Testing Serial Dependence in Time Series Models of Counts

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

This version: January 2000

Abstract

In the course of the analysis of time series of counts the need to test for the presence of a dependence structure arises regularly. Suitable tests for this purpose are analysed in this paper. Their size and power properties are evaluated under various alternatives among which the INARMA-processes play a prominent role. The results can be summarized as follows. (1) All the tests considered but one are robust against extra binomial variation in the data. (2) Newly proposed 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. (3) The tests considered are not powerful enough to distinguish between higher-order integer-valued autoregressive processes and the popular parameter-driven processes where a dynamic latent process introduces the serial dependence into the counts. The methods and findings of this study are applied to three data sets: the so called Fürth-data already analysed in the branching process literature, data on worker absenteeism and to polio incidence data.

Keywords: time series of counts; INARMA-models; 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, Newcaslte upon Tyne, NE1 7RU, England. E-mail: A.R.Tremayne@newcastle.ac.uk

The authors thank Dr. Tim Barmby, University of Newcastle, for supplying us with the absentee data.

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 or the monthly cases of rare infectious diseases in a specified area. Typically such series consist of positive (or zero) counts with a sample mean not higher than 10 rendering any continuous modelling inadequate. Several models that take the discreetness 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.

This paper focuses on the 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 a very interesting class of discrete valued processes with the ability not only to specify the dependence structure but also with a possibility to choose among a wide class of (discrete) marginal distributions. Although inherently nonlinear in nature, the INARMA-processes try to mimic the linear structure of the well known linear Gaussian ARMA-processes. There exists however a gap between the theoretical models for INARMA-processes and their practical application to time series of counts. It is the purpose of this paper to close this gap somewhat.

A natural first question in the analysis of time series of counts is whether the data exhibit a significant serial dependence or not. If that is not the case standard methods suitable in the iid-case can be applied, otherwise a more sophisticated analysis is called for. To equip the applied researcher with suitable tools to answer this question we discuss various standard tests of serial dependence and introduce a new class of tests derived from the branching process literature. The size and power properties of the various tests are examined in Monte Carlo experiments. Once a significant serial correlation is established in the data the next task is to identify the type of correlation structure. For this purpose we propose a testing strategy that supports this identification process and is especially designed to identify series that are suitable for the analysis with INARMA-processes. Again the power properties of this testing strategy is analysed in Monte Carlo experiments.

The paper is organized as follows: In Section 2 the basic INARMA-processes are described in some detail along with a parameter-driven process with an unobservable dynamic process entering the mean function serving as an alternative. 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

For the class of observation-driven models the simple integer-valued first order autoregressive (INAR(1)) and first order moving average (INMA(1)) processes are presented in greater detail as well as the INAR(2)-model in order to account for the possibility of higher order dependence. For the class of parameter-driven models the widely employed latent process model introduced by Zeger (1988) is briefly described.

2.1 The INAR(1)-process

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

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

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 variate with finite first (μ_W) and second moment (σ_W^2). W_t and X_{t-1} are supposed to be stochastically independent for all points in time.

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

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

The $Y_{i,t-1}$ are assumed to be iid 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 (Cf. Feller, 1968, ch. 12) as well as in the theory of stopped-sum distributions (Cf. Johnson, Kotz and Kemp, 1992, ch. 9). The close relationship between the Bienaymé-Galton-Watson branching processes and the INAR(1)-process will be described 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 who have been in the space at time $t-1$ and new entrants during the time span $(t-1, t]$. Each particle's probability to stay in the space is given by a. High values of a generate high correlation among subsequent observations and low values of a a low correlation.

Important properties of the INAR(1)-process are summarized below. More detailed information is provided in the papers mentioned above as well as in Alzaid and Al-Osh (1988). Due to the stationarity assumption the derivation of the first and second order moments is straightforward: $\mathrm{E}(X_t) = \mu_W/(1-a)$, $\mathrm{Var}(X_t) = (a\,\mu_W + \sigma_W^2)/(1-a^2)$ and the autocorrelation function is given by $\rho(k) = a^k$ for $k = 1, 2, \ldots$. The behaviour of the entire distribution of X_t is conveniently summarized in the probability generating function (pgf)

$$\mathcal{P}_{X_t}(s) = \mathcal{P}_{X_{t-1}}(1 - a + a\,s) \cdot \mathcal{P}_{W_t}(s) \ . \tag{3}$$

In contrast to the well known Gaussian processes the knowledge of the first and second order moments does not suffice to describe the dependence structure of the process entirely. Due to the Markovian property of the INAR(1)-process the relevant tool is the bivariate distribution function or the bivariate pgf

$$\mathcal{P}_{X_t,X_{t-1}}(s_1, s_2) = \mathcal{P}_{X_{t-1}}(s_1(1 - a - a\, s_2)) \cdot \mathcal{P}_{W_t}(s_2)\ . \tag{4}$$

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 due to Athreya and Ney (1972)

$$X_t = \sum_{i=1}^{X_{t-1}} Y_{i,t-1} + W_t\ , \qquad t = 0, \pm 1, \pm 2, \dots\ , \tag{5}$$

where $Y_{i,t-1}$ denotes lattice iid-random variables with pgf $\mathcal{P}_Y(s)$ and W_t denotes lattice iid-random variables with pgf $\mathcal{P}_W(s)$. $Y_{i,t-1}$ and W_t are assumed to be stochastically independent for all i and t. The structural equivalence is restricted to the subcritical case where $\mathrm{E}(Y_{i,t-1}) < 1$ holds. This feature can be exploited for purposes of parameter estimation and inference given the fact that a rich body of literature on BGWI-processes is available.

As mentioned in the introduction the INARMA-processes are characterized by their dependence structure and by their marginal distribution. 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) we assume that $W_t \sim \mathrm{Po}(\lambda)$ with $\lambda > 0$. The marginal distribution of the process X_t can now be derived (Cf. Al-Osh and Aly, 1992) by inserting the pgf of W_t, which is given by $\mathcal{P}_W(s) = \exp[-\lambda(1 - s)]$ into (3). The thus obtained functional difference then has to be solved iteratively:

$$\begin{aligned}\mathcal{P}_{X_t}(s) &= \exp\left[-\lambda(1 - s + a - as + a^2 - a^2 s + \ldots + a^T - a^T s)\right] \\ &\qquad \times \mathcal{P}_{X_0}(1 - a^T + a^T s) \\ &= \exp\left[-\frac{\lambda}{1-a}(1-s)(1-a^{T-1})\right] \cdot \mathcal{P}_{X_0}(1 - a^T + a^T s)\ .\end{aligned} \tag{6}$$

For $T \to \infty$ we obtain the final result:

$$\mathcal{P}_{X_t}(s) = \exp\left[-\frac{\lambda}{1-a}(1-s)\right] , \tag{7}$$

which shows that $X_t \sim \text{Po}(\lambda/(1-a))$. The resulting process will be denoted as PoINAR(1)-process.

Figure 1 depicts simulated sample paths of the PoINAR(1)-process. For all three panels $\text{E}(X_t) = 2$ was chosen while the autocorrelation parameter a and the Poisson parameter λ was allowed to vary across the panels. In the top panel a was set to 0.1 resulting in a very low autocorrelation and the sample path is quite erratic. In order to keep $\text{E}(X_t) = 2$ the value of λ had to be set to 1.8. In the center panel the parameter combination was $a = 0.5$ and $\lambda = 1$. In the bottom panel finally we set a to 0.9 and λ to 0.1. The high autocorrelation is quite evident in the bottom panel as well as the extremely low innovation rate λ. Note that the process returns back to its mean quite regularly as a consequence of the stationarity property introduced above.

2.2 The INMA(1)-process

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

$$X_t = b \circ W_{t-1} + W_t , \qquad t = 0, \pm 1, \pm 2, \dots , \tag{8}$$

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 variate with finite mean (μ_W) and variance (σ_W). The *thinning*-operation $b \circ W_{t-1}$ is defined as follows:

$$b \circ W_{t-1} \equiv \sum_{i=1}^{W_{t-1}} Y_{i,t-1} , \tag{9}$$

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

The INMA(1)-process can be illustrated using the example given already above: X_t denotes the number of particles in a well defined space at time t. According to the

INMA(1)-process this number is made up of particles who entered the space during the time span $(t-1, t]$ and survivors of the entrants of the time span $(t-2, t-1]$. Each elements has a fixed survival probability of b. In contrast to the INAR(1)-process *thinning* takes place only among the immigrants of $t-1$ not among all particles in the space. As a consequence some particles are forced to die automatically. This ensures that only two consecutive observations of X_t are correlated but not observations more than one time period apart. The resulting process is neither Markovian nor a BGWI-process.

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

$$\rho(k) = \begin{cases} \dfrac{b\,\sigma_W^2}{[b(1-b)\mu_W + (1+b^2)\sigma_W^2]} & \text{for} \quad k=1 \\ 0 & \text{for} \quad k>1 \end{cases} \tag{10}$$

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

The distribution of the INMA(1)-process can be summarized in its pgf

$$\mathcal{P}_{X_t}(s) = \mathcal{P}_{W_{t-1}}(1-b+bs) \cdot \mathcal{P}_{W_t}(s)\ . \tag{11}$$

As noted already above the INMA(1)-process is not Markovian and as a consequence the conditional distribution of $X_t|X_{t-1}, X_{t-2}, \dots, X_1$ depends not only on X_{t-1} but also on $X_{t-2}, X_{t-3}, \dots, X_1$. Due to a result of McKenzie (1988b) it is nevertheless possible to describe the entire dependence structure of INMA(1)-process by the bivariate pgf

$$\mathcal{P}_{X_{t-1},X_t}(s_1, s_2) = \mathcal{P}_{W_{t-2}}(1-b\circ + b\,s_1)\,\mathcal{P}_{W_{t-1}}\Big(s_1\,(1-b+b\,s_2)\Big)\,\mathcal{P}_{W_t}(s_2)\ . \tag{12}$$

From the pfg given in (11) it is obvious that the marginal distribution of the process is derived from the distributional assumption for the innovation process $\{W_t\}$. A natural first candidate is again the Poisson distribution. Employing $W_t \sim \mathrm{Po}(\lambda)$ the pgf of

[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} has been used in this derivation.

the PoINMA(1)-model is easily derived to be $\mathcal{P}_{X_t}(s)\exp[-\lambda(1+b)(1-s)]$. As a result $X_t \sim \text{Po}(\lambda(1+b))$.

Simulated sample paths for different parameter values of the PoINMA(1)-process are depicted in Figure 2. Again $\text{E}(X_t) = 2$ is fixed. The parameter combinations used to generate the graphs are: $b = 0.1$ and $\lambda = 1.8$ for the top panel; $b = 0.5$ and $\lambda = 1.33$ for the center panel, and $b = 1$ and $\lambda = 1$ for the bottom panel. The resulting autocorrelations are 0.09, 0.33 and 0.5. Again the effect of an increased autocorrelation leads to a smoothing of the sample path.

2.3 The INAR(2)-process

Higher order dependence in the data cannot be captured by the models discussed so far but only by higher order processes. The INAR(2)-process $\{X_t, t = 0, \pm1, \pm2, \ldots\}$, a seemingly natural extension of the INAR(1)-process, is defined in the usual manner

$$X_t = a_1 \circ X_{t-1} + a_2 \circ X_{t-2} + W_t \,, \qquad t = 0, \pm1, \pm2, \ldots \,. \tag{13}$$

The *thinning*-operation is in analogy to (2). To ensure the stationarity of the process, it is assumed 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} follows a multinomial distribution with parameters (a_1, a_2, X_{t-1}). A short description of the way the INAR(2)-process is simulated may serve to gain further inside into the structure of the model. At time t two binomial *thinning* operations are performed: $U_t \equiv a_1 \circ X_{t-1}$ and $V_t \equiv a_2 \circ (X_{t-1} - U_t)$. While U_t is employed in the next time point $t+1$ to generate the new value of the process X_{t+1}, it is not until the time point $t+2$ before V_t is involved in the generation of X_{t+2}. In general the formula for generating the process is given by $X_t = U_{t-1} + V_{t-2} + W_t$.

A direct consequence of the fact that thinned elements of X_t re-enter the process at different time points 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 contradiction to that obtained by Du and Li (1990) who do not employ an additional assumption for the *thinning* operation 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 the INAR(p)-process in general can be viewed as multitype branching process with immigration. As a consequence the classical branching process literature can be exploited for the analysis of higher order INAR-processes as well.

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

$$\rho(k) = a_1\, \rho(k-1) + a_2\, \rho(k-2) \qquad \text{for} \quad k \geq 2 \tag{14}$$

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

As is the case in the Gaussian ARMA-processes, the a_1/a_2-parameter space can be partitioned in an area, where the autocorrelation function 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 $a_2 < a_1 - a_1^2$ is satisfied, the autocorrelation function decays exponentially to zeros. The oscillating behaviour can be found in situations, 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 4 depicts simulated sample paths for the PoINAR(2)-process at different parameter combinations. For all three panels the sum of a_1 and a_2 was set to 0.9 and in order to fix the mean of the generated series to 2, the Poisson parameter λ was set to 0.2. In the top panel the $a_1 = 0.8$ and $a_2 = 0.1$. The corresponding sample autocorrelation function and sample partial autocorrelation function of this series is shown in top panel of Figure 5. The autocorrelation clearly exhibits no oscillating behaviour. This changes in the middle panel of Figure 4, where a simulated sample path is depicted that is based on the parameter values $a_1 = a_2 = 0.45$. The corresponding autocorrelation function is again in the middle panel of Figure 5. The two parts it is made up of is clearly visible. While up to $k = 5$ an oscillating behaviour is evident, for $k \geq 6$ the autocorrelation function decays exponentially. In the bottom panels of the two figures a situation, where $a_1 = 0.1$ and $a_2 = 0.8$ is shown. The corresponding autocorrelation function is oscillating an damps out.

2.4 A parameter-driven model for time series of counts

A widely used parameter-driven model for time series of counts was introduced by Zeger (1988). It is set up in the regression framework of generalized linear models and specifies the mean function of the process $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ by a log-linear predictor and a non-negative latent dynamic process $\{\varepsilon_t : t = 0, \pm 1, \pm 2, \ldots\}$. In the standard specification the counts given the latent process $\{\varepsilon_t\}$ are assumed to follow a Poisson distribution

$$X_t|\varepsilon_t \sim \text{Po}(\exp\{\mathbf{z}_t'\boldsymbol{\beta}\}\varepsilon_t) , \tag{15}$$

where $\{\boldsymbol{z}_t : t = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of possible time-varying covariates including a constant term and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)')$ is the vector of regression parameters. In a pure time series framework the covariates typically consist of a time trend and various cyclical components. The dynamic latent process, which is assumed to be independent of the process $\{X_t\}$ introduces both serial correlation as well as overdispersion. A simple specification of the latent process which prevents the mean function from becoming negative is to employ a Gaussian AR(1)-process for the logarithm of ε_t

$$\log \varepsilon_t = \delta \log \varepsilon_{t-1} + \nu_t \qquad \text{and } |\delta| < 1. \tag{16}$$

The innovations ν_t are assumed to be iid-N(μ_ν, σ_ν^2).

The mean of the process (μ_t) is given by $\mathrm{E}(X_t) = \exp(\mathbf{z}_t'\boldsymbol{\beta})$ while the variance, which is given by $\mathrm{Var}(X_t) = \mu_t + \sigma_\varepsilon^2\ \mu_t^2$ always exceeds the mean as long as $\sigma_\varepsilon^2 > 0$. The autocorrelation function

$$\rho(k) = \frac{\rho_\varepsilon(k)}{\{[1 + (\sigma_\varepsilon^2\mu_t)^{-1}][1 + (\sigma_\varepsilon^2\mu_{t-k})^{-1}]\}^{1/2}}\,, \tag{17}$$

where $\rho_\varepsilon(k) = \mathrm{Cor}(X_t, X_{t-k})$, is not independent of the regressors $\{\boldsymbol{z}_t\}$. Note that positive as well as negative autocorrelation is permitted in this model. This is in contrast to the INARMA-processes described above, where only positive autocorrelation is possible.

Additional properties of this process along with a general review of parameter-driven and observation driven-models is provided in Davis, Dunsmuir and Wang (1999). In the subsequent analysis this model which we should call the Zeger-model will be used to assess the power properties of the test of serial dependence to be introduced in the next section because of its prominent use in the literature.

A graph of the simulated data from the Zeger-model is depicted in Figure 6. The data were generated using a time trend as well as two cyclical components as regressors: $\boldsymbol{z}_t'\boldsymbol{\beta} = \beta_0 + \beta_1 t + \beta_2\cos(2\pi t/12) + \beta_3\sin(2\pi t/12)$. This is in the spirit of Zeger (1988) and Davis, Dunsmuir and Wang (1999). In both panels the mean of the process is set to 2. In the top panel the first order autocorrelation is around 0.1 and in the bottom panel it is about 0.5. Note that sample path in the top panel is quite volatile and regularly returns to zero. The higher correlation in the bottom panel is due to long runs of zeros. Correlations as high as 0.9 as given in Figures 1 and 2 are beyond the realm of the Zeger-model. This is due to the fact that the dependence is induced not directly but through the latent process. It can be shown that e.g. the first-order autocorrelation of the observed process $\{X_t\}$ is always less then or equal to the first-order autocorrelation of the underlying latent process (see e.g. Davis, Dunsmuir and Wang, 1999).

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 tests fail to take the discreetness of the data into account or break down due to the presence of multiple ties. All the tests analysed here are carefully chosen in order to cope with this very special situation.

The first test to be considered is the simple runs test. In order to apply the test to time series of counts however the original series has to be dichotomized on the basis of a specified criterion. It is often recommended the median being used for this task with observations which are identical to the sample median being discarded. But given the fact that stationary processes 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 use of the sample mean as a threshold, given the fact that the mean of discrete data will generally not be discrete.

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 nominal significance levels can regularly be attained through a randomized test design. As was confirmed by means of Monte Carlo experiments not published here from samples of size 40 onwards the much more convenient normal approximation

$$Z = \frac{R - 1 - \dfrac{2\,T_1(T-T_1)}{T}}{\left[\dfrac{2T_1(T-T_1)[2T_1(T-T_1)-T]}{T^2(T-1)}\right]^{1/2}}\,, \qquad (18)$$

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

The test design chosen here is one sided. This is motivated by the fact that under the alternative hypotheses of INAR(1), INAR(2) or INMA(1)-processes which are in the center of our interest no negative autocorrelation is possible. Hence the only sensible departure of the null is smaller number of runs than would be expected under the null.

Due to the discrete nature of the distribution of R it is often recommended (cf. e.g. Gibbons and Chakraborti 1992, p. 77) to use a continuity correction in the Z-statistic. The resulting test statistic

$$Z_{cc} = \frac{R - 0,5 - \dfrac{2T_1(T-T_1)}{T}}{\left[\dfrac{2T_1(T-T_1)[2T_1(T-T_1)-T]}{T^2(T-1)}\right]^{1/2}} \tag{19}$$

is asymptotically equivalent to the Z statistic. Due to the one sided test design the null hypothesis of serial independence is rejected in both cases if Z(resp.$Z_{cc} < z_\alpha$, where z_α is the relevant quantile of the standard normal distribution.

Another approach to test 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:

$$S = \left(\sqrt{T}\bar{x}\right)^{-1} \sum_{t=2}^{T} (x_{t-1} - \bar{x})(x_t - \bar{x}) , \tag{20}$$

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

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

$$S^{\star} = \sqrt{T} \frac{\sum_{t=2}^{T} (x_{t-1} - \bar{x})(x_t - \bar{x})}{\sum_{t=1}^{T} (x_t - \bar{x})^2} \tag{21}$$

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.

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 $S > s_\alpha$ 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 Venkatamaran (1982) and by Mills and Seneta (1989). The first statistic is based on the sample autocorrelations whereas the second one is based on the sample partial autocorrelations. Both statistics are in the spirit of portmanteau tests due to 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 series in order to test for serial dependence in the data (e.g. Pindyck and Rubinfeld, 1991, sec. 16.2) we advocate the use of goodness-of-fit tests for simple branching processes to the original series using the test statistics given below.

Restricted to first order processes the relevant implementation of the Venkatamaran (1982) statistic, given in section 6 of his paper is as follows:

$$Q_{acf}(1) = \hat{\varrho}_2^2 \, \frac{\left[\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^2}{\sum_{t=3}^{T}(x_t - \bar{x})^2(x_{t-2} - \bar{x})^2} \tag{22}$$

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 $Q_{acf}(1) \xrightarrow{d} \chi^2(1)$. To see this in a first step the limit distribution of the sample autocorrelation $\hat{\varrho}_k$ has to be derived. Since the summands in the numerator of $\hat{\varrho}_k$ are not independent but form a martingale difference with respect to the information set $\mathcal{F}_{t-1} = \{x_{t-1}, \ldots, x_1\}$ it is necessary to employ a martingale central limit theorem. It is then possible to show that $\sqrt{T}\,\hat{\varrho}_1 \xrightarrow{d} \mathrm{N}(0,1)$ and $T\,\hat{\varrho}_1^2 \xrightarrow{d} \chi^2(1)$ respectively for all $k = 1, 2, \ldots$. In a second step the probability limit of the fraction in (22) has to be derived. For the numerator a weak law of large numbers for independent variables is appropriate to establish that plim $[T^{-1}\sum_{t=1}^{T}(x_t - \bar{x})^2]^2 = \lambda^2$. The denominator consists of dependent summands who form a martingale with respect to $\mathcal{F}_{t-1}$. An appropriate weak law of large numbers helps to show that plim $T^{-1}\sum_{t=3}^{T}(x_t-\bar{x})^2(x_{t-2}-\bar{x})^2=\lambda^2$. This establishes the asymptotical neutrality of the fraction and closes the proof.

A further test based on the branching process literature is due to Mills and Seneta (1989). In the first order case the test statistic is given by

$$Q_{pacf}(1)=\hat{\phi}_2^2\,\frac{\left[\sum_{t=1}^{T}(x_t-\bar{x})^2\right]^2}{\sum_{t=3}^{T}(x_t-\bar{x})^2(x_{t-2}-\bar{x})^2}, \tag{23}$$

where $\hat{\phi}_2$ is the second order sample partial autocorrelation. Along the lines of the proof outlined above 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 first order sample autocorrelation whereas the $Q_{pacf}(1)$-test is based on the sample partial autocorrelation function. In parallel with AR- and MA-models for continuous time series it may be the case that the two tests can help us distinguish between an autoregressive and a moving average structure in the count data.

The two tests just introduced can easily be expanded to a portmanteau structure in order to capture higher order dependence in the data. It is hoped that the use of the $Q_{acf}(k)$- and the $Q_{pacf}(k)$-test with $k=1$ and $k>1$ can help distinguish between first and higher order processes. The portmanteau version of the $Q_{acf}(1)$-test is defined as follows:

$$Q_{acf}(k)=\sum_{i=1}^{k}\hat{\varrho}_{s_i+1}^2\,\frac{\left[\sum_{t=1}^{T}(x_t-\bar{x})^2\right]^2}{\sum_{t=s_i+2}^{T}(x_t-\bar{x})^2(x_{t-s_i-1}-\bar{x})^2}, \tag{24}$$

where $k\geq s_i\geq 1$ each i, $s_{i+1}-s_i\geq 2$ and k is arbitrary.

The corresponding portmanteau version of the $Q_{pacf}(1)$-test is given by

$$Q_{pacf}(k) = \sum_{i=1}^{k} \hat{\phi}_{s_i+1}^2 \frac{\left[\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^2}{\sum_{t=s_i+2}^{T} (x_t - \bar{x})^2(x_{t-s_i-1} - \bar{x})^2} \tag{25}$$

The appropriate limit distribution for both portmanteau versions of the tests is the χ^2-distribution with k degrees of freedom. This can easily seen since under the null hypothesis of independence both statistics consist of the sum of k independent $\chi^2(1)$ variables as derived above.

4 Empirical size properties of the tests

A Monte Carlo study has been conducted in order to analyse the the size properties of the various test statistics discussed in the last section. The number of Monte Carlo replications were 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 were generated for low level data ($\lambda = 1$) and higher level data ($\lambda = 5$). The sample sizes used were $50, 100, 500$ and 1000. In preliminary work we experimented with smaller sample sizes but various elements of unsatisfactory behaviour were in evidence. Add it to which additional complexities due to randomized devices of runs tests are required. For some additional information on this see Jung (1999, sec. 3.2). For the portmanteau tests $Q_{acf}(k)$ and $Q_{pacf)}(k)$ the number of summands k was set to 5.

Table 7 gives the rejection percentages of the different tests for for a nominal level of 5% as well as 1%. Numbers in boldface indicate that the value is outside the 95% confidence interval derived above. Several conclusions can be drawn. Most tests exhibit size distortions for the smallest sample size used in our simulation study. Whereas for

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

some tests the size distortions disappear for moderate sample sizes, for some test 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. Both versions of the score test behave in a similar fashion. The $Q_{acf}(1)$-test does not seem to be exhibits a slight tendency to underrejection while its portmanteau version $Q_{acf}(5)$ seems not to be affected by size distortions. The only test that tends to overreject is the $Q_{pacf}(k)$-test. Again no differences in the behaviour of the null distribution of the test statistics can be observed for different values of k. A variation in the level of the series does not seem to have systematic effects on the size behaviour of the test statistics under investigation. This analysis has been carried out for all conventional significance levels with qualitatively similar results. A full set of results is available from the authors if desired.

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 the negative binomial distribution was chosen under the null. Parameters of this distribution were set as to reflect a situation with modest overdispersion (variance-mean ratio of 1.5) and one with high extra binomial variation (variance-mean ratio 3). Table 7 gives the rejection frequencies of the different configurations of our Monte Carlo experiment again for the nominal sizes of 5% as well as 1%.

The results are qualitatively similar to those obtained under the Poisson assumption. A remarkable exception is the behaviour of the S-test. Even in the presence of modest overdispersion the test rejects the true null about twice as often as intended by the significance level. The nonstationarity of this statistic is evident from its behaviour in growing sample sizes. In the presence of high overdispersion the situation is even worse. This behaviour can be explained by looking at formula (20). In the presence of overdispersion in the data this statistic is not properly scaled anymore and therefore is of no practical use. A second test that is affected by the presence of overdispersion in the data is the $Q_{pacf}(k)$-test. For small and medium sample sizes in low level data ($\mathrm{E}(X) = 1$) especially the portmanteau version of the test exhibits quite a significant overrejection. As a consequence of this the value of k in the $Q_{pacf}(k)$-test should be chosen with great care in small and medium sample sizes.

As in the Poisson case changing the level of the data by increasing the Poisson parameter λ from 1 to 5 does not seem to have systematic effects on the size behaviour of the other test statistics under investigation. This analysis has been carried out too for all conventional significance levels with qualitatively similar results. Again a full set of results is available from the authors if desired.

Overall our findings can be summarized as follows: All the tests discussed in Section 3 exhibit only minor size distortions as long as the data are equidispersed. Overdispersion can lead to noticeable size distortions in the newly proposed $Q_{acf}(k)$- and Q_{pacf}-test and a nonstationary behaviour of the S-test. As a consequence in the presence of overdispersion the use of the S-test cannot be recommended and the for the $Q_{acf}(k)$- and $Q_{pacf}(k)$-tests size adjusted critical have to be used.

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. The design parameters for the experiments were chosen in such a way that the results for the different alternatives are comparable. 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 were used. Sample sizes of 50, 100 and 500 were chosen. No power calculations were conducted for sample size 1000 since all the power curves show high power at all significance levels for this sample size as a result of their consistency property to be shown below. The empirical power curves presented here are based on a nominal size of 5%. Calculations for other nominal sizes show no substantive differences in the results. They are available from the authors upon request.

The PoINAR(1)-process serves as the first alternative hypothesis to be analysed. Empirical power curves were calculated for series, where the Poisson parameter λ was set to 1 and series where λ was set to 5. The autocorrelation parameter a was successively raised from 0, resembling the situation under the null hypothesis to 0.90 in steps of 0.05. The results are summarized in Figures 7 and 8. Figure 7 depicts the power curves for all tests discussed in Section 3 at sample size 100. In the top panels λ was set to 1 and in the bottom panels to 5. The power curves in the left hand size panels are based on size adjusted critical values while in the right hand size panels asymptotical critical values were used.

It is quite obvious that the S^* version of the score test dominates all the other tests over the entire parameter space considered in the experiments. Freelands S-test exhibits quite good power properties too but fails to handle situations with high autocorrelation at small sample sizes satisfactorily. This non-monotonic behaviour of the empirical power of the S-test can be attributed to the to the different scaling factors used in (20) and (21) respectively. While the scaling factor employed in the S-test is the sample mean, the $S^\star$-test is scaled by the sample variance. Although both statistics provide an estimate for the Poisson parameter λ the variation of the two estimators differs markedly as soon as the dependence is very high. More details are provided in Jung (1999, sec. 3.2).

The power properties of the simple runs test is quite remarkable. Going from the left hand panels to the right hand panels it can be seen that the while the runs test with continuity correction (Z_{cc}) and the runs test without continuity correction (Z) have identical power curves based on the size adjusted critical values, the size distortions of the Z_{cc}-test leads to a slightly worse performance with respect to power based on the asymptotic critical values. As a result of this finding it may be recommended not to use the continuity correction because this can lead to inferior results. Both versions of the runs test are affected by the level of the underlying data series. The higher the level of the data, the lower the power of the runs test. 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 proposed null hypothesis and a loss of power. From comparing the upper panels of Figure 7 with the lower panels it can be inferred that only the power properties of the runs test is affected by the level of the underlying data.

Finally the power properties of the tests based on the sample partial autocorrelation differs markedly from that based on the sample autocorrelation. Both the $Q_{pacf}(1)$-test as well as its portmanteau version $Q_{pacf}(5)$ do not seem to possess any power at all and exhibit a tendency to be biased for high values of the parameter a. The $Q_{acf}(1)$-test performs slightly better as compared to its portmanteau version $Q_{acf}(5)$ but the power properties of both versions of this test are clearly inferior to the score tests and the runs test. The consistency of the $S^{\star}$-test and the two versions of the $Q_{acf}(k)$-test is demonstrated in Figure 8.

The second alternative analysed is the PoINMA(1)-process. The only difference in the design of the Monte Carlo experiments as compared to the PoINAR(1)-alternative is the fact that distance between the null hypothesis of iid-Poisson random variables is increased by successively raising the value of the parameter b of the INMA(1)-process (8). The results are summarized in Figures 9 and 10.

From the inspection of Figure 9 several points emerge. While it seems to be harder to identify a first-order moving average structure in the data the ability of the different tests to detect it is comparable to the first-order autoregressive case discussed above. Again both version of the score test tend to dominate the other tests employed in this study. The simple runs test is also doing remarkably well here. As hoped for in Section 3 the $Q_{pacf}(k)$-tests are now picking power and as shown in Figure 10 are consistent under the PoINMA(1)-alternative. In contrast to that the $Q_{acf}(k)$-tests are now virtually unaffected by the growing dependence structure. Both versions of this test are not consistent under the PoINMA(1)-alternative. While the portmanteau version of the Q_{pacf}-tests is picking up power at a slower rate as compared to the $Q_{pacf}(1)$-statistic the situation is reversed in the Q_{acf}-case.

The power curves in the left hand panels in Figure 9 are based the adjusted critical values obtained in Section 4 while in the right hand panels asymptotical critical values were used. The size distortions influence on the empirical power is restricted to the alternatives, which are very close to the null hypothesis. A noticeable difference arises in the case of the runs test. Due to the size distortions in the continuity corrected version of the runs test the empirical power of this test statistic is somewhat lower than that of of the uncorrected statistic when asymptotical critical values are employed. As under the PoINAR(1)-alternative it can be seen when comparing the top to the bottom panels that the only test affected by the change in the level of the data is runs test. The explanation for this phenomenon is already provided above.

As mentioned above, the power properties of the $Q_{acf}(k)$-tests and the $Q_{pacf}(k)$-tests differ markedly. While the former possess power against the PoINAR(1)-alternative, the latter do not. In the case of the PoINMA(1)-alternative the power properties of the two tests are reversed. This is due to the fact that the $Q_{acf}(k)$-tests are determined on the basis of the second, fourth, sixth and so on order sample autocorrelations as shown in (22) and (24) respectively. The the $Q_{pacf}(k)$-tests on the other side are based on the second, fourth, sixth and so on order sample partial autocorrelations. This is quite evident from equations (23) and (25). Under the PoINAR(1)-alternative the second, fourth, sixth and so on order sample autocorrelations pick up the dependence structure present in the data while the second, fourth, sixth and so on order sample partial autocorrelations will be close to zero under this alternative. Under the PoINMA(1)-alternative the situation is reversed. The second, fourth, sixth and so on order sample autocorrelation coefficients will not pick up any dependence due to the special kind of dependence in the this process, while the second, fourth, sixth and so on order sample partial autocorrelation coefficients are positive due to the dependence structure in the data. This very different behavior of the test statistics can subsequently be used to provide a tool to distinguish between INAR-processes and INMA-processes and will be discussed in greater detail below.

The last INARMA-process to be analysed here is the PoINAR(2)-process (13). The design of the Monte Carlo experiments is identical to the one chosen above. The distance from the null hypothesis of iid-Poisson variables is increased 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 sample autocorrelation function as described in section 2.3 above, two sets of experiments were conducted. For the first set, the combinations of a_1 and a_2 was chosen in such a way as the a_2 was kept below the parabolic boundary depicted in Figure 3. In the second set series with a oscillating autocorrelations were generated.

As long as $a_2 < a_1 - a_1^2$ the autocorrelation properties of the PoINAR(2)-process are very similar to those of the PoINAR(1)-process. This is evident from the empirical power curves depicted top panels of Figure 11. But in contrast to the first order process the second order autoregressive process exhibits a exponentially decaying partial autocorrelation function. This can bee seen in the top right panel of Figure 5. Consequently under the PoINAR(2)-alternative the $Q_{pacf}(k)$-test is picking up power as expected. All the tests considered in this study are therefore consistent under this alternative.

As soon as $a_2 > a_1 - a_1^2$ the situation changes quite dramatically as evident from the bottom panel of Figure 11. The score tests as well as both versions of the runs test are not very powerful against this constellation of the PoINAR(2)-alternative. All the other tests exhibit a superior power performance over the relevant area of the parameter space. The $Q_{acf}(1)$-test seems to be possess the test with the highest power under those circumstances. The second most powerful test is the $Q_{pacf}(1)$-test while the portmanteau versions of both tests lag somewhat behind. The explanation for these results is quite evident. The higher the value of a_2 as opposed to the value of a_1 grows the higher the probability of the elements of the second *thinning* operation that re-enter the process with a lag of 1 to survive. This can bee seen from the bottom panel in Figure 4. Every other time point the process is staying at a level for some time. It seems as if the process is jumping between two regimes. A regime of high values and a regime of low values. As a_1 is so small only a very low first order autocorrelation arises. But second order autocorrelation is very high. All the tests that rely solely on the first order autocorrelations as the score test do must inevitably fail to capture the dependence structure correctly. The failure of the simple runs test is also evident from the bottom panel of Figure 4. The process seems to be erratic with jumps across the sample mean of the process occurring quite often and consequently decreasing the probability of longer runs.

Variations in the design parameter of the Monte Carlo do not produce substantively different results with respect to the power properties of the test. Our findings up to this stage can be summarized as follows. Both versions of the score test and the simple runs test are surprisingly powerful when it comes to detecting a dependence structure that is restricted to the first order. As soon as the dependence structure is more complicated both classes of test fail to work. The specially designed $Q_{acf}(k)$- and Q_{pacf}-tests are relatively slow in picking up first order dependence alone. But as soon as higher order dependence is in the data they are certainly the more powerful alternatives. The quite distinctive power behaviour of the two versions of this class of tests can potentially be used to allow for a classification of the correlation structure among the various INARMA-processes.

To confront the tests of serial dependence described here with a data situation that is very popular among applied researchers and whose dependence is generated in a very different way as in the INARMA-processes, the Zeger-model as briefly outlined in Section 2.4 is used. The setup of the Monte Carlo experiments is chosen in such a way as to be comparable to the experiments conducted under the INARMA-alternatives. Specifically the parameters of the process, $\boldsymbol{\beta}, \sigma_\varepsilon$ and δ were set to reflect a low level data situation (with the mean of the process equal to 1 and 5 respectively under the null hypothesis of idd-Poisson variables). As evident from the autocorrelation function of the process (17) control of the autocorrelation of the observed data is not directly possible but only through the parameter δ, the autocorrelation parameter of the latent process. Consequently this parameter is varied over the relevant parameter space, that is from 0 to 0.9 in the experiments and will be serve as the abscissae in the graphs of the empirical power curves. This is in contrast to the graphs depicted above and should be kept in mind when interpreting the results of the Monte Carlo. It should also be noted that negative values for ρ although theoretically possible are not investigated here. As denoted above significant negative autocorrelation in the data is evidence that INARMA-models are not an appropriate assumption for the data generating process and therefore is not analysed here any further. As was also mentioned in Section 2.4 the latent process introduces not only autocorrelation but also overdispersion. As the value of ρ increases the dispersion in the data rises in an nonlinear fashion. In the setup of the experiments it was paid attention that the level of overdispersion was kept a reasonable level.

Along with the values of the design parameters already discussed above in the Zeger-model explanatory variables have to be fixed. In the pure time series framework of this study sensible regressors are the time trend and various cyclical components. We conducted Monte Carlo experiments for models with different regressors. One for a model with time trend only, one with a single cyclical component and one with a time trend and two cyclical components. This last model served for the generation of the data in Figure 6 and is given in Section 2.4. Since the no substantive different results for the power analysis of the different specifications of the regressor matrix emerged, we only describe the results for the Zeger-model with trend and two cyclical components as regressors here.

The empirical power curves for the various test are depicted in Figure 12. Note that the S-test is discarded from the analysis because of its nonstationary behaviour under the null hypothesis in the presence of overdispersion. It can be seen that all the tests considered in the analysis pick up power and are consistent under the Zeger-alternative. The $S^{\star}$-test is dominating over the entire parameter spaced analysed here. The runs test is also quite powerful whereas the $Q_{acf}(k)$-test is picking up power very slowly as the autocorrelation in the latent process rises. As under other alternatives discussed above, the $Q_{pacf}(k)$-test is the least powerful test examined. It emerges that no exploitable behaviour of the different tests distinguishes the power analysis under the Zeger-alternative as compared to the PoINAR(2)-alternative. Up to this stage of the analysis no testing strategy based on the tests of Section 3 can help us to distinguish between higher order INARMA-process and the Zeger-alternative. As a result of this analysis, further evidence is needed.

6 Applications

The methods and findings of this paper are applied to three data sets. The first data set is well known in the branching process literature and has been analyzed 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 the German physicist Reinhold Fürth (Cf. Fürth, 1918). The sample mean of the series is 1.59 and the sample variance is 1.51, providing no indication for the presence of overdispersion. The Q-test for the presence of overdispersion of Davis, Dunsmuir and Wang (1999) supports this intuition. The results of the various tests for serial correlation are summarized in first column of Table 7. All tests are significant at the all 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 that neither the simple PoINAR(1)-model nor the simple PoINMA(1)-model is adequate to describe the Fürth-data satisfactorily. This result is in accordance with Mills and Seneta (1989), who found a poor fit for the fitted branching process model, which is equivalent with our PoINAR(1)-model.Since no significant overdispersion can be found in the data, fitting a PoINAR(2)-model to this data set should be considered.

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 a slight hint toward the presence of overdispersion. The Q-test of Davis, Dunsmuir and Wang rejects the null hypothesis of no overdispersion at the 10% level but not on the 5% level. The results of the various tests of serial dependence are summarized in column 2 of Table 7. All tests, but the $Q_{pacf}(k)$-tests are highly significant. In accordance with the findings in the last section we conclude that a first-order autocorrelation structure is present in the data. The simple PoINAR(1)-model therefore should provide a suitable basis for a further analysis of this data set.

The last data set analysed here is an extension of the polio incidence data of Zeger (1988). The original data set covered the time period 01/1970 to 12/1983. The extended data set adds observations for 1984 and 1985 giving us a sample size of 192. The sample mean of this series is 1.24 and the sample variance 3.20. Consequently the Q-test for the presence of overdispersion of highly significant. The results for the various tests are summarized in column 3 of Table 7. Some points are worthwhile to mention. The unusual high difference between the results of the S-test and the $S^{\star}$-test is certainly due to the presence of overdispersion. As explained in Section 3 this affects the performance of the S-star test quite dramatically. The $Q_{acf}(k)$-tests are highly significant whereas the $Q_{pacf}(k)$-tests are not. According to our findings in the last section an INAR(1)-model might suffice to describe the dependence structure in the data. Due to the presence of overdispersion a suitable marginal distribution has to be chosen. Interestingly the polio incidence data have been analysed by Zeger using an parameter-driven approach as described in Section 2.4.

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 wide variety of alternative data generating process.

The main findings of the paper are: (1) The INAR(2)-process provides a parsimoniously parameterized model to capture higher-order dependence in the data. Certain parameter combinations in the INAR(2)-model lead to oscillating autocorrelation functions. (2) Extra binomial variation in the data leads to a nonstationary behaviour of the null distribution of the score test proposed by Freeland. The other tests considered in this study are only affected to a minor extend by the presence of overdispersion. (3) The often recommended continuity correction for the simple runs test leads to severe size distortions under the null resulting in inferior power properties for this test based on the asymptotic critical values. (4) Newly proposed tests based on the sample autocorrelation and the sample partial autocorrelation can be utilized to distinguish between a first-order autoregressive, a first-order moving-average and a higher-order autoregressive type of dependence structure in time series of counts. The portmanteau version of theses test do not provide additional inside in this respect. (5) The tests analysed in this study are not powerful enough to distinguish between a higher-order autoregressive INAR-structure and a dependence that is introduced through a dynamic latent process. For this purpose it is necessary to combine the tests described in this study along with tests for the presence of a latent process as e.g. discussed in Davis, Dunsmuir and Wang (1999).

References

AL-OSH, M.A. and E.-E.A.A. ALY (1992) First order autoregressive time series with negative binomial and geometric marginals. *Commun. Statist.-Theor. Meth. 21*, 2483-2492.

Al-OSH, M.A. and A.A. ALZAID (1987) First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis 8*, 261-275.

Al-OSH, M.A. and A.A. ALZAID (1988) Integer-valued moving average (INMA) process. *Statistical Papers 29*, 281-300.

ALZAID, A.A. and M.A. AL-OSH (1988) First-order integer-valued autoregressive (INAR(1)) process: distributional and regression properites. *Statistica Neerlandica 42*, 53-61.

ALZAID, A.A. and M.A. AL-OSH (1990) An integer-valued p-th order autoregressive structure (INAR(p)) process. *Journal of Applied Probability 27*, 314-324.

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

BOX, G.E.P. and D.A. PIERCE (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association 65*, 1509-1526.

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 Y. WANG (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.; G. GAUTHIER and A. LATOUR (1995) Branching processes with immigration and integer-valued time series. *Serdica Mathematical Journal 21*, 123-136.

DU, J.-G. and Y. LI (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 S. CHAKRABORTI (1992) *Nonparametric statistical inference* (3rd edn). Marcel Dekker, New York.

JOHNSON, N.L.; S. KOTZ and A.W. KEMP (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.

MCKENZIE, E. (1985) Some simple models for discrete variate time series. *Water Ressources Bulletin 21*, 645- 650.

MCKENZIE, E. (1988a) Some ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability 20*, 822-835.

MCKENZIE, E. (1988b). The distributional structure of finite moving-average processes. *Journal of Applied Probability 25*, 313-321.

MILLS, T.M. and E. SENETA (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 D.L. RUBINFELD (1998). *Econometric Models and Economic Forecasts* (4th edn). McGraw-Hill, New York.

STEUTEL F.W. and K. VAN HARN (1979) Discrete analogues of self-decomposability and stability. *The Annals of Probability 7*, 893-899.

VENKATAMARAN, 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 J. WOLFOWITZ (1940) On a test whether two samples are from the same population. *Annals of Mathematical Statistics 11*, 147-162.

ZEGER, S. (1988) A regression model for time series of counts. *Biometrika 75,* 621-629.

size	λ	test	$T = 50$	$T = 100$	$T = 500$	$T = 1000$
5%	1	Z	**5.25**	5.10	5.03	5.08
		Z_{cc}	**3.69**	**4.07**	**4.52**	**4.71**
		S	**3.49**	**4.00**	**4.51**	**4.75**
		$S^{\star}$	**3.45**	**4.02**	**4.54**	**4.78**
		$Q_{acf}(1)$	5.01	**5.18**	5.02	5.06
		$Q_{acf}(5)$	4.90	5.10	5.10	4.98
		$Q_{pacf}(1)$	**6.12**	**5.77**	**5.15**	**5.13**
		$Q_{pacf}(5)$	**6.94**	**6.15**	**5.32**	**5.13**
	5	Z	**5.28**	**4.80**	5.08	5.04
		Z_{cc}	**3.43**	**4.11**	**4.61**	**4.72**
		S	**3.33**	**3.86**	**4.54**	**4.77**
		$S^{\star}$	**3.30**	**3.87**	**4.56**	**4.76**
		$Q_{acf}(1)$	**4.86**	**4.96**	5.01	5.02
		$Q_{acf}(5)$	**4.85**	**4.81**	4.96	5.02
		$Q_{pacf}(1)$	**5.72**	**5.35**	5.09	5.04
		$Q_{pacf}(5)$	**6.29**	**5.49**	5.07	5.05
1%	1	Z	**1.10**	**1.08**	1.02	1.00
		Z_{cc}	**0.77**	**0.81**	**0.88**	**0.92**
		S	**0.81**	**0.87**	**0.94**	0.96
		$S^{\star}$	**0.61**	**0.75**	**0.91**	0.96
		$Q_{acf}(1)$	**0.71**	**0.90**	1.00	1.00
		$Q_{acf}(5)$	1.04	**1.10**	1.03	1.00
		$Q_{pacf}(1)$	1.04	**1.11**	**1.05**	1.02
		$Q_{pacf}(5)$	**1.46**	**1.28**	**1.07**	**1.02**
	5	Z	**0.89**	**1.05**	0.98	1.00
		Z_{cc}	**0.71**	**0.74**	**0.88**	**0.92**
		S	**0.70**	**0.80**	**0.92**	0.99
		$S^{\star}$	**0.50**	**0.68**	**0.91**	0.98
		$Q_{acf}(1)$	**0.69**	**0.82**	0.97	0.98
		$Q_{acf}(5)$	1.04	1.00	0.99	1.01
		$Q_{pacf}(1)$	0.98	0.96	1.00	0.99
		$Q_{pacf}(5)$	**1.21**	1.03	0.98	1.00

Table 1: *Rejection percentages of the tests under the iid-Poisson assumption at a nominal level of* 5% *and* 1%.

size	E(X)	test	modest overdispersion				high overdispersion			
			$T=50$	$T=100$	$T=500$	$T=1000$	$T=50$	$T=100$	$T=500$	$T=1000$
5%	1	Z	**5.32**	4.90	5.07	5.06	**5.31**	**5.11**	**5.13**	5.02
		Z_{cc}	**3.58**	**4.08**	**4.58**	**4.72**	**3.53**	**4.20**	**4.56**	**4.82**
		S	**9.14**	**10.62**	**12.40**	**12.88**	**18.71**	**21.70**	**25.61**	**26.83**
		$S^\star$	**3.60**	**4.18**	**4.80**	**4.83**	**3.95**	**4.59**	**5.14**	**5.27**
		$Q_{acf}(1)$	**5.16**	**5.39**	**5.25**	**5.17**	**5.66**	**6.90**	**6.45**	**5.98**
		$Q_{acf}(5)$	4.93	**5.52**	**5.66**	**5.42**	**8.53**	**6.98**	**6.51**	**7.48**
		$Q_{pacf}(1)$	**6.75**	**6.26**	**5.44**	**5.29**	**8.69**	**8.66**	**6.88**	**6.23**
		$Q_{pacf}(5)$	**8.81**	**7.57**	**6.08**	**5.58**	**14.43**	**13.23**	**9.61**	**8.04**
	5	Z	**5.23**	4.80	5.03	4.95	**5.21**	4.91	4.95	4.94
		Z_{cc}	**3.42**	**4.17**	**4.58**	**4.66**	**3.56**	**4.06**	**4.47**	**4.61**
		S	**9.49**	**10.97**	**12.58**	**12.77**	**21.97**	**24.29**	**27.09**	**27.75**
		$S^\star$	**3.33**	**3.98**	**4.59**	**4.57**	**3.56**	**4.02**	**4.65**	**4.82**
		$Q_{acf}(1)$	4.93	4.97	5.02	5.03	**4.85**	**5.14**	**5.13**	5.07
		$Q_{acf}(5)$	**4.78**	4.90	4.94	5.09	**4.71**	**5.14**	**5.28**	**5.22**
		$Q_{pacf}(1)$	**5.91**	**5.41**	**5.11**	5.06	**6.09**	**5.80**	**5.29**	**5.16**
		$Q_{pacf}(5)$	**6.58**	**5.67**	**5.11**	**5.19**	**7.66**	**6.56**	**5.55**	**5.34**
1%	1	Z	**1.05**	**1.09**	**1.08**	1.00	**1.14**	**1.15**	1.03	**1.06**
		Z_{cc}	**0.74**	**0.72**	**0.94**	**0.91**	**0.74**	**0.84**	**0.92**	**0.94**
		S	**4.07**	**4.86**	**5.73**	**5.80**	**13.23**	**15.82**	**19.16**	**20.13**
		$S^\star$	**0.72**	**0.90**	**1.09**	1.02	0.97	**1.28**	**1.42**	**1.38**
		$Q_{acf}(1)$	**0.67**	**0.94**	**1.11**	**1.08**	**0.56**	**1.33**	**1.79**	**1.62**
		$Q_{acf}(5)$	0.97	**1.15**	**1.23**	**1.16**	**0.84**	**1.63**	**2.44**	**2.10**
		$Q_{pacf}(1)$	**1.16**	**1.26**	**1.21**	**1.14**	**1.52**	**2.11**	**2.01**	**1.75**
		$Q_{pacf}(5)$	**2.05**	**1.74**	**1.39**	**1.23**	**4.47**	**3.98**	**2.95**	**2.35**
	5	Z	**0.86**	1.03	0.99	0.99	**0.95**	0.99	0.98	1.02
		Z_{cc}	**0.69**	**0.69**	**0.90**	**0.91**	**0.71**	**0.75**	**0.87**	**0.92**
		S	**4.06**	**4.77**	**5.52**	**5.57**	**15.69**	**17.71**	**20.21**	**20.85**
		$S^\star$	**0.52**	**0.70**	**0.88**	**0.90**	**0.65**	**0.81**	0.98	**1.05**
		$Q_{acf}(1)$	**0.69**	**0.80**	0.96	0.98	**0.61**	**0.87**	1.03	1.02
		$Q_{acf}(5)$	1.01	1.03	0.97	1.04	**0.94**	**1.06**	**1.08**	1.03
		$Q_{pacf}(1)$	0.99	0.98	0.99	1.00	0.99	**1.11**	**1.09**	**1.06**
		$Q_{pacf}(5)$	**1.35**	**1.07**	1.02	1.04	**1.70**	**1.40**	**1.15**	**1.10**

Table 2: *Rejection percentages of the tests under overdispersion at a nominal level of* 5% *and* 1%.

test	Fürth data	Absentee data	Polio data
Z	-10.71^{***}	-12.40^{***}	-3.27^{***}
Z_{cc}	-10.67^{***}	-12.36^{***}	-3.18^{***}
S	14.12^{***}	17.78^{***}	10.37^{***}
$S^{\star}$	14.93^{***}	16.34^{***}	4.05^{***}
$Q_{acf}(1)$	40.02^{***}	108.86^{***}	7.30^{***}
$Q_{acf}(5)$	72.41^{***}	315.10^{***}	12.13^{**}
$Q_{pacf}(1)$	17.51^{***}	$3.65^{\star}$	1.70
$Q_{pacf}(5)$	24.77^{***}	$14.84^{\star}$	3.74

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

Table 3: *Results of the various tests applied to three data sets*

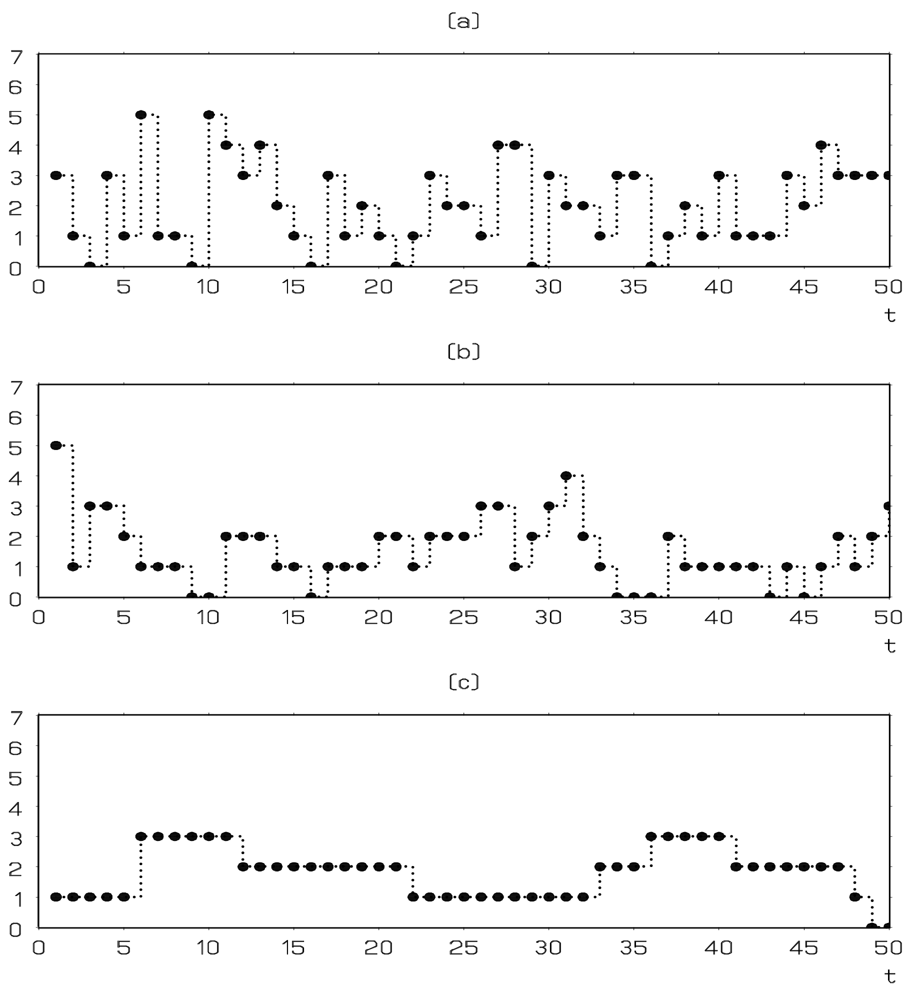

Figure 1: Simulated sample paths for the PoINAR(1)-process

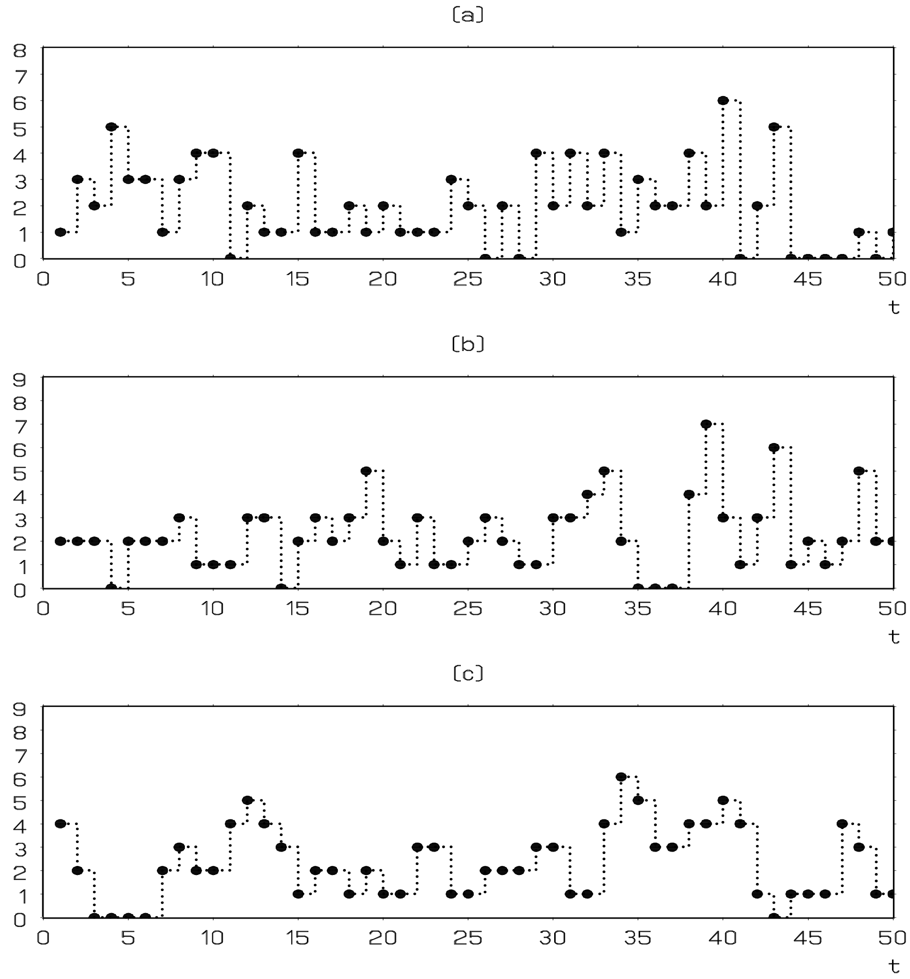

Figure 2: Simulated sample paths for the PoINMA(1)-process

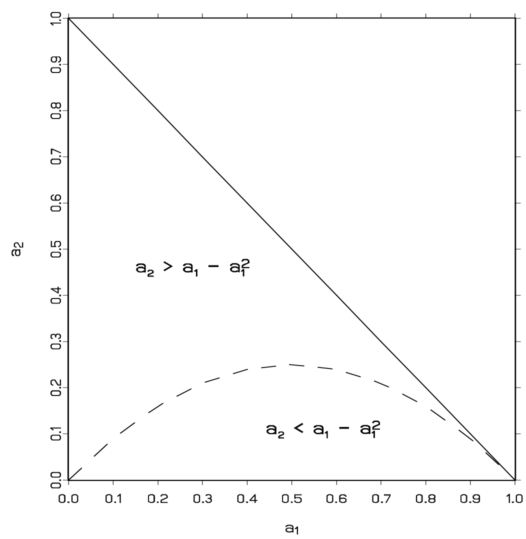

Figure 3: Partition of the a_1/a_2-parameter space in the PoINAR(2)-process

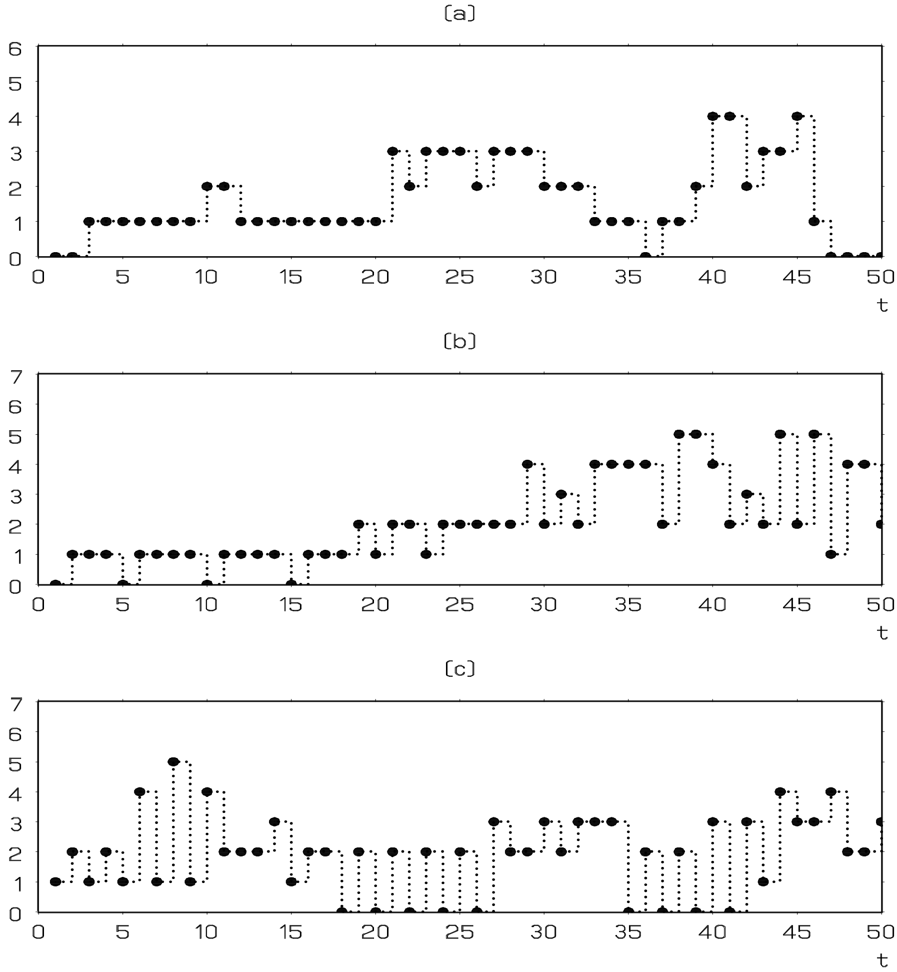

Figure 4: Simulated sample paths for the PoINAR(2)-process

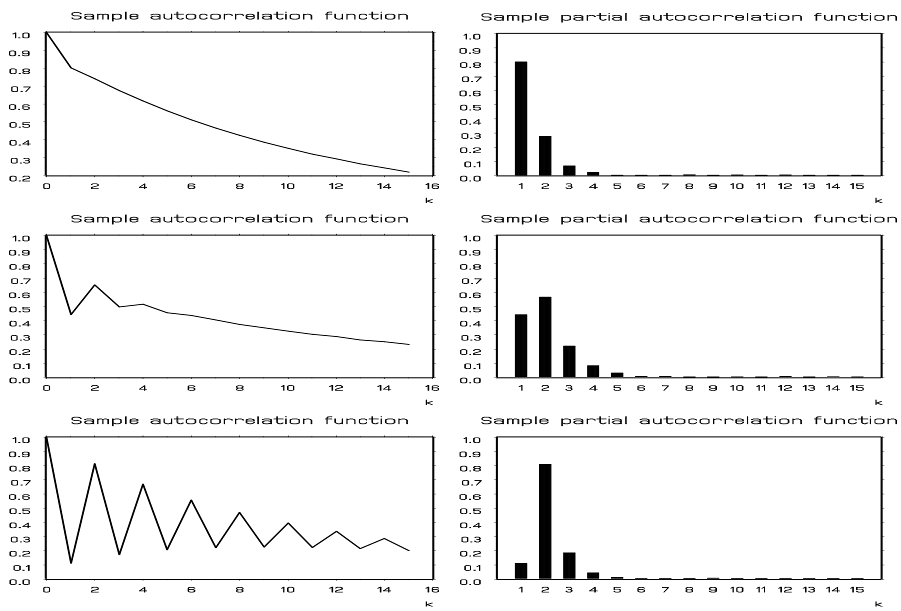

Figure 5: Sample autocorrelation and partial autocorrelation functions for the simulated sample paths of the PoINAR(2)-process

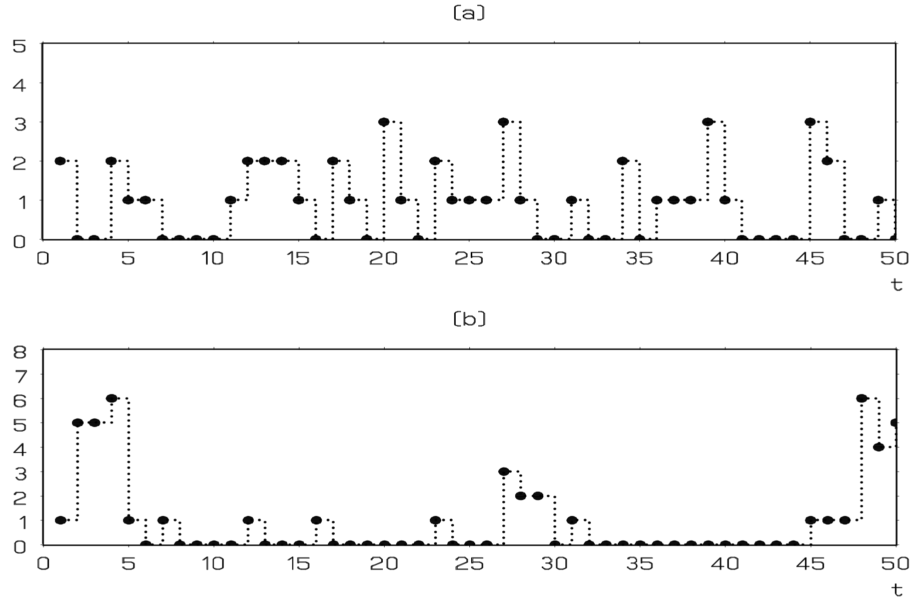

Figure 6: Simulated sample paths for the Zeger-model

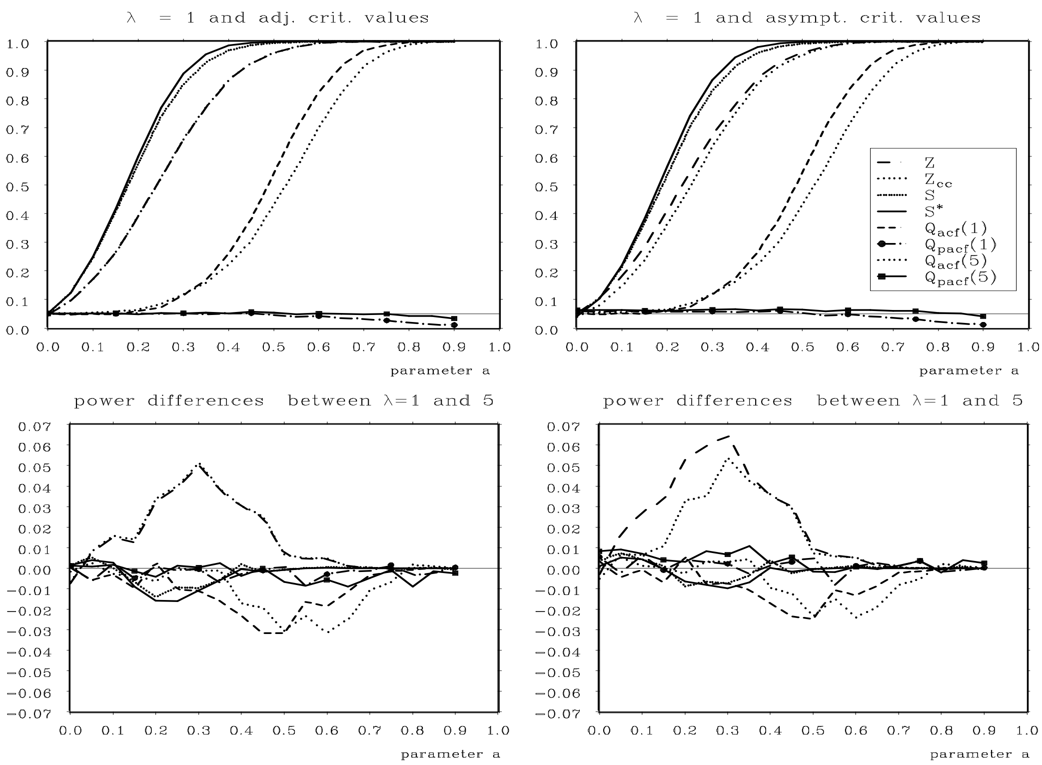

Figure 7: Empirical power curves for the various tests under the PoINAR(1)-alternative at sample size $T = 100$.

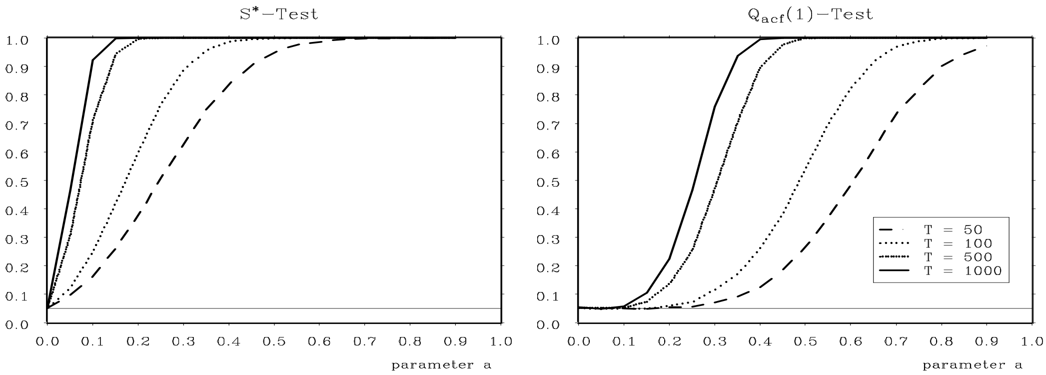

Figure 8: Empirical power curves for the $S^{\star}$-test and the $Q_{acf}(1)$-test under the PoINAR(1)-alternative at various sample sizes.

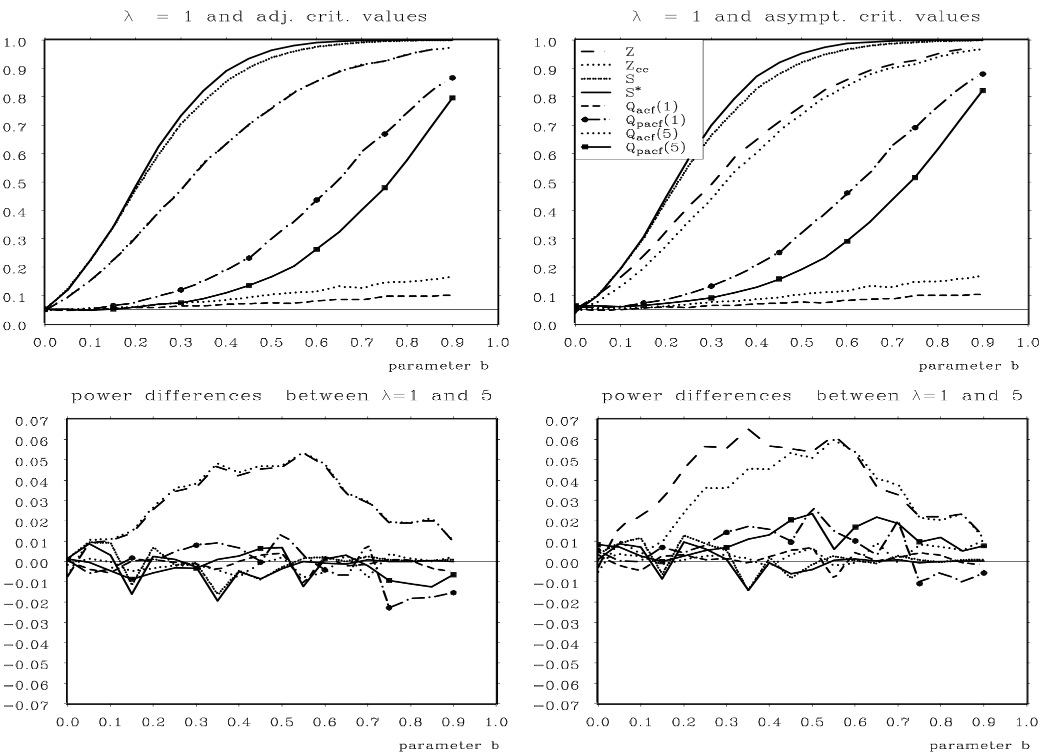

Figure 9: Empirical power curves for the various tests under the PoINMA(1)-alternative at sample size $T = 100$.

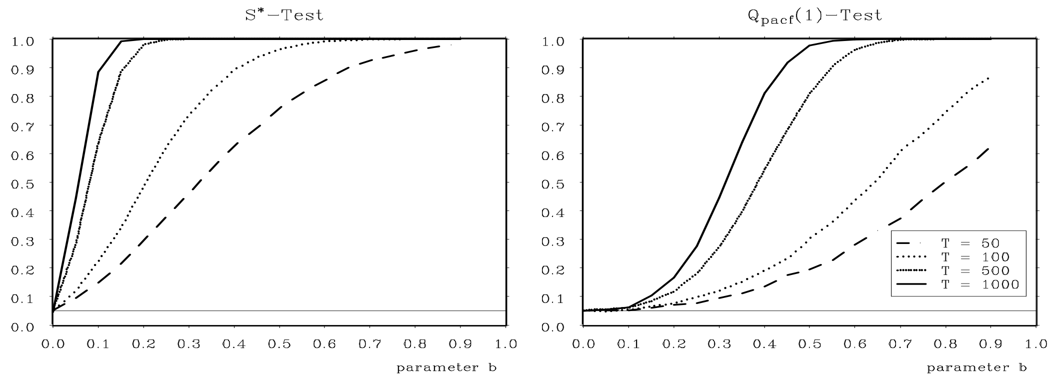

Figure 10: Empirical power curves for the $S^{\star}$-test and the $Q_{pacf}(1)$-test under the PoINMA(1)-alternative at various sample sizes.

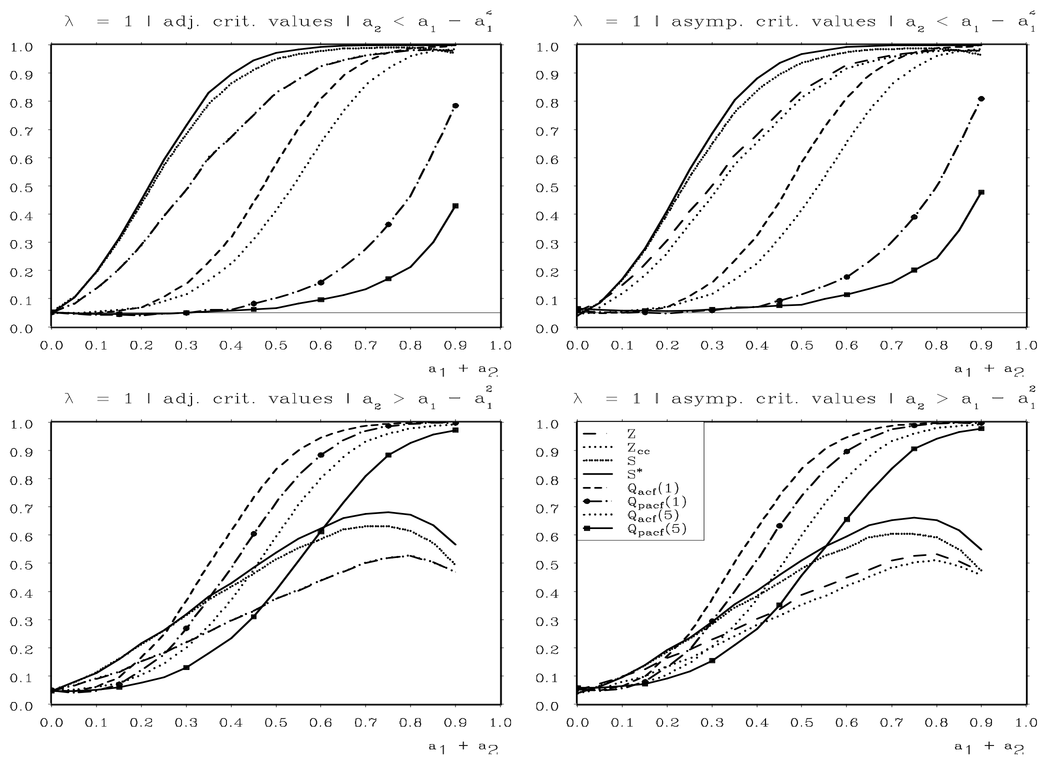

Figure 11: Empirical power curves for various tests under the PoINAR(2)-alternative at sample size 100.

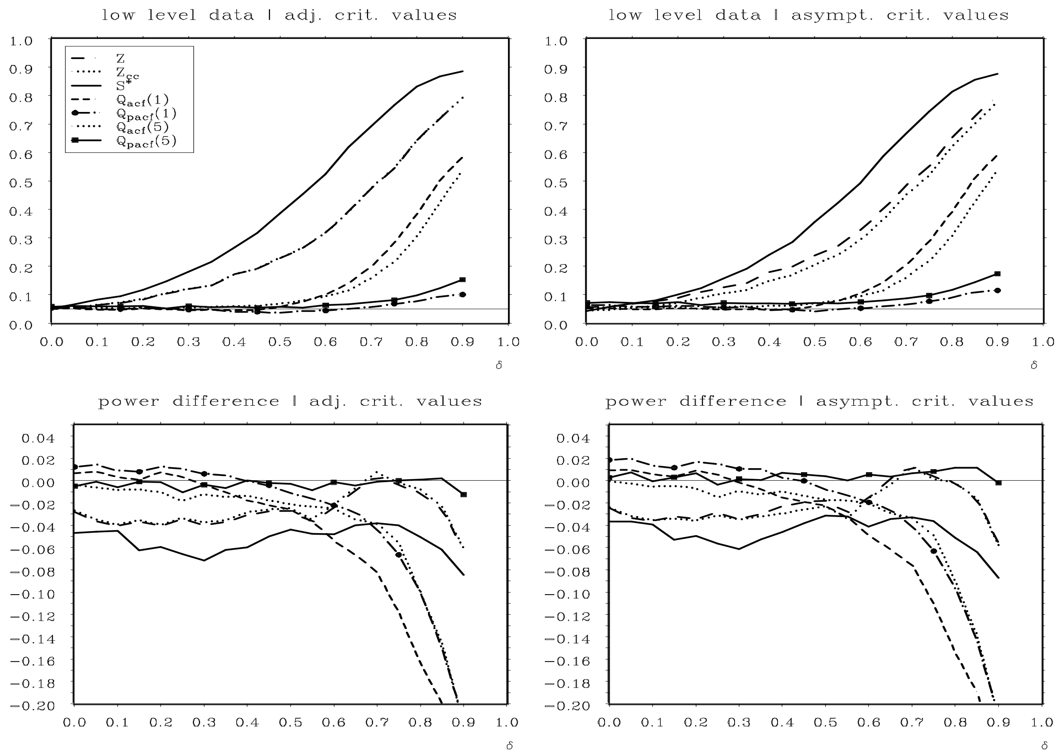

Figure 12: Empirical power curves for the various tests under the Zeger-alternative at sample size $T = 100$.