Proceedings of the 8th WSEAS International Conference on APPLIED MATHEMATICS, Tenerife, Spain, December 16-18, 2005 (pp359-362)

On AR(1) versus MA(1) Models for Non-stationary Time Series of Poisson Counts: Part I (Theory)

VANDNA JOWAHEER
University of Mauritius
Department of Mathematics
Reduit, Mauritius
MAURITIUS

BRAJENDRA C. SUTRADHAR
Memorial University of Newfoundland
Department of Mathematics and Statistics
St. John's, NL, A1C5S7
CANADA

Abstract: Analysis of time series of counts is an important research topic in many bio-medical and socio-economic sectors. For example, analyzing the yearly number of patients of a particular disease in a country is an important problem for health economics. Similarly, analyzing the monthly number of tourists for a city/country and the yearly number of patents awarded to a firm are important economic problems. Unlike in the Gaussian time series case, the analysis of this type of count data is, however, not easy due to the difficulty of modelling the correlated count data recorded over a long period of time. The problem becomes much more difficult if the counts are non-stationary over time, which is likely to be the case in many practical situations. Recently, some authors have developed Gaussian type non-stationary AR(1) (auto-regressive of order 1) models to fit the time series of count data. But, as in practice, there may be situations where Gaussian type moving average (MA) models may fit the count data better than the AR models, this paper develops a non-stationary MA(1) model and compare its basic properties with those of the AR(1) model. For the purpose of statistical inference, the parameters of the proposed models are estimated through an efficient quasi-likelihood (QL) approach.

Key–Words: Consistency, Efficiency, Lag correlations, Non-stationary counts, Observation-driven MA(1) models.

1 Introduction

As opposed to the modelling of non-stationary Gaussian time series data, the modelling of time series of non-stationary counts is not easy. This is because, unlike the Gaussian case, it is not easy to write a multivariate count distribution with a suitable correlation structure. To model the correlation structure of the count data recorded over time, some authors such as Zeger [10] (see also Harvey and Fernandes [4], Davis et al [2]) assume that conditional on a dependent sequence of a stationary Gaussian random effects with auto-correlation structure, the time series data follow independent Poisson distributions so that unconditionally the observations are correlated. Note that in this approach, even though the random effects have a Gaussian correlation structure, they however yield a complicated correlation structure for the count responses.

It is interesting to point out that during the late eighties, some researchers indeed tried to construct a Gaussian type correlation model for the count data. For example, one may refer to McKenzie [6,7] and Sim and Lee [8]. For more recent works in this direction, the readers are referred to Al-osh and Aly [1] and Freeland and McCabe [3], for example. These authors have mainly modelled the stationary count data with auto-regressive order 1 (AR(1)) Gaussian type correlation structure. Note that as the count data in a time series rarely follow stationarity in practice, recently, Mallick and Sutradhar [5] have extended the stationary AR(1) models for the negative binomial counts to the non-stationary case that include the non-stationary Poisson case as a sub-model. In Section 2.1, this non-stationary AR(1) Poisson model is described in brief. Further note that in practice there may be situations where time series of count data may be better fitted by a Gaussian type moving average order 1 (MA(1)) model as compared to the AR(1) type model. For this reason, in this paper, a non-stationary MA(1) model for count data is developed and its basic properties are compared to that of the non-stationary AR(1) model. This is done in the next section, specifically in Sections 2.1 and 2.2. In Section 3, a generalized QL (GQL) approach is discussed for the estimation of the regression and the MA(1) correlation parameters. The paper is concluded in Section 4.

Proceedings of the 8th WSEAS International Conference on APPLIED MATHEMATICS, Tenerife, Spain, December 16-18, 2005 (pp359-362)

2 Non-stationary Counts for Poisson Counts

2.1 Non-stationary AR(1) Models

Let y_t $(t = 1, \cdots, T)$ be the count response recorded at time t and x_t $(t = 1, \cdots, T)$ be the corresponding $p \times 1$ vector of covariates. Further let $\beta = (\beta_1, \cdots, \beta_p)'$ be the p-dimensional vector of regression effects. Suppose that y_1 follows the Poisson distribution with mean parameter $\mu_1 = e^{x_1'\beta}$, that is $y_1 \sim P(\mu_1 = e^{x_1'\beta})$. For $t = 2, \ldots, T$, one may follow the stationary model due to McKenzie [7] (see also Sutradhar [9]), and write the relationship of y_t with y_{t-1} as

$$y_t = \rho * y_{t-1} + d_t, \tag{1}$$

but unlike the stationary case it is assumed in (1) that $y_{t-1} \sim P(\mu_{t-1})$, and $d_t \sim P(\mu_t - \rho\mu_{t-1})$, with $\mu_t = e^{x_t'\beta}$ (see Mallick and Sutradhar [5]). Here d_t and y_{t-1} are independent. Also in (1), for given count y_{t-1}, $\rho * y_{t-1} = \sum_{j=1}^{y_{t-1}} b_j(\rho)$, where $b_j(\rho)$ stands for a binary variable with $pr(b_j(\rho) = 1) = \rho$ and $pr(b_j(\rho) = 0) = 1 - \rho$. This operation in (1), i.e., $\rho * y_{t-1}$ is known as the so called binomial thinning operation. It then follows that $y_t \sim P(\mu_t)$ so that

$$E(Y_t) = var(Y_t) = \mu_t = e^{x_t'\beta}, \tag{2}$$

for all $t = 1, \ldots, T$. Furthermore, by using (1), one can compute the $E(Y_t Y_{t-\ell})$ which yields the lag $\ell = 1, \ldots, T-1$ correlations as

$$\rho_y(l) = corr(Y_t, Y_{t-l}) = \rho^l \sqrt{\frac{\mu_{t-l}}{\mu_t}}. \tag{3}$$

Note that for the non-stationary case, for $\mu_t - \rho\mu_{t-1}$ to be non-negative ρ must satisfy the range restriction

$$0 < \rho < min\left[1, \frac{\mu_t}{\mu_{t-1}}\right], t = 2, \cdots, T. \tag{4}$$

Mallick and Sutradhar [5] also discussed the estimation of the regression parameter β and the AR(1) correlation parameter ρ.

2.2 Non-stationary MA(1) Models for Count Data

Similar to the Gaussian model, the MA(1) model for the count responses may be expressed as

$$y_t = \rho * d_{t-1} + d_t, \tag{5}$$

where '*' denotes the same binomial thinning operation as in the AR(1) case. Note that under the MA(1) model (5), y_t is a function of discrete errors that occur at the present time point t and at the lag 1 past time point $t-1$, whereas under the AR(1) model (1), y_t is a function of the discrete error at time point t and the lag 1 past count response. Suppose that d_t and d_{t-1} follow the Poisson distributions given as

$$d_t \sim P(\mu_t/(1+\rho)), \text{ and } d_{t-1} \sim P(\mu_{t-1}/(1+\rho)), \tag{6}$$

respectively. Similar to the AR(1) case, y_1 in (5), is assumed to follow the Poisson distribution with parameter μ_1. For $t = 2, \ldots, T$, and using the notation $z_{t-1} = \rho * d_{t-1},\ t = 2, \ldots, T$, one can compute the mean $\nu_t = E(Y_t)$ and the variance $\sigma_{tt} = \text{var}(Y_t)$, as

$$\nu_t = E_{d_{t-1}} E[z_{t-1}] + E[d_t] = [\rho\mu_{t-1} + \mu_t]/(1+\rho), \tag{7}$$

and

$$\begin{aligned} \sigma_{tt} &= \text{var}_{d_{t-1}} E[z_t|d_{t-1}] + E_{d_{t-1}} \text{var}[z_t|d_{t-1}] \\ &\quad + \text{var}[d_{ijt}] \\ &= \text{var}_{d_{t-1}}[\rho d_{t-1}] + E_{d_{t-1}}[\rho(1-\rho)d_{t-1}] \\ &\quad + [\mu_t/(1+\rho)] \\ &= [\rho\mu_{t-1} + \mu_t]/(1+\rho), \end{aligned} \tag{8}$$

respectively. Note that $\sigma_{tt} = \nu_t$, for $t = 2, \ldots, T$, whereas y_1 has the mean and variance as $\nu_1 = \sigma_{11} = \mu_1$ only. Further note that the mean and the variance computed in (7) and (8) also follow from the fact that for $t = 2, \ldots, T$, y_t in (5)) has the Poisson distribution with parameter $\nu_t = [\rho\mu_{t-1} + \mu_t]/(1+\rho)$ where $\mu_t = e^{x_t'\beta}$. Furthermore, it follows from (7) and (8) that the ρ parameter must satisfy the range restriction $\max[-\mu_t/\mu_{t-1}] < \rho < 1$.

Next, under the MA(1) model (5)-(6), it can be shown that for $t = 2, \ldots, T$, and $\ell = 1, \ldots, T-1$, the auto-covariances are given by

$$\text{cov}(Y_t, Y_{t-\ell}) = \begin{cases} \rho\mu_{t-\ell}/(1+\rho) & \text{for } \ell = 1 \\ 0 & \text{for } \ell > 1. \end{cases} \tag{9}$$

3 Estimation of Parameters

This section deals with the estimation of the regression effect β, and the correlation parameter ρ. More specifically, the regression effect β is estimated by using the so-called GQL (see Sutradhar [9] and Zeger [10]) approach, whereas ρ is estimated by the method of moments.

3.1 Estimation of the Regression Effects β

The GQL approach exploits the mean vector and the covariance structure of the data. To be specific, let $y = (y_1, \cdots, y_t, \cdots, y_T)'$ be the T-dimensional vector of all responses and $\nu = (\nu_1, \cdots, \nu_t, \cdots, \nu_T)'$ be the mean vector of y, where ν_t by (7) is given as $\nu_1 = \mu_1$, and for $t = 2, \ldots, T$, $\nu_t = [\mu_t + \rho\mu_{t-1}]/(1+\rho)$. Furthermore, let $\Sigma = (\sigma_{tt'})$ be the $T \times T$ covariance matrix of y, where

$$\sigma_{tt'} = \begin{cases} \sigma_{tt}, & \text{if } t = t' \\ \frac{\rho\mu_t}{1+\rho}, & \text{if } t < t' \end{cases}. \quad (10)$$

with σ_{tt} as given by (8). It then follows that for known ρ, one may write the GQL estimating equation for β as

$$\frac{\partial \nu'}{\partial \beta}\Sigma^{-1}(y-\nu) = 0, \quad (11)$$

(Sutradhar [9]) which may be solved iteratively by Newton-Raphson iterative technique. To be specific, (11) is solved for β iteratively by using

$$\begin{aligned} \hat{\beta}(r+1) &= \hat{\beta}(r) \\ &+ \left[\left(\{X'A + Z'B\}\Sigma^{-1}\{AX+BZ\}\right)^{-1} \right. \\ &\left. \times\{X'A + Z'B\}\Sigma^{-1}(y-\nu)\right]_{[r]}, \quad (12) \end{aligned}$$

where

$$X' = (x_1, \ldots, x_t, \ldots, x_T),\ Z' = (1_p,\ x_1, \ldots, x_{T-1}),$$

$$A = \text{diag}(\mu_1,\ \frac{\mu_2}{1+\rho}, \cdots, \frac{\mu_t}{1+\rho}, \cdots, \frac{\mu_T}{1+\rho}),$$

$$B = \text{diag}(0, \frac{\rho\mu_1}{1+\rho}, \frac{\rho\mu_2}{1+\rho}, \cdots, \frac{\rho\mu_t}{1+\rho}, \cdots, \frac{\rho\mu_{T-1}}{1+\rho}),$$

and $[.]_r$ denotes the fact that the expression within the brackets is evaluated at $\hat{\beta}(r)$. Let $\hat{\beta}_{GQL}$denote the solution obtained from (12). Under mild regularity conditions it may be shown that $\hat{\beta}_{GQL}$ has the asymptotic as $(T \to \infty)$ normal distribution given as

$$\begin{aligned} T^{\frac{1}{2}}(\hat{\beta}_{GQL} - \beta) &\sim N(0,\ T\left[\{X'A + Z'B\}\Sigma^{-1} \right. \\ &\left. \times\{AX+BZ\}\right]^{-1}). \quad (13) \end{aligned}$$

3.2 Estimation of the Correlation Parameter ρ

As far as the ρ parameter is concerned, this will be estimated by using the well known method of moments. For the purpose, one first observes that

$$\begin{aligned} E\left[\frac{(Y_t - \nu_t)}{\sqrt{\nu_t}}\right]^2 &= 1 \\ E\left[\frac{(Y_t - \nu_t)}{\sqrt{\nu_t}}\frac{(Y_{t-1} - \nu_{t-1})}{\sqrt{\nu_{t-1}}}\right] &= \frac{\rho}{1+\rho}\frac{\mu_{t-1}}{\sqrt{\nu_t\nu_{t-1}}} \quad (14) \end{aligned}$$

Consequently, one may obtain a consistent estimator of ρ by solving the moment equation

$$\frac{a(\rho)}{b(\rho)} = \frac{\rho}{1+\rho}c(\rho), \quad (15)$$

where

$$\begin{aligned} a(\rho) &= \frac{1}{T-1}\sum_{t=2}^{T}\frac{(Y_t - \nu_t)}{\sqrt{\nu_t}}\frac{(Y_{t-1} - \nu_{t-1})}{\sqrt{\nu_{t-1}}}, \\ b(\rho) &= \frac{1}{T}\sum_{t=1}^{T}\left[\frac{(Y_t - \nu_t)}{\sqrt{\nu_t}}\right]^2, \end{aligned}$$

and

$$c(\rho) = \frac{1}{T-1}\sum_{t=2}^{T}\frac{\mu_{t-1}}{\sqrt{\nu_t\nu_{t-1}}}. \quad (16)$$

Note that solving (15) for ρ is complicated as ν_t contains ρ for all $t = 1, \ldots, T$. One may however obtain an approximate solution based on an iterative technique by using an initial value of ρ, say ρ_0, in all ν_t, and solving (15) for ρ as

$$\rho_1 = \frac{a(\rho_0)}{b(\rho_0)c(\rho_0) - a(\rho_0)}. \quad (17)$$

Next one may improve the estimate of ρ by using ρ_1 in place of ρ_0 in (17). That is, the new solution of ρ is obtained as

$$\rho_2 = \frac{a(\rho_1)}{b(\rho_1)c(\rho_1) - a(\rho_1)}. \quad (18)$$

This iteration continues until convergence.

4 Conclusion

This paper has introduced two non-stationary time series models, namely the AR(1) and MA(1) models for the analysis of a time series of counts. It is also shown how to obtain consistent and efficient estimates for the parameters of these two models. In a separate paper, these two models have been fitted to the US polio count data and it was seen that the MA(1) model fits this data better than the AR(1) model. This data set was analyzed earlier by Mallick and Sutradhar [5] using the AR(1) model and by Davis et al [2] and Zeger [10] using the so-called random effects approach. Note that the proposed models should be useful for the forecasting of a future count, such as (1) the tourist number in a future month for a city or country, (2) the number of patient from a disease at a future time. This forecasting aspect is however beyond the scope of the present paper.

Acknowledgements: This research was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank the reviewers for helpful comments and suggestions.

References:

[1] M. A. Al-osh and E. A. A. Aly, First order autoregressive time series with negative binomial and geometric marginals, *Commun. Statist.-Theory Meth.* 21, 1992, pp. 2483-2492.

[2] R. A. Davis, W. T. M. Dunsmuir, and Y. Wang, On autocorrelation in a Poisson regression model, *Biometrika* 87, 2000, pp. 491-505.

[3] R. K. Freeland and B. P. M. McCabe, Forecasting discrete valued low count time series, *International Journal of Forecasting* 20, 2004, pp. 427-434.

[4] A. C. Harvey and C. Fernandes, Time series models for count or qualitative observations, *Journal of Business and Economic Statistics* 7, 1989, pp. 407-417.

[5] T. Mallick and B. C. Sutradhar, Analyzing time series of non-stationary negative binomial series, *Technical Report No. 4, Department of Mathematics and Statistics, Memorial University of Newfoundland, Canada*, 2005.

[6] E. McKenzie, Autoregressive moving average processes with negative binomial and geometric marginal distributions, *Adv. Appl. Probab.* 18, 1986, pp. 679-705.

[7] E. McKenzie, Some ARMA models for dependent sequences of Poisson counts, *Adv. in Appl. Probab.* 20, 1988, pp. 822-835.

[8] C. H. Sim and P. Lee, Simulation of negative binomial processes, *J. Statist. Comput. Simul.* 34, 1989, pp. 29-42.

[9] B. C. Sutradhar, An overview on regression models for discrete longitudinal responses, *Statistical Science* 18, 2003, pp. 377-393.

[10] S. L. Zeger, A regression model for time series of counts, *Biometrika* 75, 1988, pp. 621-629.