Testing Serial Dependence in Time Series Models of Counts Against Some INARMA Alternatives

Robert C. Jung[*] *and Andrew R. Tremayne*[†]

This version: January 2001

Abstract

In analysing time series of counts, the need to test for the presence of a dependence structure routinely arises. Suitable tests for this purpose are considered in this paper. Their size and power properties are evaluated under various alternatives taken from the class of INARMA processes. We find that all the tests considered except one are robust against extra binomial variation in the data and that tests based on the sample autocorrelations and the sample partial autocorrelations can help to distinguish between integer-valued first-order and second-order autoregressive as well as first-order moving average processes. Applications pertaining to two real world data sets are provided.

Keywords: Time series of counts; INARMA models; Autocorrelation; Partial Autocorrelation; Score test; Monte Carlo; Size and power properties.

[*]Wirtschaftswissenschaftliche Fakultät, Universität Tübingen, Mohlstr. 36, 72074 Tübingen, Germany. E-mail: Robert.Jung@uni-tuebingen.de

[†]Department of Economics, University of Newcastle, Newcastle upon Tyne, NE1 7RU, England. E-mail: A.R.Tremayne@newcastle.ac.uk

An earlier version of the paper was presented at the 8th World Congress of the Econometric Society August 2000, Seattle, WA. The authors thank Dr. Tim Barmby, University of Newcastle, for supplying the absentee data used in Section 6.

1 Introduction

Time series of (small) counts arise in various fields of statistics. Examples are: the number of customers waiting to be served at a counter recorded at discrete points in time; the daily number of absent workers in a firm; and the monthly cases of rare infectious diseases in a specified area. Typically such series consist of positive (or zero) counts with a sample mean perhaps not higher than 10. This renders any consideration of continuous modelling inappropriate. Several models that take the discretness of the data explicitly into account have been developed in the literature. Following a proposal of Cox (1981) they are divided into two broad categories: observation-driven; and parameter-driven models. While the latter rely on a latent process connecting the observations, the former specify a direct link between current and past observations. In their important monograph Cameron and Trivedi (1998) provide an overview over the recent literature in this area.

This paper focuses on a special class of observation driven models, the so called integer-valued autoregressive-moving average (INARMA) processes introduced by McKenzie (1985) and Al-Osh and Alzaid (1987). They provide an interesting class of discrete valued processes with the ability not only to specify the dependence structure but also to choose among a wide class of (discrete) marginal distributions. Although inherently nonlinear in nature, INARMA models are often specified so as to mimic the linear structure of the well known linear Gaussian ARMA processes. There exists, however, a gap between theoretical models for INARMA processes and their practical application to time series of counts. It is the purpose of this paper to address this gap by assessing methods for determining whether or not the INARMA class should be entertained for a given set of data.

A natural first question in the analysis of time series of counts is whether the data exhibit a significant serial dependence or not. If this is not the case, standard methods suitable for independent data can be applied, otherwise a more sophisticated analysis is called for. To equip the applied researcher with suitable tools to answer the question pertaining to serial dependence, we discuss various standard tests and introduce a new class of tests obtained from the branching process literature. The size and power properties of the various tests are examined in Monte Carlo experiments.

Once significant serial correlation is established in data the next task is to attempt to identify the type of correlation structure and specify a time series model suitable for it. In this context, we suggest a testing strategy that may be useful with regard to particular INARMA processes. The device proves useful in applications with real data reported below.

The paper is organized as follows: In Section 2 three basic INARMA processes are described in some detail. In Section 3 well known and newly designed tests for serial dependence are discussed. Section 4 provides the results of Monte Carlo studies on the size properties of the various testing procedures under the assumption of equidispersed iid Poisson variables as well as overdispersed iid negative binomial variables. The empirical power of the tests against various alternatives is analysed in Section 5. Section 6 provides some applications and Section 7 concludes.

2 Time Series Models for Counts

Three basic observation-driven time series models for counts are presented here. They cover a wide range of possible dependence regimes. First order dependence can either be modelled by the so-called integer-valued first order autoregressive (INAR(1)) model or by the first order moving average (INMA(1)) model. These have been considered extensively elsewhere, see, for example Al-Osh and Alzaid (1987,1988), McKenzie (1988), Brännäs (1994) and Brännäs and Hall (1998). We briefly review their basic properties and structure below. To study higher order dependence the INAR(2) model proves to be a very interesting starting point.

2.1 The INAR(1) process

The INAR(1) process $\{X_t; t = 0, \pm 1, \pm 2, \ldots\}$ is defined by the difference equation

$$X_t = a \circ X_{t-1} + W_t \ , \qquad t = 0, \pm 1, \pm 2, \ldots \ , \tag{1}$$

with the state space of the process being $\mathbb{N}_0$. It is assumed that $a \in [0,1)$ and W_t is an iid discrete random variable sequence with finite first (μ_W) and second (central) moment (σ_W^2). W_t and X_{t-1} are presumed to be stochastically independent for all points in time. The process generated by (1) is stationary.

The process closely resembles the familiar Gaussian AR(1) process but is nonlinear due to the $\circ$-operation replacing the usual scalar multiplication in continuous models. The purpose of this operation, which goes back to the work of Steutel and van Harn (1979), is to ensure the integer discreteness of the process. Following McKenzie (1988) it will be called the *binomial thinning* or simply *thinning* operator and is defined as follows:

$$a \circ X_{t-1} \equiv Y_{1,t-1} + Y_{2,t-1} + \ldots + Y_{X_{t-1},t-1} = \sum_{i=1}^{X_{t-1}} Y_{i,t-1} \ , \tag{2}$$

where the $Y_{i,t-1}$ are assumed to be iid Bernoulli random variables with $P(Y_{i,t-1} = 1) = a$ and $P(Y_{i,t-1} = 0) = 1 - a$. It is important to note that subsequent thinning operations are performed independently of each other with a constant probability a and that thinning is a random operation with an associated probability distribution. Although not as rigorously defined as above, the concept of thinning is nevertheless well known in classical probability theory and has been in use in the Bienaymé-Galton-Watson branching processes (see Feller, 1968, ch. 12) as well as in the theory of stopped-sum distributions (see, for example, Johnson, Kotz and Kemp, 1992, ch. 9). The close relationship between the Bienaymé-Galton-Watson branching processes and the INAR(1) process will be highlighted below and subsequently exploited in the next section.

An illustrative example of the process described by equations (1) and (2) is as follows. Consider X_t to be the number of particles in a well defined space at time t. According to the INAR(1) process this number is made up of: particles, some or all of which were in the space at time $t-1$; and new entrants during the time span $(t-1, t]$. Each particle's probability of staying within the space from one time period to the next is given by a. High (low) values of a generate high (low) correlation among subsequent observations.

Important properties of the INAR(1) process are summarised below. More detailed information is provided in the papers mentioned above as well as in Alzaid and Al-Osh (1988). Grunwald et al. (2000) introduce the class of conditional linear AR(1) models in which the INAR(1) process is nested and discuss some of the properties of this class of models. Due to stationarity, the derivation of the first and second order (unconditional) moments is straightforward: $\mathrm{E}(X_t) = \mu_W/(1-a)$; and $\mathrm{Var}(X_t) = (a\,\mu_W + \sigma_W^2)/(1-a^2)$. The autocorrelation function (ACF) of the process, which is given by $\rho(k) = a^k$ for $k = 1, 2, \ldots$, is identical to the ACF of a linear Gaussian process with the qualification that only positive autocorrelation is allowed here. In contradistinction to Gaussian processes, however, a knowledge of the first and second order moments does not suffice to describe the dependence structure of the process entirely. Note that due to the Markovian property of the INAR(1) model the relevant tool for this purpose is the bivariate distribution function or the bivariate probability generating function (pgf). This is explored in greater detail in a companion paper by Jung, Ronning and Tremayne (2001) (henceforth JRT). The structural equivalence between the INAR(1) process and the well known Bienaymé-Galton-Watson branching process with immigration (BGWI process) can easily been seen when (1) and (2) are compared with a definition of the BGWI process exposited in Athreya and Ney (1972), for example,

$$X_t = \sum_{i=1}^{X_{t-1}} Y_{i,t-1} + W_t\ , \qquad t = 0, \pm 1, \pm 2, \ldots\ , \tag{3}$$

where $Y_{i,t-1}$ and W_t denote stochastically independent lattice iid random variables. The structural equivalence is restricted to the case where $\mathrm{E}(Y_{i,t-1}) < 1$ holds, i.e. to subcritical branching processes. This feature can be exploited for purposes of parameter estimation and inference given the rich body of literature on BGWI processes.

As mentioned in the introduction the INARMA processes are characterized by their dependence structure and by their marginal distributions. So far no assumption about this marginal distribution has been made. A natural first choice in the analysis of counting processes is the Poisson distribution. Following Al-Osh and Alzaid (1987) one assumption is that $W_t \sim \text{Po}(\lambda)$ with $\lambda > 0$. The marginal distribution of the process X_t can than be shown (see Al-Osh and Aly, 1992 and JRT) to be $\text{Po}(\lambda/(1-a))$. The resulting INAR(1) process will henceforth be denoted PoINAR(1).

Figure 1 depicts simulated sample paths of the PoINAR(1) process. For all three panels $\text{E}(X_t) = 2$, while the autocorrelation parameter a and the Poisson parameter λ vary across the panels. In panel $[a]$ the parameter a was set equal to 0.1 resulting in a sample path that looks quite erratic because most of the existing count is thinned and the current value derives mainly from new entrants. In order to obtain $\text{E}(X_t) = 2$ the value of λ is 1.8. In panel $[b]$ we use $a = 0.5$ and $\lambda = 1$ and in panel $[c]$ we set a to 0.9 and λ to 0.1. In applications it seems a often assumes quite a large value, see, for example Hellström (2001). The high autocorrelation is quite evident in the bottom panel as well as the extremely low innovation rate λ. Note that in all three panels the process reverts back to its mean quite regularly as a consequence of the stationarity property.

<Figure 1 about here>

2.2 The INMA(1) process

A different type of dependence structure can be induced using the first order integer-valued moving average (INMA(1)) process $\{X_t, t = 0, \pm 1, \pm 2, \ldots\}$

$$X_t = b \circ W_{t-1} + W_t \ , \qquad t = 0, \pm 1, \pm 2, \ldots \ , \tag{4}$$

with the state space of the process being again $\mathbb{N}_0$. It is assumed that $b \in [0,1]$ and that W_t is a lattice iid random variable with finite mean (μ_W) and variance (σ_W). The thinning operation $b \circ W_{t-1}$ is defined as follows:

$$b \circ W_{t-1} \equiv \sum_{i=1}^{W_{t-1}} Y_{i,t-1} \ , \tag{5}$$

where $Y_{i,t-1}$ is an iid Bernoulli random variate with $P(Y_{i,t-1} = 1) = b$ and $P(Y_{i,t-1} = 0) = 1 - b$.

A physical interpretation of the INMA(1) process can be given using the example above where X_t denotes the number of particles in a well defined space at time t. According to the INMA(1) model this number is made up of particles which entered the space during the time span $(t-1, t]$ and survivors (of the thinning operation) of the entrants during the time span $(t-2, t-1]$. Each element is endowed with a fixed survival probability b. Note that, in contrast to the INAR(1) process, thinning takes place only among the immigrants at time $t-1$, not among all particles present in the space at that time. Further, all particles have a maximum life of two periods and are forced to die automatically after surviving one thinning. This ensures that only adjacent observations of X_t are correlated, whereas observations that are more than one time period apart are not. The resulting process is neither Markovian nor is it a BGWI process.

The first and second order moments of X_t can be derived without assumptions about its marginal distribution. The mean of the process is given by $\mathrm{E}(X_t) = (1+b)\mu_W$ and the variance[1] by $\mathrm{Var}(X_t) = (1+b^2)\sigma_W^2 + b(1-b)\mu_W$. The ACF

$$\rho(k) = \begin{cases} \dfrac{b\,\sigma_W^2}{[b(1-b)\mu_W + (1+b^2)\sigma_W^2]} & \text{for} \quad k=1 \\ 0 & \text{for} \quad k>1 \end{cases} \tag{6}$$

is analogous to the Gaussian MA(1) process. It is straightforward to show that, with b being restricted to $[0,1]$, $\rho(1)$ is allowed to vary in the interval $[0, 0.5]$ only.

A natural first candidate for the marginal distribution of the INMA(1) process is again the Poisson distribution. Assuming $W_t \sim \mathrm{Po}(\lambda)$ it is shown in Al-Osh and Alzaid (1988) or Jung (1999) that $X_t \sim \mathrm{Po}\left(\lambda(1+b)\right)$. The resulting process is a PoINMA(1) process.

Simulated sample paths for different parameter values with this model are depicted in Figure 2. Again $\mathrm{E}(X_t) = 2$ is fixed. The parameter combinations used to generate the graphs are: $b = 0.1$ and $\lambda = 1.8$ for panel $[a]$; $b = 0.5$ and $\lambda = 1.33$ for panel $[b]$; and $b = 1$ and $\lambda = 1$ for panel $[c]$. The resulting autocorrelations are 0.09, 0.33 and 0.5. As before

[1]The property that $b \circ W_{t-1} | W_{t-1}$ follows a binomial distribution with scale parameter b and index parameter W_{t-1} is used in this derivation.

the effect of an increased autocorrelation (less severe thinning) leads to a smoothing of the sample path.

<Figure 2 about here>

2.3 The INAR(2) process

Higher order dependence in the data is not captured by the models discussed so far and so we now explicitly consider one higher order process. The INAR(2) process $\{X_t, t = 0, \pm 1, \pm 2, \ldots\}$, a seemingly natural extension of the INAR(1) process, is defined in the usual manner

$$X_t = a_1 \circ X_{t-1} + a_2 \circ X_{t-2} + W_t \ , \qquad t = 0, \pm 1, \pm 2, \ldots \ . \tag{7}$$

The thinning operations are analogous to (2). To ensure the stationarity of the process, we require that $a_1 + a_2 < 1$. It turns out that, without additional assumptions regarding the thinning operation, no sensible and easily interpretable processes result. Following Alzaid and Al-Osh (1990) we assume that the vector $(a_1 \circ X_{t-1}, a_2 \circ X_{t-1})'$ given X_{t-1} is multinomially distributed with parameters (a_1, a_2, X_{t-1}). JRT discuss the probability generating function and other details pertaining to such a process.

A short description of the way the INAR(2) process can be simulated may provide further insight into the structure of the model. At time t two *binomial thinning* operations are performed: $U_t \equiv a_1 \circ X_t$ and $V_t \equiv a_2 \circ (X_t - U_t)$. While U_t is employed at time point $t+1$ to help generate the new value of the process X_{t+1}, it is not until time point $t+2$ that V_t becomes involved in the generation of the observable process. In general the formula for generating the count for the process at time t is given by $X_t = U_{t-1} + V_{t-2} + W_t$. A direct consequence of the two thinning operations being involved, but appearing at different time points in the process of generating the X_t's, is that a moving average structure is induced. As shown by Alzaid and Al-Osh (1990, p. 320), the autocovariance function of a general INAR(p) process defined as above is similar to that of a Gaussian ARMA($p, p-1$) process. This result is in contrast with that obtained by Du and Li

(1990) who do not employ an additional assumption for the thinning operations in their definition of an INAR(p) process. As a consequence, a sensible physical interpretation of their process is not readily available. Dion, Gauthier and Latour (1995) are able to demonstrate that INAR(p) processes can generally be viewed as multitype branching processes with immigration. Using this link it is, therefore, possible to exploit the classical branching process literature for the analysis of INAR processes of higher order than unity.

Employing the Poisson assumption $W_t \sim \text{Po}(\lambda)$ for the innovation process, an INAR(2) process (PoINAR(2)) with a Poisson marginal distribution results. The first and second order moments of this process can be derived based on the assumption that a_1 and a_2 are independent of one another and of the past history of the process. The mean and variance of the process are equal to $\text{E}(X_t) = \text{Var}(X_t) = \lambda/(1 - a_1 - a_2)$. The ACF satisfies the second-order difference equation

$$\rho(k) = a_1\,\rho(k-1) + a_2\,\rho(k-2) \qquad \text{for} \quad k \geq 2\ , \tag{8}$$

with the starting values $\rho(0) = 1$ and $\rho(1) = a_1$. Note that the first order autocorrelation of this process depends solely on the parameter a_1 while higher order autocorrelations depend on both a_1 and a_2.

As is the case with Gaussian ARMA processes, the a_1/a_2 parameter space can be partitioned into an area where the ACF decays exponentially to zero for all lags $k \geq 2$ and an area where it oscillates before it damps out. This is depicted in Figure 3. For processes, where $a_2 < a_1 - a_1^2$, the ACF decays exponentially to zero. Oscillatory behaviour is found when $a_2 > a_1 - a_1^2$. If a_2 happens to be equal to $a_1 - a_1^2$ the first and second order autocorrelations are equal.

<Figure 3 about here>

Figure 4 depicts simulated sample paths for the PoINAR(2) process for different parameter combinations. In all three panels the sum of a_1 and a_2 is 0.9 and, in order to fix the mean of the generated series at 2, the Poisson parameter λ is set to 0.2. The realisation in panel [a] is generated using $a_1 = 0.8$ and $a_2 = 0.1$. The corresponding sample autocorrelation function (SACF) and sample partial autocorrelation function (SPACF) of the series is shown in top panel of Figure 5. The autocorrelation exhibits no oscillatory behaviour, as is to be expected. In panel [b] of Figure 4, a simulated sample path is depicted that is based on the parameter values $a_1 = a_2 = 0.45$. The corresponding, qualitatively different, SACF is depicted in the middle panel of Figure 5. The two parts of which it is made up are clearly evident. For lags up to $k = 5$ oscillating behaviour is seen, while for $k \geq 6$ the SACF decays exponentially. In the bottom panels of the two figures a situation where $a_1 = 0.1$ and $a_2 = 0.8$ is shown. Again, the corresponding SACF oscillates and subsequently damps out.

<Figure 4 about here>

<Figure 5 about here>

3 Tests of serial dependence

There are several parametric as well as nonparametric tests available in the literature to test for the presence of serial dependence in an ordered sample $(x_1, \ldots, x_T)$ of counts. Many nonparametric tests fail to take the discrete nature of the data into account, or break down due to the presence of multiple ties. All the tests analysed here are selected specifically because of their acknowledgement of this special data situation. Cameron and Trivedi (1998) exposit the use of standard autocorrelation tests routinely employed in the analysis of time series of continuous data. The exact relationships between these tests and the tests proposed here is the subject of separate research; our focus here is on tests specially designed for time series of counts and in the context of the models of Section 2.

The first test considered is the simple runs test. In order to apply the test to time series of counts, the original series has to be dichotomized on the basis of some specific criterion. It is often recommended that the median be used for this task, with observations that are identical to the sample median being discarded. But, given that fact that stationary models for low value counts will often return to the median, many observations would have to be discarded. This leads in some cases to a significant reduction in the power of the test. Following Gibbons and Chakraborti (1992, p. 77) we, therefore, make use of the sample mean, since even with discrete integer data the sample mean will rarely be integer valued. In this way, no data is likely to be lost because of ties with the threshold value used to define the dichotomy.

Under the null hypothesis of no serial dependence the distribution of the number of runs R can be derived using combinatorics (see e.g. Gibbons and Chakraborti, 1992). The resulting test statistic is discrete and therefore conventional nominal significance levels can only usually be attained through a randomized test design. We confirmed by means of Monte Carlo experiments not reported here that, from samples of size 40 upwards, the much more convenient normal approximation

$$Z = \frac{R - 1 - \dfrac{2\,T_1(T-T_1)}{T}}{\left[\dfrac{2T_1(T-T_1)[2T_1(T-T_1)-T]}{T^2(T-1)}\right]^{1/2}}\,, \tag{9}$$

of the runs test can safely be recommended. Wald and Wolfowitz (1940) show that $Z \xrightarrow{d} \mathrm{N}(0,1)$ under the null hypothesis of independence of the observable series of counts.

Since the distribution of R is discrete, the use of a continuity correction is often recommended (see, for example, Gibbons and Chakraborti 1992, p. 77) in the Z-statistic. The resulting test statistic

$$Z_{cc} = \frac{R - 0,5 - \dfrac{2\,T_1(T-T_1)}{T}}{\left[\dfrac{2T_1(T-T_1)[2T_1(T-T_1)-T]}{T^2(T-1)}\right]^{1/2}} \tag{10}$$

is, of course, asymptotically equivalent to the Z statistic.

The test design chosen here is one-sided. This is motivated by the fact that, under the alternative hypotheses of conventional INAR(1), INAR(2) or INMA(1) models explicitly entertained in this paper, no negative autocorrelation is permitted. Hence the only meaningful departure from the null is a smaller number of runs than would be expected under the null. Notice that this state of affairs will apply in particular in the presence of first-order autocorrelation. Due to the one-sided test design the null hypothesis of serial independence is rejected if Z (resp. Z_{cc}) $< z_\alpha$, where z_α is the relevant quantile of the standard normal distribution.

Another approach to testing for the presence of serial dependence in a time series of counts is provided by the score test of Freeland (1998). The test statistic denoted by S is defined as follows:

$$S = \left(\sqrt{T}\bar{x}\right)^{-1} \sum_{t=2}^{T} (x_{t-1} - \bar{x})(x_t - \bar{x}) \ , \tag{11}$$

where $\bar{x} = 1/T \sum_{t=1}^{T} x_t$. Under the null hypothesis of the x_t's being iid Poisson with parameter $\lambda > 0$, Freeland (1998) shows that $S \xrightarrow{d} \mathrm{N}(0,1)$.

A modified version of Freeland's test can be derived utilizing the mean-variance equality property of the Poisson distribution. The modified statistic

$$S^\star = \sqrt{T}\, \frac{\sum_{t=2}^{T} (x_{t-1} - \bar{x})(x_t - \bar{x})}{\sum_{t=1}^{T} (x_t - \bar{x})^2} \tag{12}$$

is asymptotically equivalent to the S-statistic based on the fact that under the null hypothesis of iid Poisson random variables the probability limits of the sample mean and the sample variance are equal. Moreover, it seems reasonable to expect that this version of the test will be better equipped to cope with situations where the Poisson distribution is over-restrictive. This is likely to be of practical relevance in view of the widespread incidence of overdispersion in data.

Again we advocate the use of a one-sided test. A rejection of the null is in order if the measured dependence is higher that would be expected under the null, that is when S or $S^\star > s_\alpha$.

A third class of tests to be considered is derived from goodness-of-fit tests of simple branching processes as proposed by Venkataraman (1982) and by Mills and Seneta (1989). One type of statistic is based on the sample autocorrelations, whereas a second one is based on the sample partial autocorrelations. Both statistics may be invoked in the framework of portmanteau tests originating from the work of Box and Pierce (1970) in the context of models for continuous time series. Parallel to the application of the Box-Pierce test to the original data series in order to test for serial dependence in it (e.g. Pindyck and Rubinfeld, 1991, sec. 16.2 and Cameron and Trivedi, 1998, sec. 7.3.2 in the count data context) we advocate the use of goodness-of-fit tests for simple branching processes to the original series using the test statistics of Venkataraman (1982) and Mills and Seneta (1989) and some of their special features.

Restricting consideration to a statistic based on one term of Mills and Seneta (1989), expression (14) (i.e. the case for their *T = 1*) the relevant implementation of the Venkataraman's (1982) statistic, given in Section 6 of his paper, is as follows:

$$Q_{acf}(1) = \hat{\varrho}_2^2 \frac{\left[\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^2}{\sum_{t=3}^{T}(x_t - \bar{x})^2(x_{t-2} - \bar{x})^2} \tag{13}$$

where $\hat{\varrho}_2 = \sum_{t=3}^{T}(x_t - \bar{x})(x_{t-2} - \bar{x}) / \sum_{t=1}^{T}(x_t - \bar{x})^2$.

Under the null of iid Poisson variables x_t it can be shown that the statistic $Q_{acf}(1) \xrightarrow{d} \chi^2(1)$. An outline of the reasoning is as follows. The limit distribution of the sample autocorrelation $\hat{\varrho}_2$ is needed. Since the summands in the numerator of $\hat{\varrho}_2$ are not independent but form a martingale difference with respect a suitable information set, a martingale central limit theorem is required. This shows that $\sqrt{T}\,\hat{\varrho}_2 \xrightarrow{d} \mathrm{N}(0,1)$ and $T\,\hat{\varrho}_2^2 \xrightarrow{d} \chi^2(1)$. Similar arguments extend this result to all $k = 1, 2, \ldots$. Then the probability limit of the ratio in (13) is needed. For the numerator a weak law of large numbers shows that plim $[T^{-1} \sum_{t=1}^{T}(x_t - \bar{x})^2]^2 = \lambda^2$. The denominator consists of dependent summands, with an appropriate weak law of large numbers showing that plim $T^{-1} \sum_{t=3}^{T}(x_t - \bar{x})^2(x_{t-2} - \bar{x})^2 = \lambda^2$. This establishes that the statistic of (13) is asymptotically equivalent to one based on $T\,\hat{\varrho}_2^2$.

A further test based on the branching process literature is due to Mills and Seneta (1989) and exploits the structure of the sample partial autocorrelations. Based upon a single term the test statistic is given by

$$Q_{pacf}(1) = \hat{\phi}_2^2 \frac{\left[\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^2}{\sum_{t=3}^{T}(x_t - \bar{x})^2(x_{t-2} - \bar{x})^2}, \tag{14}$$

where $\hat{\phi}_2$ is the second order sample partial autocorrelation. It is easy to show that $Q_{pacf}(1) \xrightarrow{d} \chi^2(1)$.

Note that the $Q_{acf}(1)$ test is based on the second order sample autocorrelation whereas the $Q_{pacf}(1)$ test is based on the second order sample partial autocorrelation. In parallel with AR and MA models for continuous time series it may be the case that the two tests implemented in this way can help us distinguish between an INAR(1) and an INMA(1) structure in count data.

The two tests just introduced can easily be expanded to a portmanteau structure in order to assess higher order dependence in the data. The portmanteau version of the $Q_{acf}(1)$ test is defined as follows:

$$Q_{acf}(k) = \sum_{i=1}^{k} \hat{\varrho}_{i+1}^2 \frac{\left[\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^2}{\sum_{t=i+2}^{T}(x_t - \bar{x})^2(x_{t-i-1} - \bar{x})^2}, \tag{15}$$

where $k \geq 2$ is an arbitrary integer. Notice here and again below that these implementations ignore first order lag sample correlations, unlike the suggestions of Cameron and Trivedi (1998). The corresponding portmanteau version of the $Q_{pacf}(1)$ test is given by

$$Q_{pacf}(k) = \sum_{i=1}^{k} \hat{\phi}_{i+1}^2 \frac{\left[\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^2}{\sum_{t=i+2}^{T}(x_t - \bar{x})^2(x_{t-i-1} - \bar{x})^2}, \tag{16}$$

where again $k \geq 2$ is an arbitrary integer. The appropriate limit distribution for both portmanteau versions of the tests is the χ^2-distribution with k degrees of freedom.

4 Empirical size properties of the tests

A Monte Carlo study is used to analyse the size properties of the various test statistics discussed in the last section. The number of Monte Carlo replications is set to 200 000 to provide reasonably narrow confidence intervals for the nominal sizes under investigation. Using the normal approximation[2] the 95% confidence interval for the tests at e.g. a nominal level of 5% is given by $[4.90; 5.10]$ and for a nominal level of 1% by $[0.96; 1.04]$. Under the null, independent Poisson variables are generated for low level data ($\lambda = 1$) and higher level data ($\lambda = 5$). The sample sizes reported are generally $50, 100, 500$ and 1000 (though a response surface analysis employs $T = 40$ and certain intermediate sample sizes). In preliminary work we experimented with smaller sample sizes but various elements of unsatisfactory behaviour were in evidence. Additionally, extra complexities due to randomization devices with runs tests are required. For some further information on this see Jung (1999, sec. 3.2). For the portmanteau tests $Q_{acf}(k)$ and $Q_{pacf)}(k)$ the number of terms, k, on which they are based is 1, 5 and 10.

Table 1 gives the rejection percentages of the different tests for for a nominal level of 5% and 1% (we do, in fact, have comprehensive results for all experiments reported in this Section for all significance points from 1 to 10% in steps of 1% so as to model the whole of the tail of the relevant distribution, but they are not presented to save space). Numbers in boldface indicate that the value is outside the relevant 95% confidence interval. Several conclusions can be drawn. Most tests exhibit size distortions for the smallest sample size used in our simulation study (a fortiori, this applies to the smaller sample sizes experimented with but not reported). Whereas for some tests the size distortions disappear for moderate sample sizes, for some tests they do not, even for T as high as 1000. The simple runs test behaves satisfactorily under the null while the runs test using the continuity correction underrejects even for sample size 1000. This defect leads us to prefer the former and to exclude the test embodying the continuity correction in the analysis of Section 5. Both versions of the score test also behave in this fashion, though empirical rejection frequencies at the nominal 5% level are about 4.8% at the largest sample size.

[2]The relevant formula is: $2 \times 1.96\,[\text{nominal size} \times\ (1 - \text{nominal size})/(200\,000)]^{1/2}$.

The asymptotic theory provides a good guide to the finite sample behaviour of $Q_{acf}(k)$ test, k = 1, 5, but for sample sizes up to 200 there is some tendency to over-reject (with associated size distortion) on the part of $Q_{acf}(10)$. The only test that seems to overreject in all cases in small and moderate samples is the $Q_{pacf}(k)$ test. Only marginal differences in the behaviour of the empirical null distribution of the test statistics can be observed for different values of k. A variation in the level of the series has no more than minimal systematic effects on the observed size behaviour of the test statistics under investigation.

<Table 1 about here>

Overdispersion is a phenomenon often encountered in the analysis of count data. To assess the empirical size of the tests in the presence of overdispersion, Monte Carlo experiments using the negative binomial distribution under the null are also reported. Parameters of this distribution are set so as to reflect a situation with modest overdispersion (variance-mean ratio of 1.5) and another one with higher extra binomial variation (variance-mean ratio = 3). Table 2 gives the rejection frequencies for the different configurations of our Monte Carlo experiment, again for the nominal sizes of 5% and 1%.

<Table 2 about here>

The results are qualitatively similar to those obtained under the Poisson assumption with certain exceptions. An expected one concerns the behaviour of the S test. Since its scaling is based on the mean, which is generally not a consistent estimator of the norming required in the Central Limit Theorem used to establish its limit distribution, it cannot be expected to behave well. Even in the presence of modest overdispersion the test rejects the true null about twice as often as intended by the nominal significance level and in the presence of higher overdispersion the situation is even worse. Clearly this problem is exacerbated by increasing sample size. A second class of tests that is affected by overdispersion in the data are the $Q_{acf}(k)$ and $Q_{pacf}(k)$ tests. For small and medium sample sizes in low level data (mean process level = 1) the portmanteau versions of these tests exhibits some overrejection, and this problem is more evident when overdispersion is more extreme. When the mean of the process increases to 5, all tests behave similarly, except that the performance of S is even worse. In view of the regular occurrence of overdispersion in applied work, this test is excluded from the empirical power analysis which follows.

To provide a convenient summary of all of the results of our Monte Carlo experiments and to allow applied workers to obtain size adjusted critical values for the tests of their choice we performed a response surface analysis. We omitted the Z_{cc} test and the S test from this exercise because we feel there are prima facie reasons why applied workers will not want to use these. All response surface regressions are based upon the empirical rejection frequencies obtained under the null hypothesis of iid Poisson random variables for sample sizes T equal to $40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000$ and for values of the Poisson parameter λ of 1 and 5. This setup gives a sample size of 22 for each regression, so that the results are best interpreted as primarily descriptive. After some experimentation, the following sensible looking functional form was found to work satisfactorily for all the test statistics investigated under the three different sizes $(10\%, 5\%, 1\%)$ chosen.

$$c_\bullet(p) = \beta_\infty + \beta_1 T^{-0.5} + \beta_2 T^{-1} + \beta_3 \lambda T^{-0.5} , \tag{17}$$

where $c_\bullet(p)$ denotes the estimate for the p percent quantile obtained from the Monte Carlo experiments for the various tests as indicated by the bullet. The results are summarised in Table 3, which structure follows closely that used by MacKinnon (1991, p. 275).

<Table 3 about here>

The foregoing implies that we would advocate the use of the simple runs test, the modified score test, or one of the variants of the Q statistics in applications. Table 3 can be used to calculate size adjusted critical values for any test contained therein. In particular, users may wish to size adjust the $S^\star$ test. It will be seen below that this test exhibits excellent power properties when this approach is adopted. We would also generally recommend the use of size adjustment in the context of $Q_{acf}(k)$ and $Q_{pacf}(k)$ tests. In order to use the Table, the value of λ required may be estimated by means of the sample mean of the series.

5 Monte Carlo power properties

The ability of the various tests of serial dependence introduced in Section 3 to distinguish between the alternative data generating processes discussed in Section 2 is now evaluated on the basis of Monte Carlo experiments. All rejection frequencies are calculated on the basis of 10 000 Monte Carlo replications. Due to size distortions found in Section 4, size adjusted as well as asymptotic critical values are used. Sample sizes of 50, 100 and 500 form the main basis of the analysis. Power calculations are also sometimes used for sample size 1000, but most consistents exhibit high power when $T = 500$, therby rendering them of only limited importance. The empirical power curves presented are based on a nominal size of 5%, on occasion corrected as discussed in Section 4. Calculations for other significance levels show no substantive differences in the results.

The PoINAR(1) model serves as the first alternative hypothesis to be analysed. Empirical power curves are calculated for series where the Poisson parameter λ of the innovation process is set to 1 and 5. The autocorrelation parameter a rises from 0, corresponding to the null of serial independence, to 0.90 in steps of 0.05. The results are summarised in Figures 6 and 7. Figure 6 depicts the power curves for all tests discussed in Section 3 surviving the size analysis of the last Section, with the qualification that only a representative of each of the two types of Q tests is depicted. This is because the power of tests is known to be a declining function of the degrees of freedom index for a given noncentrality parameter; see Poskitt and Tremayne (1981, 1986) for related discussion. The sample size used for the graph is 100. In the top panels the results for $\lambda = 1$ are depicted whereas the bottom panels indicate the power differences between processes with $\lambda = 1$ and processes with $\lambda = 5$. The power curves in the left hand panels are based on size adjusted critical values, whilst those to the right use asymptotic critical values.

$<$Figure 6 about here$>$

It is clear that the S^* version of the score test dominates all the other tests over the entire parameter space considered in the experiments. The power of Freeland's original score test S is also good, except when T is small and autocorrelation is high. But in view of the poorer size properties of S under departures from the Poisson model, we would recommend use of the more robust S^*.

The power properties of the basic runs test Z are of interest, for they are quite impressive, being qualitatively similar, though inferior, to those of S^*. However, the runs test is affected by the level of the underlying data, as can be seen from the lower panels in Figure 6 which indicate a power reduction as λ increases. Intuitively, this is due to the fact that the higher the level of the data the less skewed is the marginal distribution of the counts. As a result the probability of shorter runs of observations above or below the sample mean is raised leading to a lower rejection rate for the specified null hypothesis and a concomitant loss of power. Inspection of the lower panels of Figure 6 indicates that the power properties only of the runs test are affected in this way.

The last notable feature of this Figure is that the empirical power of the tests based on sample partial autocorrelations differ markedly from those based on sample autocorrelations. The $Q_{pacf}(1)$ test (as well as its portmanteau versions $Q_{pacf}(5)$ and $Q_{pacf}(10)$) possess virtually no power at all and indeed exhibit a tendency to be biased for high values of the parameter a. Of course, in view the special choice of these statistics omitting a contribution from the first order sample partial autocorrelation used in this paper this is unsurprising. The $Q_{acf}(1)$ test performs a little better than its portmanteau versions $Q_{acf}(5)$ and $Q_{acf}(10)$, but the power properties of all versions of this test are clearly inferior to the score and runs tests. The tendency towards consistency of the $S^\star$ test and the Q_{acf} test is illustrated in Figure 7, with the superiority of the former at all sample sizes being clearly in evidence.

<Figure 7 about here>

The second alternative analysed is the PoINMA(1) model. Departures from the null hypothesis are now characterized by non-zero values of b. Again we conduct Monte Carlo experiments based on increasing values of b in the range 0 to 0.9 in (4). The results are summarised in Figures 8 and 9.

<Figure 8 about here>

Several points emerge from an inspection of Figure 8. While it seems to be harder to identify a first-order moving average structure in the data (as might be expected by analogy with continuous models), the ability of the different tests to detect it is broadly comparable to the first-order autoregressive case discussed above in many cases. Again the score test dominates the other tests employed in this study, with the simple runs test continuing to mimic its behaviour, albeit in an inferior manner. The roles of the $Q_{pacf}(k)$ and $Q_{acf}(k)$ tests are now reversed, as is to be anticipated. The latter is not now a consistent test and exhibits minimal rejection frequencies; the portmanteau versions of the statistics are not depicted for reasons of clarity, but the power of Q_{pacf} is a declining function of the degrees of freedom index. The strikingly different behaviour of these two types of test statistics will subsequently be exploited to produce a device for determinining if one or other, or neither, of the two models discussed in this Section thus far are likely to be adequate for a given data set. Figure 9 serves as a companion to Figure 7 and provides pointers to the consistency of S^* and Q_{pacf} (and, by implication, Z).

<Figure 9 about here>

The power curves in the left hand panels in Figure 8 are based on the adjusted critical values obtained in Section 4. The influence of size distortions introduced by use of asymptotic critical values is restricted to small departures from the null and no useful purpose would be served by additionally depicting that case. As under the PoINAR(1) alternative, it can be inferred from the right hand panel of Figure 8 that the only test affected by the change in the level of the data is the runs test. The explanation for this phenomenon has already been provided above.

The final INARMA process to be analysed here is the PoINAR(2) process (7). The design of the Monte Carlo experiments is analogous to that used above. Departures from the null hypothesis of iid Poisson variables are controlled by a vector of parameter values for a_1 and a_2 with $0 \leq a_1 + a_2 \leq 0.9$. Based on the very distinct behaviour of the ACF as described in Section 2.3 above, two sets of experiments are conducted. For the first set, the combination of a_1 and a_2 is chosen in such a way that a_2 is kept below the parabolic boundary depicted in Figure 3. In the second set of experiments, series with oscillating autocorrelations are generated. The results are graphed in Figure 10.

<Figure 10 about here>

As long as $a_2 < a_1 - a_1^2$ the autocorrelation properties of a PoINAR(2) process are quite similar to those of a PoINAR(1) process. The empirical power curves shown in the top panels of Figure 10 reflect this (the right hand panels are presented to show the limited influence of the sometimes inadvisable use of asymptotic critical values). But in contrast to the first order process the second order autoregressive process always exhibits a non-zero second order partial autocorrelation (see the top right panel of Figure 5). Consequently under the PoINAR(2) alternative the $Q_{pacf}(k)$ test has power. All the tests considered are now consistent under this alternative.

Once $a_2 > a_1 - a_1^2$ the situation changes noticeably, as is evident from the bottom left hand panel of Figure 10. The score test and the runs test are not very powerful against this parameter constellation of the PoINAR(2) alternative. The other tests now exhibit superior power performance over the relevant area of the parameter space for which the sum of the thinning parameters is large. The $Q_{acf}(1)$ test seems to be the test with the highest power under these circumstances. The second most powerful test is the $Q_{pacf}(1)$ test while the portmanteau versions behave as expected. The explanation for these results can be inferred by referring to Figure 5, which illustrates that the second order sample autocorrelation and partial autocorrelation will generally dominate their first order counterparts; the presence of the only first order autocorrelation in the S (and its major influence in the Z) statistics, together with its absence (but presence of second order lag statistics) in the Q-type tests, is clearly reflected in the empirical rejection frequencies.

Variations in the design parameters of the Monte Carlo do not produce substantively different results with respect to the power properties of the tests, though an increase in the level of births in the data has a tendency to decrease the power of Z slightly and, in fact, to increase that of certain portmanteau statistics marginally. Our findings based on the power analyses can be summarised as follows. The (modified) score test exhibits good empirical power properties when it comes to detecting a dependence structure that is restricted to first order Poisson integer time series models. As soon as the dependence structure is more complicated the other tests advocated in this paper come into their own, at least under certain parameter configurations.

Finally, the specially designed $Q_{acf}(k)$ and $Q_{pacf}(k)$ tests in conjunction with their quite distinctive power behaviour can potentially be used to permit a classification of the correlation structure among certain INARMA processes. The device is simple and requires consideration of a multiplicity of statistics (we suggest three, viz. S^*, $Q_{acf}(1)$ and $Q_{pacf}(1)$). If no statistic results in a significant value, the series under consideration may have no dependence structure (though this tentative conclusion must be tempered with the caveat that some higher order dependence not readily detected by any of the three is present). If S^* rejects the null hypothesis of serial independence, but the others do not (perhaps because of their inferior power properties) there appears a dependence structure but further investigation may need to be made to confirm its form. If $Q_{acf}(1)$ rejects (as frequently would S^*), but $Q_{pacf}(1)$ does not, an INAR(1) process may be tentatively determined, with an INMA(1) being indicated if the behaviour of the two Q tests is interchanged. Finally, if the Q tests suggest rejection of the null hypothesis, but S^* does not, a higher order model might be entertained.

6 Applications

The methods and findings of this paper are applied to two data sets. The first data set is well known in the branching process literature and has been analysed *inter alia* in Mills and Seneta (1989). The data consist of 505 counts of pedestrians on a city block observed every five seconds and have been compiled by and used in Fürth (1918). The sample mean of the series is 1.59 and the sample variance is 1.51, thus raising no question of the presence of overdispersion. The results of the various tests for serial correlation are summarised in the second column of Table 3. All tests are significant at the conventional significance levels. Since both the $Q_{acf}(k)$ tests and the $Q_{pacf}(k)$ tests provide evidence for the presence of serial correlation in the data, it can be inferred, adopting the strategy outlined at the end of Section 5, that neither the simple PoINAR(1) model nor the simple PoINMA(1) model adequately characterise the Fürth data. This result is in accordance with Mills and Seneta (1989), who found a poor fit for their BGWI model which is equivalent to the PoINAR(1) model. Since no significant overdispersion can be found in the data, fitting a PoINAR(2) model to this data set might be entertained.

The second data set consists of a daily count of the number of absentees in a specific firm. The sample size is 616, the sample mean is 5.04 and the sample variance is 5.49, providing slight evidence of overdispersion. The Q-test of Davis, Dunsmuir and Wang (1999, p. 80) is applied to investigate this feature more rigorously. These authors indicate that this Q-test may be more suitable than other tests with lagged variables. They provide Monte Carlo evidence of unsatisfactory size properties with smaller sample sizes than we have with iid Poisson random variables.

Since the limiting standard normal distribution of the statistic will not be appropriate if dependence is present in the data, we report the results of a pilot Monte Carlo experiment to assess the evidence of overdispersion in our data. Using a sample size of 600 and generated series at, or very close to, 5, we conducted 100000 replications of the Q-test under Poisson independence. We find excellent agreement with the asymptotic distribution, which may not be surprising. When dependent random variables are used in the test, there is a noticeable deterioration in the usefulness of the asymptotic standard normal approximation as dependence increases. We experimented with PoINAR(1) models with $a = 0.5$ and 0.9 and an PoINMA(1) model with $b = 0.9$. Although not reported in any detail here, the empirical rejection rate at a nominal 5% level when $a = 0.5$ and $a = 0.9$ are, respectively 9.34% and 21.58%. Computing the Q-test for the absentee data yields a statistic $Q = 1.54$. As the sample autocorrelation function indicates substantial dependence structure in the data, we feel it is too draconian to apply the usual asymptotic distribution to assess the evidence of overdispersion, though if one did, a Prob-value of 0.05 would eventuate. We evaluated the Prob-value for the test under certain plausible dependence structures also and found it to be over 0.2 with high dependence and 0.1 for moderate dependence (by which we mean $a = 0.5$ and $b = 0.9$). We, therefore, judge that the evidence against the null hypothesis of no overdispersion is by no means strong. In the spirit of Cameron and Trivedi (1998) p. 79 we, at any rate, feel the overdispersion is at most modest and conclude that the Poisson assumption for the absentee data is not unreasonable.

The results of the various tests of serial dependence are summarised in column 3 of Table 4. All tests except the $Q_{pacf}(1)$ test have Prob-values of less than 0.05. In accordance with the strategy of the last section we tentatively conclude that a first-order autocorrelation structure may be present in the data (though there may also be higher order dependencies outwith the ambit of this paper). The simple PoINAR(1) model therefore could provide a suitable basis for a further analysis of this data set.

7 Summary and conclusions

This paper presented different kind of tests for serial correlation applicable to time series of counts. In a Monte Carlo study both the size and power properties of the tests are evaluated thoroughly against a range of alternative data generating process.

The set of statistics of serial independence originally advanced is reduced somewhat by undesirable behaviour of finite sample approximations to their standard asymptotic null distributions arising from over-restrictiveness of the Poisson assumption and other features. The power properties of the remaining tests indicate that dependence structure in the data can often expect to be detected. A simple strategy to help determine an appropriate model for data is proposed and shown to be of potential value in applications. The strategy exploits the special structure of test procedures originally advanced in the branching process literature. A model of higher order than first order within the INARMA class is also considered and appears to be a useful benchmark for moving to higher order dependence structures.

References

AL-OSH, M.A. and ALY, E.-E.A.A. (1992) First Order Autoregressive Time Series with Negative Binomial and Geometric Marginals. *Commun. Statist.-Theor. Meth.* 21, 2483-2492.

Al-OSH, M.A. and ALZAID, A.A. (1987) First-Order Integer-Valued Autoregressive (INAR(1)) Process. *Journal of Time Series Analysis* 8, 261-275.

Al-OSH, M.A. and ALZAID, A.A. (1988) Integer-Valued Moving Average (INMA) Process. *Statistical Papers* 29, 281-300.

ALZAID, A.A. and AL-OSH, M.A. (1988) First-Order Integer-Valued Autoregressive (INAR(1)) Process: Distributional and Regression Properties. *Statistica Neerlandica* 42, 53-61.

ALZAID, A.A. and AL-OSH, M.A. (1990) An Integer-Valued p-th Order Autoregressive Structure (INAR(p)) Process. *Journal of Applied Probability* 27, 314-324.

ATHREYA, K.B. and NEY, P.E. (1972) *Branching Processes.* Springer Verlag, Heidelberg.

BOX, G.E.P. and PIERCE, D.A. (1970) Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association* 65, 1509-1526.

BRÄNNÄS, K. (1994) Estimating and Testing in Integer-Valued AR(1) Models. *Working Paper No. 335, Department of Economics, University of Umeå, Sweden*

BRÄNNÄS, K. and HALL, A. (1998) Estimation in Integer-Valued Moving Average Models. *Working Paper No. 477, Department of Economics, University of Umeå, Sweden.*

CAMERON, A.C. and TRIVEDI, P.K. (1998) *Regression Analysis of Count Data.* Cambridge University Press, Cambridge.

COX, D.R. (1981) Statistical Analysis of Time Series: Some Recent Developments. *Scandinavian Journal of Statistics* 8, 93-115.

DAVIS, R.A. DUNSMUIR, W.T.M. and WANG, Y. (1999) Modeling Time Series of Count Data. In: GOSH, S. (Ed.) *Asymptotics, Nonparametrics and Time Series: A Tribute to Madan Lal Puri.* Marcel Dekker, New York

DION, J.-P. GAUTHIER, G. and LATOUR, A. (1995) Branching Processes with Immigration and Integer-Valued Time Series. *Serdica Mathematical Journal* 21, 123-136.

DU, J.-G. and LI, Y. (1991) The Integer-Valued Autoregressive (INAR(p)) Model. *Journal of Time Series Analysis* 12, 129-142.

FELLER, W. (1968) *An Introduction to Probability Theory and Its Application, Volume I.* Third edition. John Wiley and Sons, New York.

FÜRTH, R. (1918) Statistik und Wahrscheinlichkeitsnachwirkung. *Physikalische Zeitschrift* 19, 421-426.

FREELAND, K. (1998) *Statistical Analysis of Discrete Time Series with Application to the Analysis of Worker's Compensation Claims Data.* Ph.D. thesis, University of British Columbia.

GIBBONS, J.D. and CHAKRABORTI, S. (1992) *Nonparametric Statistical Inference* (3rd edn). Marcel Dekker, New York.

GRUNWALD, G.K. HYNDMAN, R. TEDESCO, L. and TWEEDIE, R.L. (2000) Non-Gaussian Conditional Linear AR(1) Models. *Australian and New Zealand Journal of Statistics* 42, 479-495.

HELLSTRÖM, J. (2001) Unit Root Testing in Integer-Valued AR(1) Models. *Economics Letters* 70, 9-14.

JOHNSON, N.L. KOTZ, S. and KEMP, A.W. (1992) *Univariate Discrete Distributions* (2nd edn). Wiley, New York.

JUNG, R. (1999) *Zeitreihenanalyse für Zähldaten. Eine Untersuchung ganzzahliger Autoregressiver-Moving-Average-Prozesse.* Josef Eul Verlag, Lohmar.

JUNG, R., RONNING, G. and TREMAYNE, A.R. (2001) Estimation of Low Order INAR Models For Time Series of Counts. Mimeo.

MACKINNON, J.G. (1991) Critical Values for Cointegration Tests. In: ENGLE, R.F. and GRANGER, C.W.J. (eds.) *Long-Run Economic Relationships.* Oxford University Press, Oxford. Chapter 13.

MCKENZIE, E. (1985) Some Simple Models for Discrete Variate Time Series. *Water Ressources Bulletin* 21, 645-650.

MCKENZIE, E. (1988) Some ARMA Models for Dependent Sequences of Poisson Counts. *Advances in Applied Probability* 20, 822-835.

MILLS, T.M. and SENETA, E. (1989) Goodness-of-fit for a Branching Process with Immigration Using Sample Partial Autocorrelations. *Stochastic Processes and their Applications* 33, 151-161.

PINDYCK, R.S. and RUBINFELD, D.L. (1998) *Econometric Models and Economic Forecasts* (4th edn). McGraw-Hill, New York.

POSKITT, D.S. and TREMAYNE, A.R. (1981) An Approach to Testing Linear Time Series Models. *The Annals of Statistics* 9, 974-986.

POSKITT, D.S. and TREMAYNE, A.R. (1986) Some Aspects of the Performance of Diagnostic Checks in Bivariate Time Series Models. *Journal of Time Series Analysis* 7, 217-233.

STEUTEL, F.W. and VAN HARN, K. (1979) Discrete Analogues of Self-Decomposability and Stability. *The Annals of Probability* 7, 893-899.

VENKATARAMAN, K.N. (1982) A Time Series Approach to the Study of the Simple Subcritical Galton–Watson Process with Immigration. *Advances in Applied Probability* 14, 1-20.

WALD, A. and WOLFOWITZ, J. (1940) On a Test Whether Two Samples are From the Same Population. *Annals of Mathematical Statistics* 11, 147-162.

size	λ	test	$T=50$	$T=100$	$T=500$	$T=1000$
5%	1	Z	**5.30**	**5.11**	4.99	4.99
		Z_{cc}	**3.72**	**4.08**	**4.50**	**4.63**
		S	**3.41**	**4.11**	**4.57**	**4.74**
		$S^\star$	**3.44**	**4.08**	**4.64**	**4.73**
		$Q_{acf}(1)$	4.99	**5.11**	5.08	5.02
		$Q_{acf}(5)$	5.01	**5.19**	**5.22**	5.08
		$Q_{acf}(10)$	**6.05**	**5.86**	**5.36**	**5.26**
		$Q_{pacf}(1)$	**6.13**	**5.71**	**5.21**	5.07
		$Q_{pacf}(5)$	**6.11**	**5.75**	**5.32**	**5.15**
		$Q_{pacf}(10)$	**6.53**	**6.05**	**5.32**	**5.22**
	5	Z	**5.25**	**4.81**	5.00	4.96
		Z_{cc}	**3.46**	**4.13**	**4.55**	**4.64**
		S	**3.37**	**3.84**	**4.47**	**4.73**
		$S^\star$	**3.38**	**3.82**	**4.47**	**4.74**
		$Q_{acf}(1)$	**4.80**	4.96	5.02	5.09
		$Q_{acf}(5)$	4.98	4.97	5.04	4.92
		$Q_{acf}(10)$	**5.95**	**5.66**	**5.14**	5.07
		$Q_{pacf}(1)$	**5.69**	**5.39**	**5.11**	**5.11**
		$Q_{pacf}(5)$	**5.60**	**5.24**	5.05	4.95
		$Q_{pacf}(10)$	**5.86**	**5.33**	5.06	5.02
1%	1	Z	**1.10**	**1.12**	**1.06**	1.04
		Z_{cc}	**0.74**	**0.84**	**0.94**	0.96
		S	**0.84**	**0.93**	0.97	1.02
		$S^\star$	**0.59**	**0.79**	**0.92**	1.00
		$Q_{acf}(1)$	**0.73**	**0.91**	1.00	1.02
		$Q_{acf}(5)$	1.04	1.02	1.02	1.04
		$Q_{acf}(10)$	**1.72**	**1.49**	**1.15**	**1.10**
		$Q_{pacf}(1)$	**1.07**	**1.13**	**1.05**	**1.05**
		$Q_{pacf}(5)$	**1.15**	**1.12**	1.02	1.03
		$Q_{pacf}(10)$	**1.34**	**1.23**	**1.06**	**1.06**
	5	Z	**0.88**	1.03	1.02	1.01
		Z_{cc}	**0.73**	**0.71**	**0.91**	**0.92**
		S	**0.78**	**0.81**	**0.89**	**0.91**
		$S^\star$	**0.52**	**0.67**	**0.87**	**0.90**
		$Q_{acf}(1)$	**0.71**	**0.85**	0.98	1.00
		$Q_{acf}(5)$	**1.08**	**1.05**	1.04	1.03
		$Q_{acf}(10)$	**1.71**	**1.47**	**1.10**	**1.05**
		$Q_{pacf}(1)$	0.99	1.00	1.00	1.02
		$Q_{pacf}(5)$	**1.05**	**0.95**	0.98	1.01
		$Q_{pacf}(10)$	**1.15**	1.02	1.00	1.01

Table 1: *Rejection percentages of the tests under the iid Poisson assumption at nominal 5% and 1% significance levels.*

			modest overdispersion				high overdispersion			
size	E(X)	test	$T=50$	$T=100$	$T=500$	$T=1000$	$T=50$	$T=100$	$T=500$	$T=1000$
5%	1	Z	**5.48**	4.97	**5.17**	5.10	**5.49**	4.98	**5.17**	5.01
		Z_{cc}	**3.68**	**4.15**	**4.64**	**4.77**	**3.67**	**4.10**	**4.57**	**4.81**
		S	**9.28**	**10.78**	**12.62**	**12.91**	**18.67**	**21.70**	**25.79**	**26.88**
		S^*	**3.71**	**4.27**	**4.87**	**4.80**	**3.96**	**4.62**	**5.22**	**5.18**
		$Q_{acf}(1)$	**5.15**	**5.51**	**5.28**	5.07	**5.57**	**7.05**	**6.34**	**5.97**
		$Q_{acf}(5)$	5.10	**5.76**	**5.59**	**5.42**	**5.70**	**8.07**	**8.37**	**7.51**
		$Q_{acf}(10)$	**6.22**	**6.53**	**5.91**	**5.64**	**7.05**	**9.95**	**9.93**	**8.53**
		$Q_{pacf}(1)$	**6.78**	**6.43**	**5.49**	**5.18**	**8.63**	**8.88**	**6.80**	**6.21**
		$Q_{pacf}(5)$	**7.31**	**6.93**	**5.87**	**5.55**	**11.08**	**11.17**	**9.08**	**7.88**
		$Q_{pacf}(10)$	**8.71**	**7.77**	**6.12**	**5.70**	**15.86**	**14.45**	**10.71**	**8.93**
	5	Z	**5.24**	**4.86**	4.98	5.06	**5.11**	**4.77**	4.99	4.96
		Z_{cc}	**3.42**	**4.25**	**4.53**	**4.77**	**3.53**	**3.97**	**4.53**	**4.62**
		S	**9.47**	**10.84**	**12.62**	**13.04**	**22.09**	**24.17**	**27.22**	**27.83**
		S^*	**3.36**	**3.93**	**4.57**	**4.76**	**3.53**	**4.06**	**4.65**	**4.72**
		$Q_{acf}(1)$	**4.83**	4.95	4.97	5.01	4.97	**5.16**	**5.12**	5.03
		$Q_{acf}(5)$	**4.87**	5.01	5.04	4.95	**4.85**	**5.21**	**5.21**	**5.20**
		$Q_{acf}(10)$	**5.91**	**5.61**	**5.11**	5.01	**5.83**	**5.82**	**5.45**	**5.29**
		$Q_{pacf}(1)$	**5.87**	**5.43**	5.07	5.06	**6.24**	**5.86**	**5.29**	**5.11**
		$Q_{pacf}(5)$	**5.75**	**5.38**	5.09	5.01	**6.51**	**6.04**	**5.38**	**5.29**
		$Q_{pacf}(10)$	**6.16**	**5.48**	5.06	5.01	**7.51**	**6.50**	**5.56**	**5.26**
1%	1	Z	**1.08**	**1.09**	**1.07**	**1.07**	**1.18**	**1.14**	**1.06**	**1.08**
		Z_{cc}	**0.80**	**0.72**	**0.93**	0.98	**0.76**	**0.83**	**0.94**	0.96
		S	**4.19**	**4.91**	**5.78**	**5.79**	**13.23**	**15.83**	**19.36**	**20.20**
		S^*	**0.74**	**0.94**	**1.08**	**1.05**	1.03	**1.30**	**1.40**	**1.35**
		$Q_{acf}(1)$	**0.68**	0.98	**1.08**	**1.05**	**0.51**	**1.35**	**1.81**	**1.61**
		$Q_{acf}(5)$	0.98	**1.16**	**1.24**	**1.15**	**0.86**	**1.62**	**2.44**	**2.14**
		$Q_{acf}(10)$	**1.61**	**1.61**	**1.36**	**1.28**	**1.59**	**2.39**	**2.98**	**2.46**
		$Q_{pacf}(1)$	**1.18**	**1.33**	**1.18**	**1.09**	**1.47**	**2.10**	**2.03**	**1.73**
		$Q_{pacf}(5)$	**1.55**	**1.51**	**1.30**	**1.21**	**2.93**	**3.03**	**2.76**	**2.32**
		$Q_{pacf}(10)$	**2.18**	**1.84**	**1.41**	**1.24**	**5.42**	**4.62**	**3.33**	**2.61**
	5	Z	**0.87**	**1.05**	0.98	1.03	**0.93**	0.99	0.97	1.00
		Z_{cc}	**0.70**	**0.71**	**0.90**	**0.94**	**0.70**	**0.72**	**0.87**	**0.91**
		S	**4.03**	**4.71**	**5.53**	**5.75**	**15.72**	**17.61**	**20.31**	**20.83**
		S^*	**0.54**	**0.72**	**0.91**	0.97	**0.61**	**0.83**	0.97	1.00
		$Q_{acf}(1)$	**0.64**	**0.83**	0.99	0.99	**0.64**	**0.84**	1.02	1.03
		$Q_{acf}(5)$	0.98	1.01	1.01	0.97	0.97	1.01	**1.09**	**1.09**
		$Q_{acf}(10)$	**1.60**	**1.40**	**1.12**	**1.06**	**1.57**	**1.43**	**1.17**	**1.13**
		$Q_{pacf}(1)$	**0.94**	0.97	1.02	1.01	1.01	**1.10**	**1.10**	**1.06**
		$Q_{pacf}(5)$	**1.08**	1.01	1.02	0.98	**1.29**	**1.18**	**1.08**	**1.10**
		$Q_{pacf}(10)$	**1.28**	1.03	0.99	1.01	**1.73**	**1.38**	**1.12**	**1.07**

Table 2: *Rejection percentages under independence of the tests under overdispersion at nominal* 5% *and* 1% *significance levels.*

test	size(%)	β_∞	(S.E.)	β_1	β_2	β_3
Z	10	−1.260	(0.013)	−0.590	2.598	0.013
	5	−1.607	(0.020)	−1.199	5.875	0.027
	1	−2.335	(0.016)	−0.427	0.227	0.114
$S^\star$	10	1.284	(0.005)	−0.951	−1.793	−0.006
	5	1.647	(0.007)	−0.773	−2.814	−0.029
	1	2.327	(0.013)	−0.345	−5.802	−0.079
$Q_{acf}(1)$	10	2.685	(0.013)	1.198	−2.069	−0.090
	5	3.841	(0.017)	1.103	−6.725	−0.121
	1	6.648	(0.042)	1.425	−32.110	−0.141
$Q_{acf}(5)$	10	9.201	(0.037)	2.667	−12.406	−0.178
	5	11.048	(0.056)	2.585	−13.783	−0.184
	1	15.097	(0.081)	1.684	−7.592	−0.015
$Q_{acf}(10)$	10	15.864	(0.054)	6.501	−15.634	−0.285
	5	18.140	(0.060)	9.240	−16.085	−0.263
	1	22.809	(0.123)	18.615	−2.639	−0.098
$Q_{pacf}(1)$	10	2.681	(0.013)	1.606	7.269	−0.166
	5	3.819	(0.019)	2.195	4.101	−0.243
	1	6.593	(0.057)	4.472	−21.542	−0.361
$Q_{pacf}(5)$	10	9.212	(0.029)	3.258	9.923	−0.425
	5	11.048	(0.043)	3.517	7.360	−0.536
	1	15.236	(0.091)	−1.828	32.090	−0.707
$Q_{pacf}(10)$	10	15.937	(0.038)	5.271	16.802	−0.722
	5	18.317	(0.064)	4.461	21.342	−0.849
	1	23.526	(0.141)	−2.832	62.486	−1.067

Table 3: *Descriptive results from response surface analysis*

test	Fürth data	Absentee data
Z	$-10.71^{\star\star\star}$	$-12.40^{\star\star\star}$
$S^\star$	$14.93^{\star\star\star}$	$16.34^{\star\star\star}$
$Q_{acf}(1)$	$40.02^{\star\star\star}$	$108.86^{\star\star\star}$
$Q_{acf}(5)$	$87.12^{\star\star\star}$	$379.31^{\star\star\star}$
$Q_{acf}(10)$	$116.97^{\star\star\star}$	$579.00^{\star\star\star}$
$Q_{pacf}(1)$	$17.51^{\star\star\star}$	$3.65^{\star}$
$Q_{pacf}(5)$	$27.77^{\star\star\star}$	$17.95^{\star\star\star}$
$Q_{pacf}(10)$	$37.27^{\star\star\star}$	$22.52^{\star\star}$

$\star\star\star$ denotes significance at the 1% level
$\star\star$ denotes significance at the 5% level
$\star$ denotes significance at the 10% level

Table 4: *Results of the various tests applied to two data sets*

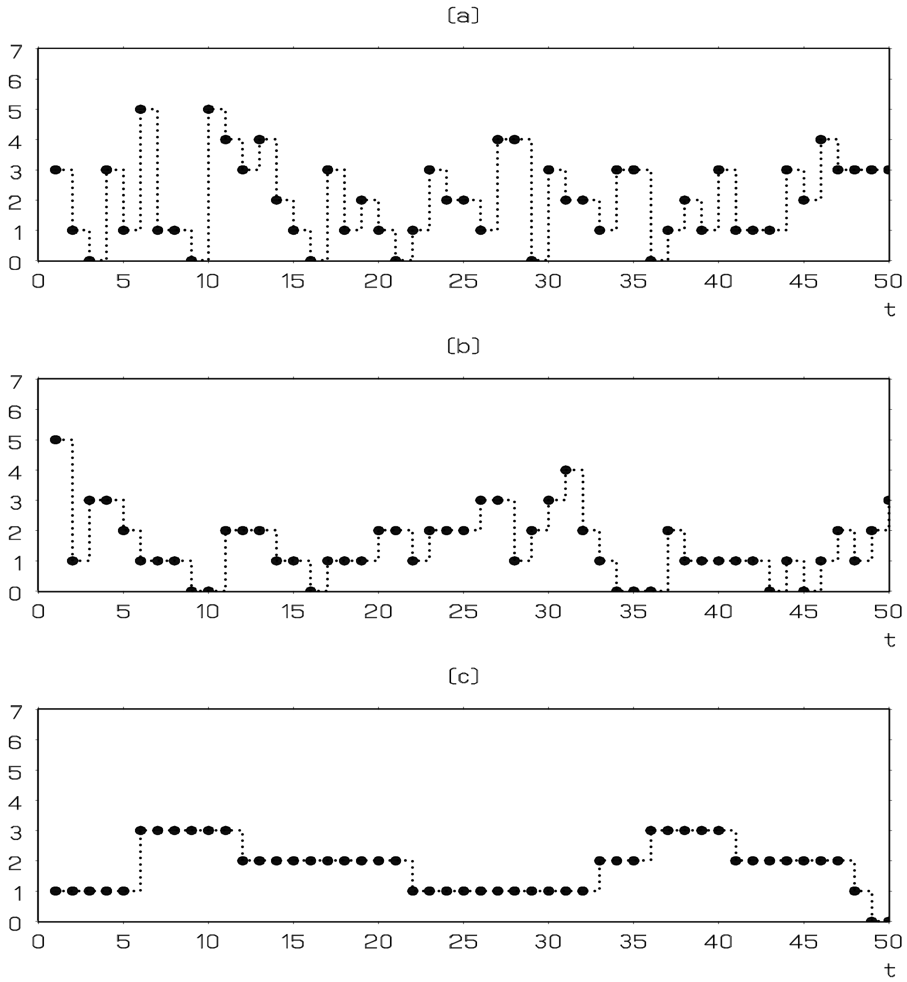

Figure 1: *Simulated sample paths for a PoINAR(1) process*

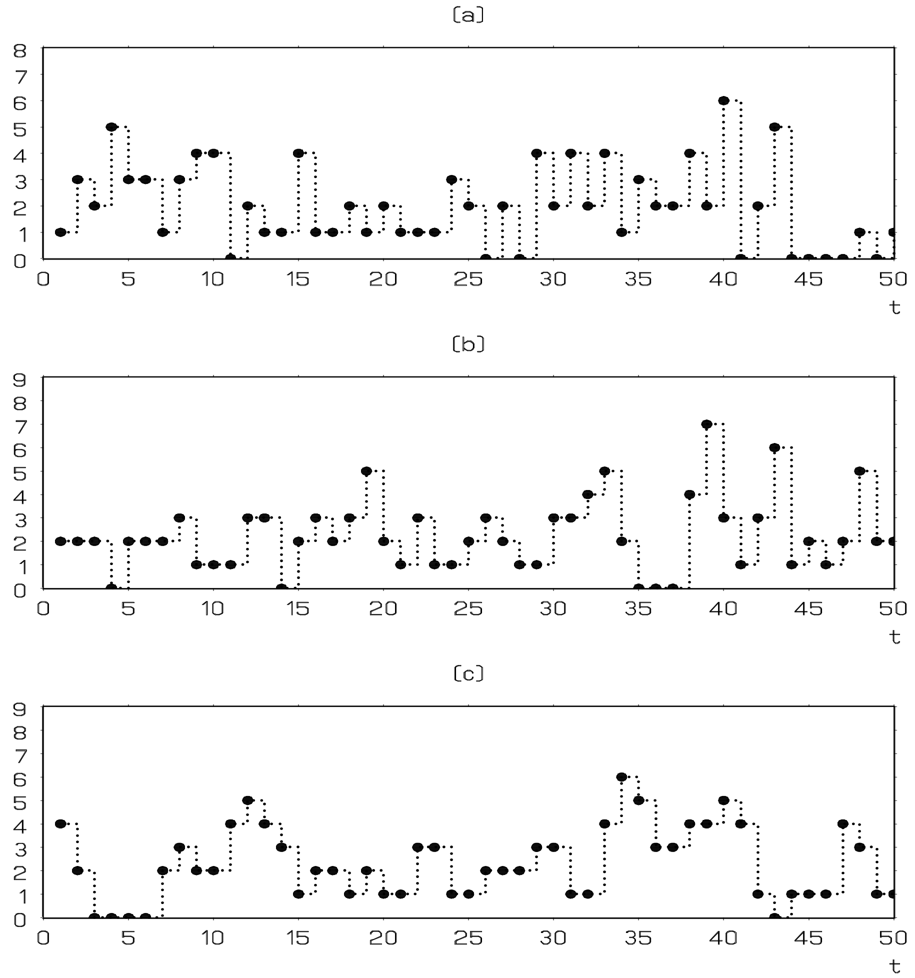

Figure 2: *Simulated sample paths for a PoINMA(1) process*

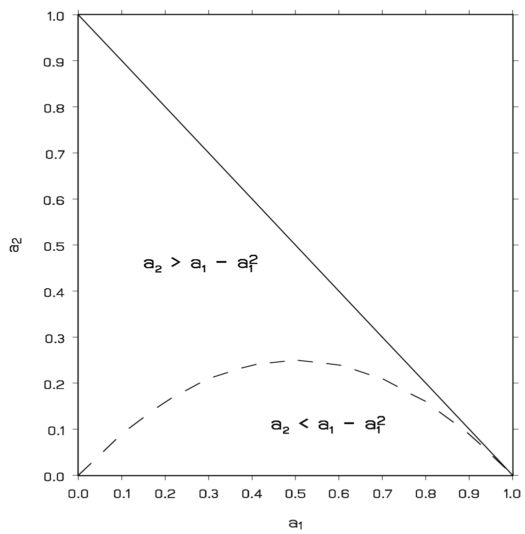

Figure 3: *Partition of the a_1/a_2-parameter space in the PoINAR(2) process*

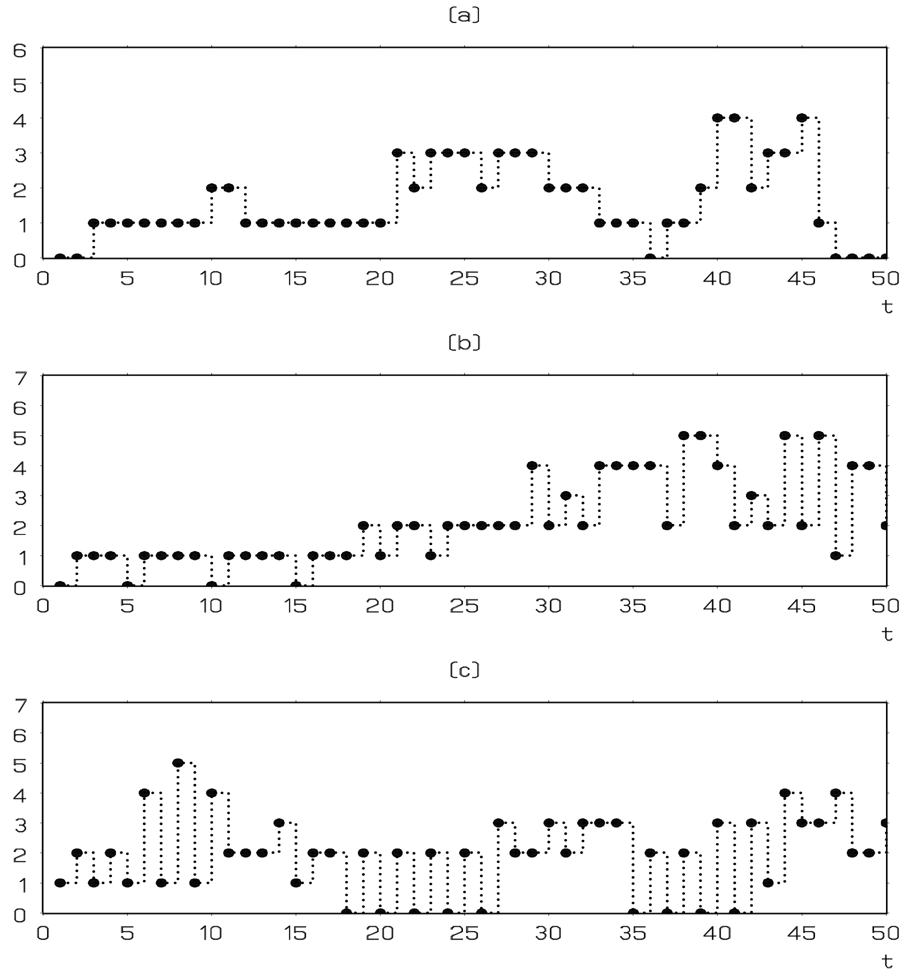

Figure 4: *Simulated sample paths for a PoINAR(2) process*

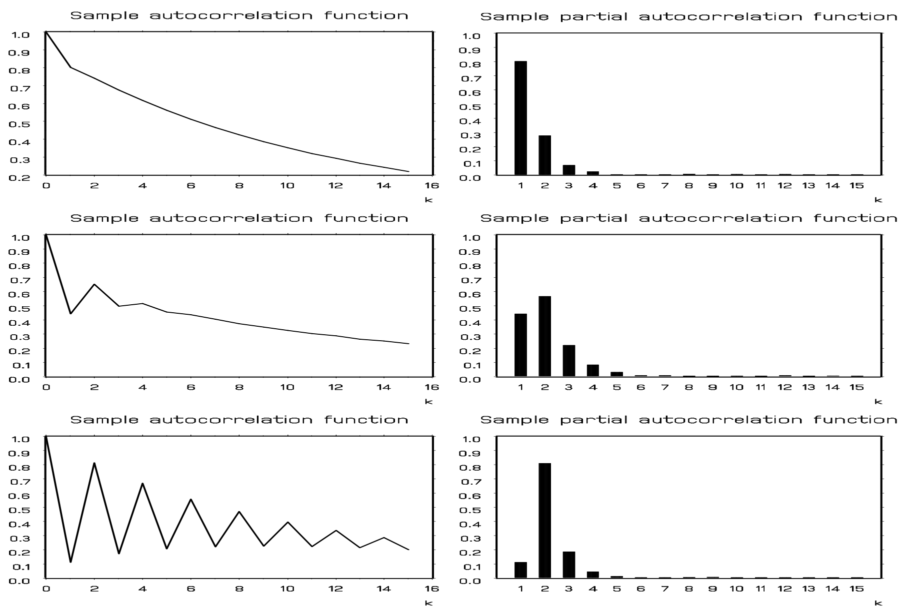

Figure 5: *Sample autocorrelation and partial autocorrelation functions for simulated sample paths of a PoINAR(2) process*

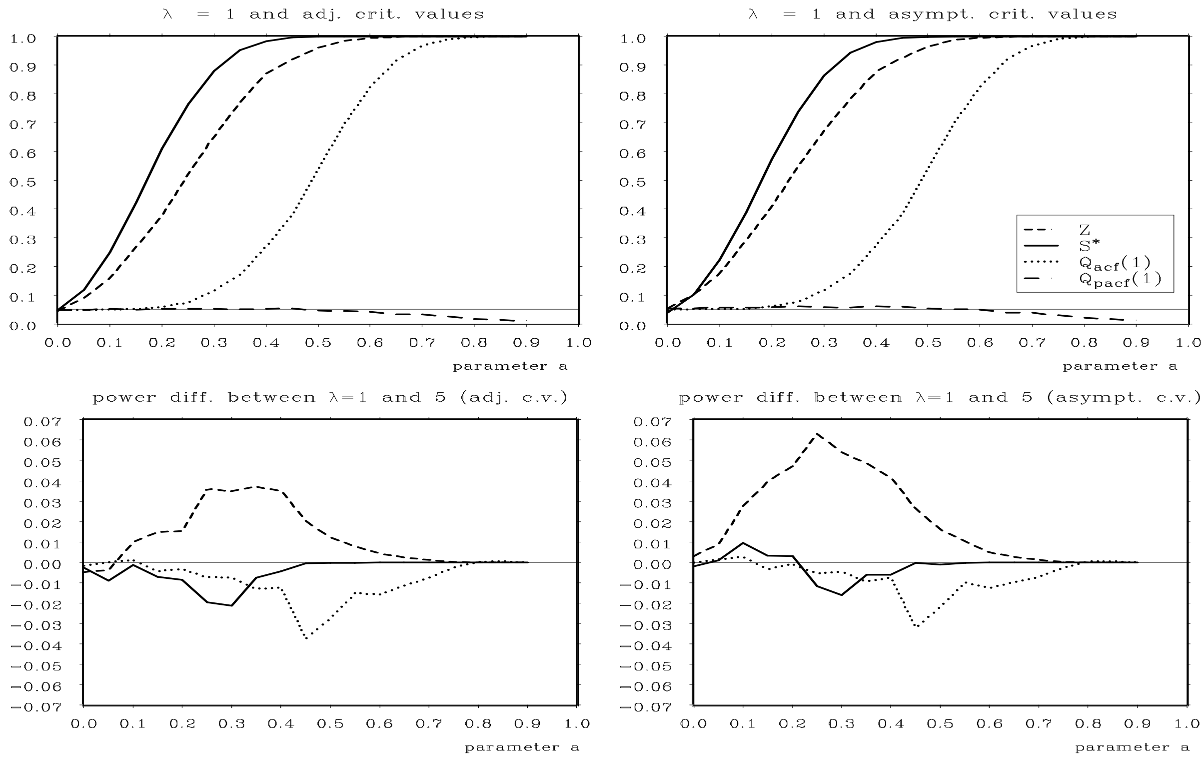

Figure 6: Empirical power curves for the various tests under the PoINAR(1) alternative at sample size $T = 100$ and a test level of 5%.

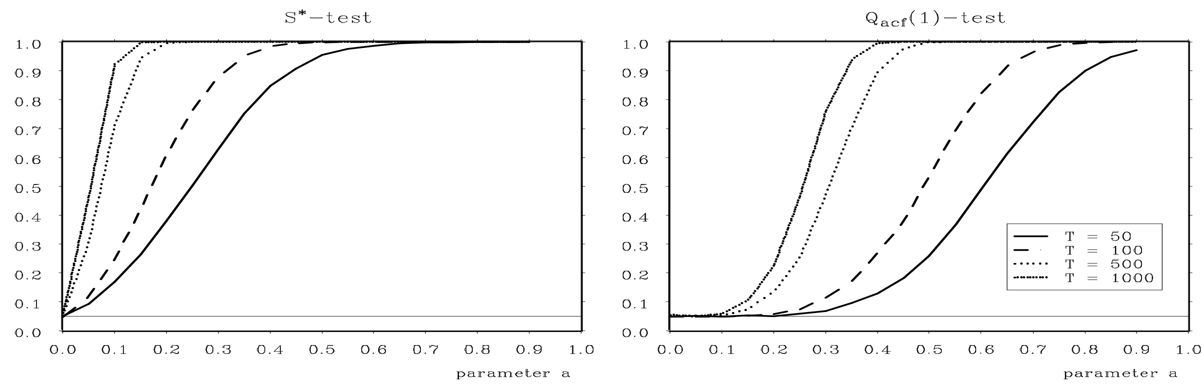

Figure 7: Empirical power curves for the $S^{\star}$ test and the $Q_{acf}(1)$ test under the PoINAR(1) alternative at various sample sizes and a test level of 5%.

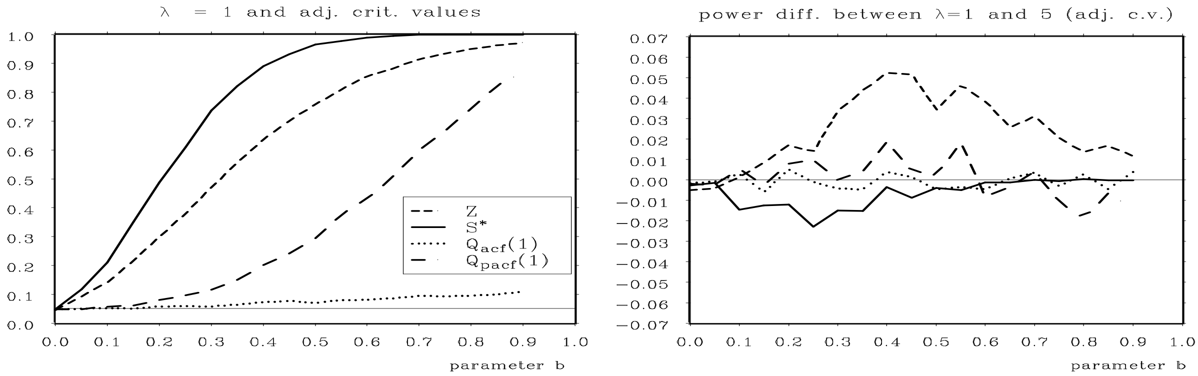

Figure 8: Empirical power curves for the various tests under the PoINMA(1) alternative at sample size $T = 100$ and a test level of 5%.

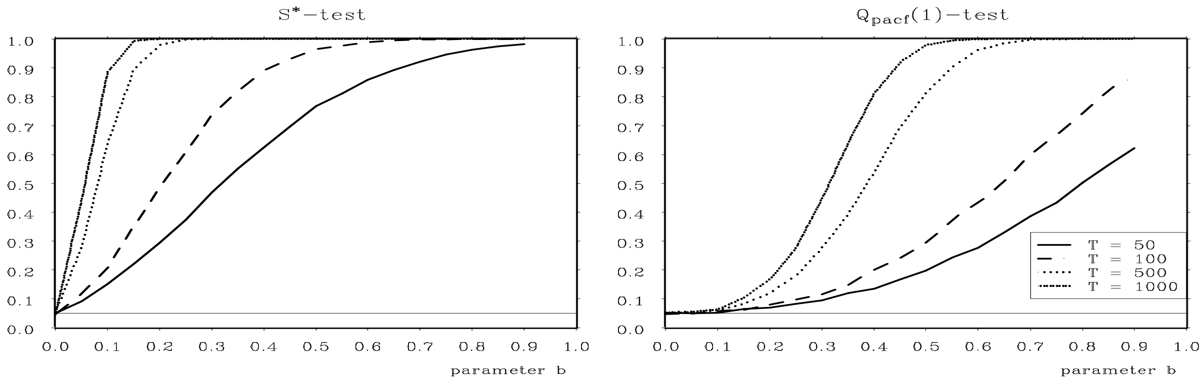

Figure 9: Empirical power curves for the $S^{\star}$ test and the $Q_{pacf}(1)$ test under the PoINMA(1) alternative at various sample sizes and a test level of 5%.

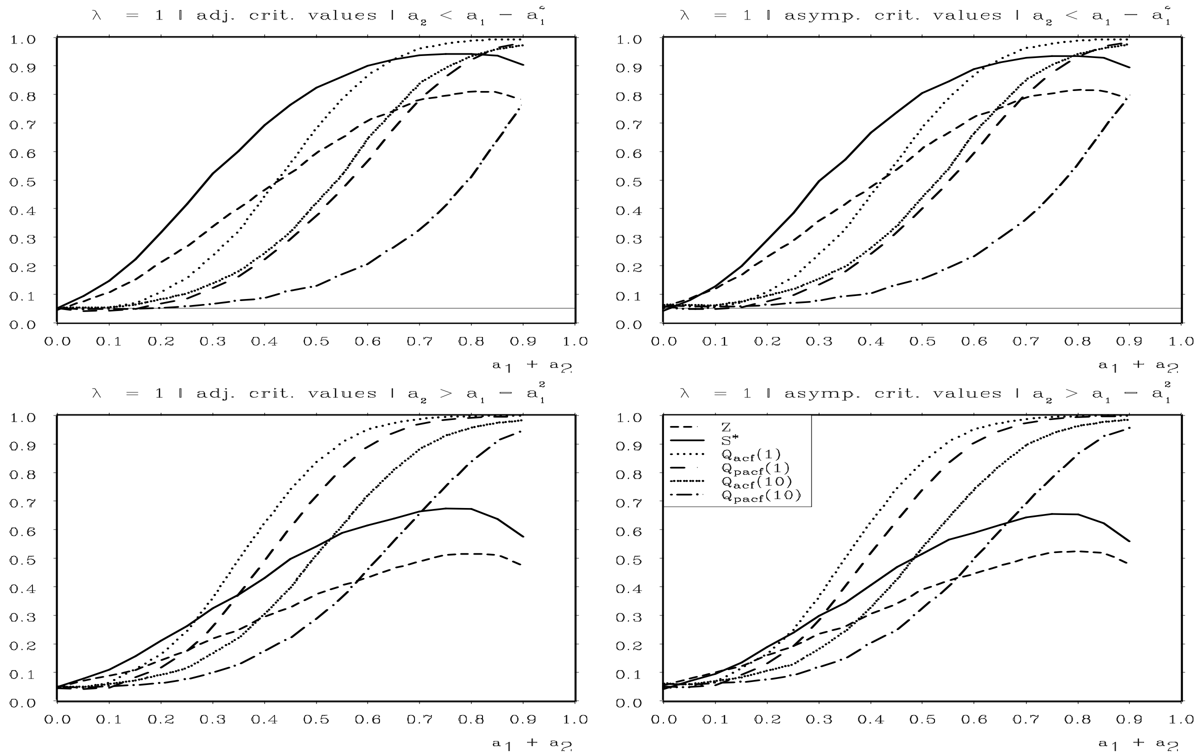

Figure 10: Empirical power curves for various tests under the PoINAR(2) alternative at sample size 100 and a test level of 5%.