data yields the CML estimates $\hat{\alpha} = 0.24(0.00)$ and $\hat{\lambda} = 1.01(0.01)$. The parametric bootstrap exercise with $M = 1000$ and the residual analysis represented in figure 2 indicate that the model captures the dynamics of the data. In fact, the variance of Pearson residuals is 1.05 and the acf of the component residuals in (c) indicate that the residuals are white noise. However, figure 2(b) indicates that at time $t = 224$ the residual is unusually large with a large arrival component. This may suggest the occurrence of an additive outlier meaning that $X_{t=224}$ may be contaminated by an
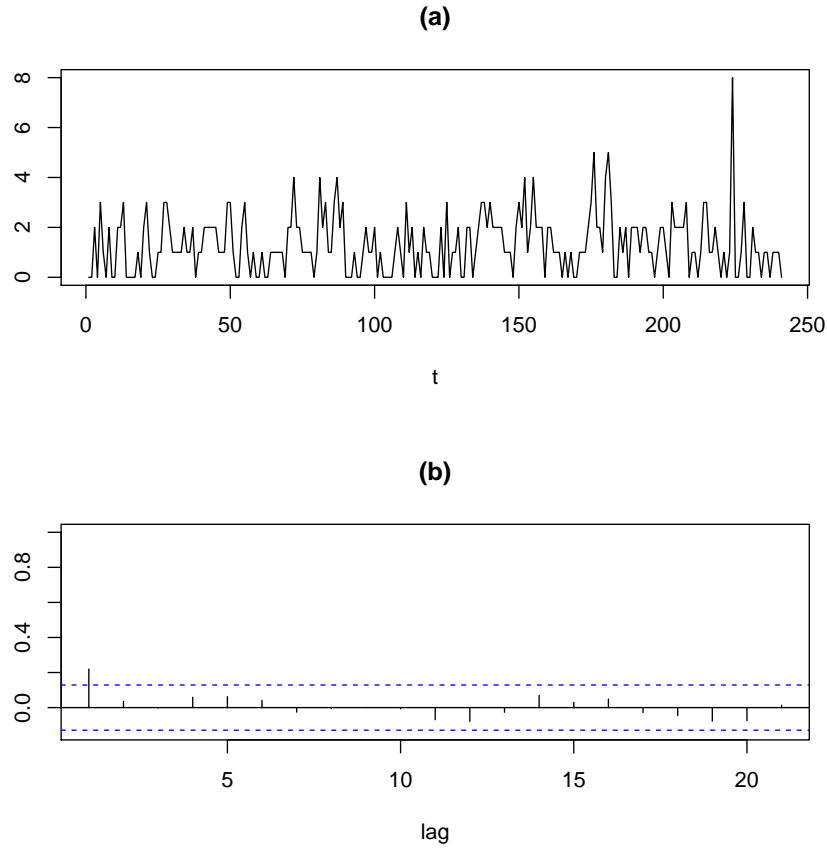
(a)

(b)

FIGURE 1. Number of different IP adresses accessing the server of the pages of the Department of Statistics of the University of Würzburg between 10 am and 6 pm on 29 November 2005 (a) and corresponding sample autocorrelation (b).

exogenous source but the effect is not carried over to subsequent observations by the dynamics.

[42] propose a Bayesian approach to model such outliers assuming that the observed process Y_t is obtained from the unobservable clean process X_t contaminating each X_t with probability δ_t with an outlier of random size η_t. Thus
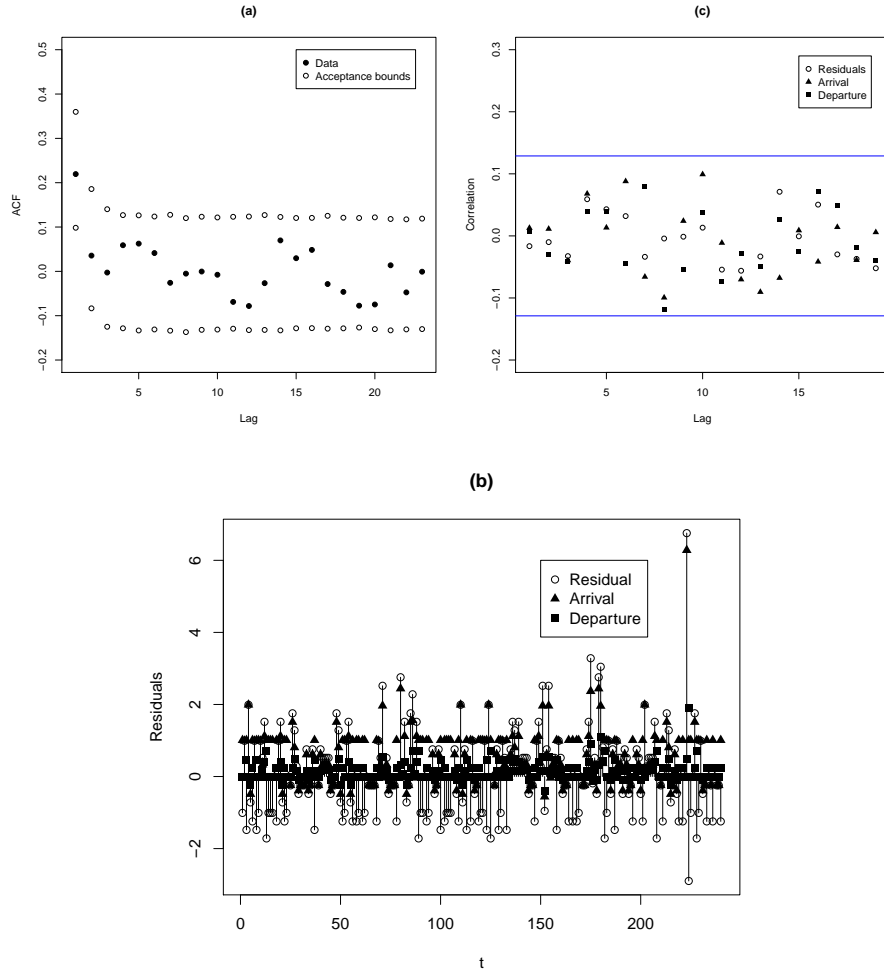
FIGURE 2. Parametric bootstrap exercise (a), component residuals (b) and corresponding autocorrelations (c) for IP data set.

$$Y_t = X_t + \eta_t \delta_t,$$

(4.1)   with $X_t = \alpha \diamond X_{t-1} + e_t$ and $\delta_t \sim Be(p_t)$

where $\delta_1, \eta_1, \ldots, \delta_n, \eta_n$ are independent and independent of the latent process X_t and η_t, the random size of the outlier at time t is a random variable with the same support as X_t and mean $\beta : \eta_t \sim Po(\beta)$. The Bayesian approach to estimate model (4.1) requires apriori distribution for the parameters of interest. For the parameters $0 < \alpha < 1$ and $\lambda > 0$ the traditionally weakly informative priors for the PoINAR(1) ([45]) are chosen: a non-informative Beta prior with parameters $a = 0.01, b = 0.01$ and a a non-informative Gamma prior with parameters $c = 0.01, d = 0.01$, respectively. The prior specifications for p_t, the probability of contamination is a Beta distribution with parameters ($g = 5, h = 95$), with expectation $E(p_t) = 0.05$, reflecting the belief that outliers occur occasionally. The prior for the mean size of the outliers β is a non informative Gamma distribution. The result of applying the outlier detection methodology is represented in figure 3 indicating the occurrence of an outlier at time $t = 224$ with high probability.

The parameter estimates for model (4.1) are $\hat{\alpha}_{Bayes} = 0.27$ and $\hat{\lambda}_{Bayes} = 0.89$, with posterior distributions represented in figure 4 and $\hat{\eta} = 7$ for the size of the outlier, leading to the following model:

$$Y_t = X_t + 7I_{224},$$
(4.2) $$X_t = 0.27 \circ X_{t-1} + e_t, \quad e_t \sim Po(0.89)$$

Figure 5 represents the residuals resulting from the fit of (4.2). Note that the largest residual reduces from 6.8 to 3.3, indicating a better fit. A further indication of the better fit is based on the prediction sum of squares $\sum_{t=2}^{n}(y_t - \hat{y}_t)^2$, where $\hat{y}_t = E(y_t | y_{t-1} = y_{t-1}; \text{parameter estimates})$ which drops from 317.9 to 264.0 when the outlier is included in the model.

5. Final remarks

Time series of counts arise in a wide variety of fields. The need to analyse such data adequately led to a multiplicity of approaches and a diversification of models that explicitly account for the discreteness of the data. One approach is based on the generalized linear models theory for dependent data [29]. Another point of view into the problem is given by parameter driven models which postulate that the observed process is driven by an unobserved process [13]. Yet another approach to the problem of modelling dependent count data is based on the use of renewal processes for generation of a correlated sequence of Bernoulli trials [11]. Here we focused on the INAR(1) models which are a class of observation-driven models particularly suited for stock type data. We illustrated the modelling of a time series with INAR(1) models.
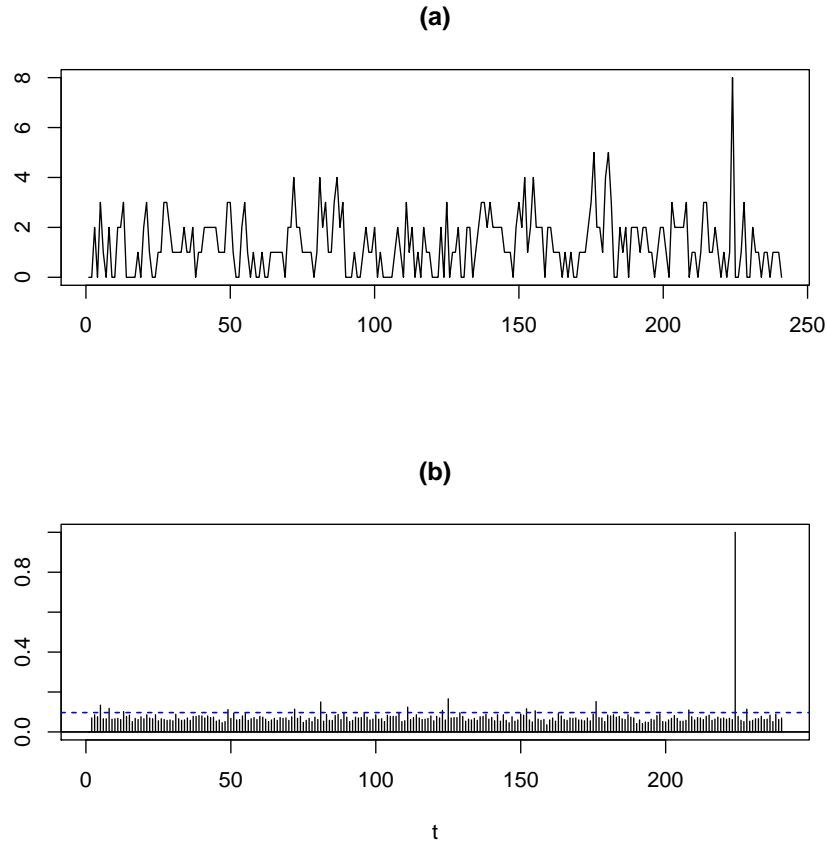
(a)

(b)

t

FIGURE 3. IP time series (a) and posterior probability of outlier occurrence at each time (b).

Several generalizations of the INAR(1) models are available in the literature, namely: INAR(p) models [23, 9], models with moving average components, INARMA, [33], periodic models [35] and bivariate models [37].

REFERENCES

[1] Ahn, S., Lee, G., Jeon, J., 2000. *Analysis of the m/d/1-type queue based on an integer-valued first-order autoregressive process*. Operations Research Letters, 27, 235–241.

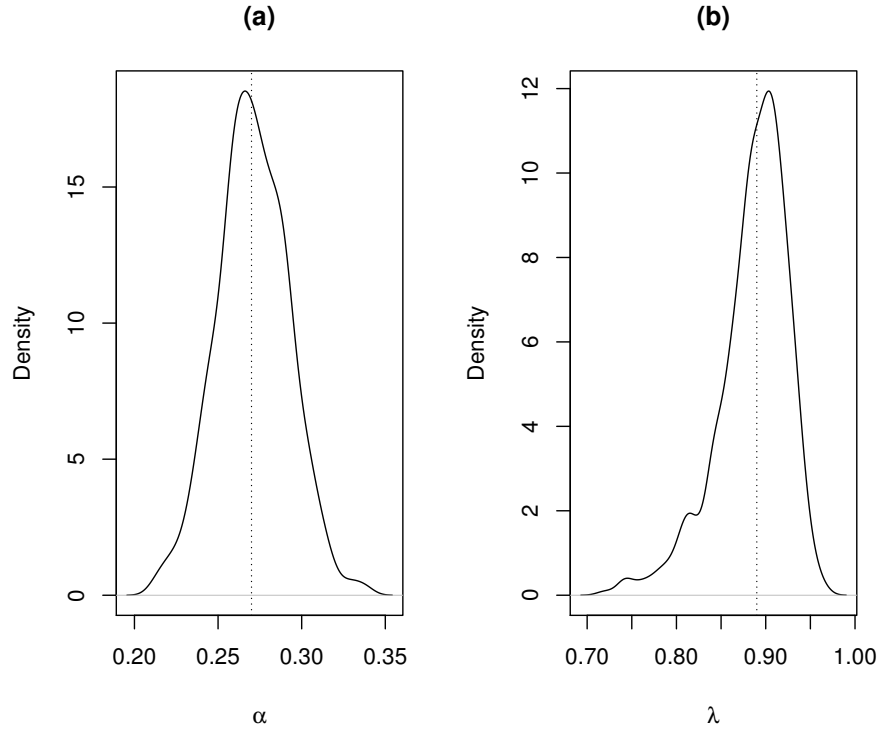**(a)**   **(b)**

FIGURE 4. Posterior distribution of α and λ. The dotted lines
represent the estimates $\hat{\alpha}_{Bayes} = 0.27$ and $\hat{\lambda}_{Bayes} = 0.89$.

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

[3] Alosh, M., 2009. *The impact of missing data in a generalized integer-valued autoregression model for count data.* Journal of Biopharmaceutical Statistics, 19, 1039–1054.

[4] Alzaid, A., Al-Osh, M., 1988. *First-order integer-valued autoregressive (INAR(1)) process: distributional and regression properties.* Statistica Neerlandica, 42, 53–61.

[5] Billingsley, P., 1961. *Statistical Inference for Markov Processes.* University of Chicago Press.

[6] Blundell, R., Griffith, R., Windmeijer, F., 2002. *Individual effects and dynamics in count data models.* Journal of Econometrics, 108, 113–131.

[7] Brännäs, K., Hellström, J., Nordström, J., 2002. *A new approach to modelling and forecasting monthly guest nights in hotels.* International Journal of Forecasting, 18, 19 – 30.
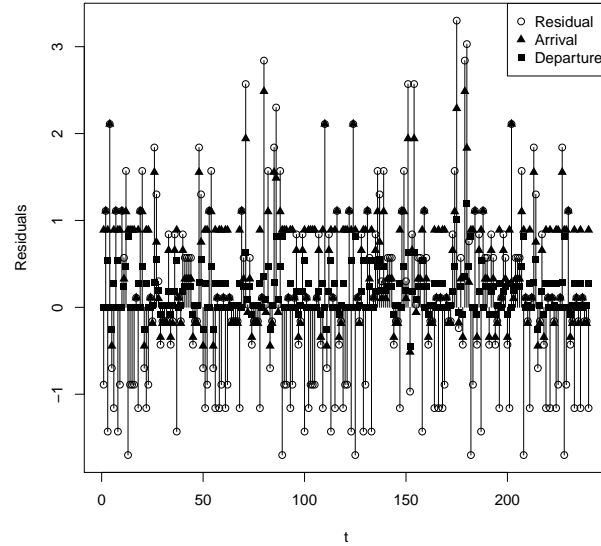
FIGURE 5. IP data set: residuals of a PoINAR(1) with an outlier at time $t = 224$.

[8] Brännäs, K., Nordström, J., 2006. *Tourist accommodation effects of festivals.* Tourism Economics, 12, 291–302.

[9] Bu, B., McCabe, B., 2008. *Maximum likelihood estimation of higher-order integer-valued autoregressive process.* Journal Time Series Analysis, 24, 973–994.

[10] Cardinal, M., Roy, R., Lambert, J., 1999. *On the application of integer-valued time series models for the analysis of disease incidence.* Statistics in Medicine, 18, 2015–2039.

[11] Cui, Y., Lund, R., 2009b. *A new look at time series of counts.* Biometrika, 96, 781–792.

[12] Czado, C., Gneiting, T., Held, L., 2009. *Predictive model assessment for count data.* Biometrics, 65, 1254–1261.

[13] Davis, R., Wu, R., 2009. *A negative binomial model for time series of counts.* Biometrika, 96, 735–749.

[14] Dawid, A., 1984. *Statistical theory: The prequential approach.* Journal of the Royal Statistical Society. Series A, 147, 278–292.

[15] Fokianos, K., Rahbek, A., Tjostheim, D., 2009. *Poisson autoregression.* Journal of the American Statistical Association, 104, 1430–1439.

[16] Freeland, R.K., McCabe, B.P.M., 2004. *Analysis of low count time series data by Poisson autoregression.* Journal of Time Series Analysis, 25, 701–722.

[17] Garcia, G., Ferrer, A., Queralt, 1997. *A note on forecasting international tourism deman in Spain.* International Journal of Forecasting, 13, 539–549.

[18] Gneiting, T., Balabdaoui, F., Raftery, A.E., 2007. *Probabilistic forecasts, calibration and sharpness.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69, 243–268.

[19] Gneiting, T., Raftery, A.E., 2007. *Strictly proper scoring rules, prediction, and estimation.* Journal of the American Statistical Association, 102, 359–378.

[20] Grunwald, G., Hyndman, R., Tedesco, L., Tweedie, R., 2000. *Non-Gaussian conditional linear AR(1) models.* Australian and New Zealand Journal of Statistics, 42, 479–495.

[21] Hannan, E.J., Quinn, B.G., 1979. *The determination of the order of an autoregression.* Journal of the Royal Statistical Society. Series B (Methodological), 41.

[22] Hurvich, C.M., Tsai, C.L., 1989. *Regression and time series model selection in small samples.* Biometrika, 76, 297–307.

[23] Jin-Guan, D., Yuan, L., 1991. *The integer-valued autoregressive INAR(p) model.* Journal of Time Series Analysis, 12, 129–142.

[24] Scotto, M., Weiß, C., Silva, M.E., Pereira, I., 2014. *Bivariate binomial autoregressive models.* Journal of Multivariate Analysis, 125, 233 – 251.

[25] Joe, H., 1996. *Time series model with univariate margins in the convolution-closed infinitely divisible class.* Appl. Probab. ,33, 664–677.

[26] Joe, H., 1997. *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.

[27] Jung, R., Tremayne, A.R., 2011a. *Useful models for time series of counts or simply wrong ones?* AStA Advances in Statistical Analysis, 95, 59–91.

[28] Jung, R.C., Tremayne, A.R., 2011b. *Convolution-closed models for count time series with applications.* Journal of Time Series Analysis, 32, 268–280.

[29] Kedem, B., Fokianos, K., 2002. *Regression Models for Time Series Analysis.* Wiley, Hoboken, NJ.

[30] Lambert, D., Liu, C., 2006. *Adaptive thresholds: monitoring streams of network counts.* Journal of American Statistical Society, 101, 78–88.

[31] McCabe, B., Martin, G., 2005. *Bayesian predictions of low count time series.* International Journal of Forecasting, 21, 315 – 330.

[32] McKenzie, E., 1988. *Some ARMA models for dependent sequences of Poisson counts.* Advances in Applied Probability, 20, 822–835.

[33] McKenzie, E., 2003. *Discrete variate time series,* in: *Handbook of Statistics*, Rao, C., Shanbhag, D. (Eds.), Elsevier Science, Amsterdam, 573–606.

[34] Monteiro, M., Pereira, I., Scotto, M.G., 2008. *Optimal alarm systems for count processes.* Communications in Statistics - Theory and Methods, 37, 3054–3076.

[35] Monteiro, M., Pereira, I., Scotto, M.G., 2010. *Integer-valued autoregressive processes with periodic structure.* Journal Statistical Planning and Inference, 140, 1529–1541.

[36] Silva, Nélia, 2005. *Análise bayesiana de séries temporais de valor inteiro,* Tese de Doutoramento em Matemática, Universidade de Aveiro, Portugal.

[37] Pedeli, X., Karlis, D., 2011. *A bivariate INAR(1) process with application.* Statistical Modelling, 11, 325–349.

[38] Psaradakis, Z., Sola, M., Spagnolo, F., Spagnolo, N., 2009. *Selecting nonlinear time series models using information criteria.* Journal of Time Series Analysis, 30, 369–394.

[39] Quoreshi, A., 2006. *Bivariate time series modeling of financial count data.* Communications in Statistics-Theory and Methods, 35, 1343–1358.

[40] Rosenblatt, M., 1952. *Remarks on a multivariate transformation.* Annals of Mathematical Statistics, 23, 470–472.

[41] Silva, I., 2012. *Analysis of discrete-valued time series.* LAP LAMBERT Academic Publishing, ISBN: 978-3659282706.

[42] Silva, Maria Eduarda, Pereira, Isabel, 2013. *Detection of additive outliers in Poisson INAR(1) time series*, in: *Mathematics of Planet Earth: energy and climate change*, J. P. Bourguignon, R. Jeltsch, A. Pinto, M. Viana (Eds.), 2014, 419–430.

[43] Silva, M.E., Oliveira, V., 2004. *Difference equations for the higher-order moments and cumulants of the INAR(1) model*. Journal of Time Series Analysis, 25, 317–333.

[44] Silva, M.E., Oliveira, V., 2005. *Difference equations for the higher-order moments and cumulants of the INAR(p) model*. Journal of Time Series Analysis, 26, 17–36.

[45] Silva, Isabel, Silva, Maria Eduarda, Pereira, Isabel, Silva, Nélia, 2005. *Replicated INAR(1) Processes*. Methodology and Computing in Applied Probability, 7, 517–542.

[46] Steutel, F., Van Harn, K., 1979. *Discrete analogues of self-decomposability and stability*. The Annals of Probability, 7, 893–899.

[47] K. F. Turkman and M. G. Scotto and P. Zea Bermudez, 2014. *Non-linear Time Series*, Springer-Verlag, Switzerland.

[48] Thyregod, P., Carstensen, N., Madsen, H., Arnbjerg-Nielsen, K., 1999. *Integer-valued autoregressive models for tipping bucket rainfall measurements*. Environmetrics, 10, 395–411.

[49] Tsay, R.S., 1992. *Model checking via parametric bootstraps in time series analysis*. Journal of the Royal Statistical Society. Series C (Applied Statistics), 41, 1–15.

[50] Weiß, C.H., 2007. *Controlling correlated processes of Poisson counts*. Quality and Reliability Engineering International, 23, 741–754.

[51] Weiß, C.H., 2008. *Thinning operations for modelling time series of counts: a survey*. AStA - Advances in Statistical Analysis, 92, 319–341.

[52] Ye, N., Giordano, J., Feldman, J., 2001. *A process control approach to cyber attack detection*. Communications of the ACM ,44, 76–82.

[53] Yu, X., Baron, M., Choudhary, P.K., 2013. *Change-point detection in binomial thinning processes, with applications in epidemiology*. Sequential Analysis, 32, 350–367.

[54] Zhou, J., Basawa, I., 2005. *Least-squares estimation for bifurcating autoregressive processes*. Statistics & Probability Letters, 74, 77 – 88.

Faculdade de Economia, Universidade do Porto & CIDMA

E-mail address: mesilva@fep.up.pt