stats

MDPI

Article

INARMA Modeling of Count Time Series

Christian H. Weiß [1,*], Martin H.-J. M. Feld [1], Naushad Mamode Khan [2] and Yuvraj Sunecher [3]

1 Department of Mathematics and Statistics, Helmut Schmidt University, 22043 Hamburg, Germany; martin27748@yahoo.de
2 Department of Economics and Statistics, University of Mauritius, Reduit 80837, Mauritius; n.mamodekhan@uom.ac.mu
3 School of Business, Management and Finance, University of Technology, La Tour Koenig 11134, Mauritius; yuvisun@yahoo.com
* Correspondence: weissc@hsu-hh.de; Tel.: +49-40-6541-2779

Received: 29 March 2019; Accepted: 31 May 2019; Published: 3 June 2019

check for updates

Abstract: While most of the literature about INARMA models (integer-valued autoregressive moving-average) concentrates on the purely autoregressive INAR models, we consider INARMA models that also include a moving-average part. We study moment properties and show how to efficiently implement maximum likelihood estimation. We analyze the estimation performance and consider the topic of model selection. We also analyze the consequences of choosing an inadequate model for the given count process. Two real-data examples are presented for illustration.

Keywords: INARMA models; maximum likelihood estimation; model selection; model adequacy

1. Introduction

When dealing with stationary real-valued time series, the autoregressive moving-average (ARMA) models constitute a popular baseline model [1]. Besides the special case of purely autoregressive models, also full ARMA models are commonly used in practice, i.e., ARMA models that include a moving-average part (of typically low order). The integer-valued counterpart to the ARMA model, the INARMA model for ARMA-like count time series, dates back to McKenzie [2], Al-Osh and Alzaid [3], see Chapters 2 and 3 in Weiß [4] for a recent survey. In contrast to the ordinary ARMA models, research and applications nearly exclusively concentrate on the purely autoregressive INAR models (a few works also consider pure moving-average-type models, i.e., INMA models, like Brännäs and Hall [5], Brännäs and Quoreshi [6], Aleksandrov and Weiß [7]), but autoregressive INARMA models with an additional MA-component are rarely used in practice. This might be caused by the fact that there is only little work on stochastic properties of full INARMA models which, in turn, complicates the application of such models in practice. However, it seems to be mainly due to practical issues, e.g., it is not clear how to efficiently estimate the model parameters or to select the model order. While maximum likelihood (ML) estimation is easily done for the Markovian INAR(p) models, it is assessed until now that "the inclusion of a moving average component renders maximum likelihood estimation infeasible" [8] (p. 1). The only likelihood-related approach discussed until now in the literature seems to be the Markov chain Monte Carlo (MCMC) method for INARMA processes developed by Neal and Subba Rao [9], Enciso-Mora et al. [10], which has been used for conducting inference on model parameters and forecast distributions as well as for model selection, also see Alzahrani et al. [11]. Furthermore, Enciso-Mora et al. [10] propose to use the Expectation Maximization (EM) algorithm for ML estimation, but only for the purely autoregressive INAR models. Wheatley et al. [12] conjecture the asymptotic equivalence between an INARMA process and an ARMA point process, which, in turn, would allow to adapt their EM scheme for ARMA point processes to the INARMA case. In the present article, however, we show that a direct numerical maximization of the log-likelihood function is tractable by considering the proposed efficient implementation of likelihood computation.

In Section 2, we give the definition of INARMA models and a more detailed discussion of those models, which are relevant for our study, namely INAR(p) and INARMA(1,1) models. Then we turn to the question of model fitting. In Section 3, we show that it is possible to efficiently implement ML estimation also for the INARMA(1,1) model. Furthermore, we compare the performance of the ML approach for INAR(p) and INARMA(1,1) models with simulations. Section 4 discusses the task of model selection across different types of INARMA models, and Section 5 analyzes the consequences of fitting the wrong type of INARMA model to the actual data-generating process (DGP). Our findings are also illustrated with two real-data examples in Section 6. Finally, we conclude in Section 7.

2. INARMA Models for Count Time Series

The basic idea behind INARMA models is to modify the ordinary ARMA recursion by replacing multiplications by so-called 'binomial thinning' operations $\{\rho \circ Z\}$ with $\rho \in [0,1]$, see [13], where

$$\rho \circ Z | Z \ \sim \mathrm{Bin}(Z, \rho). \tag{1}$$

The conditional binomial distribution immediately implies that

$$E(\rho \circ Z) = \rho\, E(Z), \quad \mathrm{Var}(\rho \circ Z) = \rho(1-\rho)\, E(Z) + \rho^2\, \mathrm{Var}(Z). \tag{2}$$

Using this integer-valued operation as a substitute for the ordinary multiplication, the formal definition of INARMA(p,q) models is given by the recursion [14]

$$Y_t = \alpha_1 \circ Y_{t-1} + \ldots + \alpha_p \circ Y_{t-p} + R_t + \beta_1 \circ R_{t-1} + \ldots + \beta_q \circ R_{t-q}, \tag{3}$$

where (R_t) constitutes a sequence of independent and identically distributed (i.i.d.) random variables, commonly referred to as the 'innovations'. We denote the innovations' mean by $E(R_t) = \lambda$ and the variance by $\mathrm{Var}(R_t) = \nu\lambda$. So $\nu = \mathrm{Var}(R_t)/E(R_t)$ expresses the dispersion ratio, which equals 1 in the case of Poisson innovations (equidispersion).

Regarding Equation (3), some caution is necessary. Since the binomial thinning operations are random operations, one has to carefully think about the joint distribution among the thinnings, and the joint distribution of the thinnings to the other random variables in Equation (3). To obtain feasible stochastic properties, one commonly formulates several independence assumptions. This is exemplified in the sequel by discussing the purely autoregressive INAR(p) models as well as the INARMA(1,1) model.

2.1. A Brief Survey on INAR Models

After its introduction by McKenzie [2], Al-Osh and Alzaid [3], the INAR(1) model was extended to a pth-order INAR(p) model by Alzaid and Al-Osh [15], Du and Li [16], but in two different ways. Both extensions follow the recursion $Y_t = \alpha_1 \circ Y_{t-1} + \ldots + \alpha_p \circ Y_{t-p} + R_t$, but while the INAR$(p)$ model by Du and Li [16] assumes independence among all thinnings, independence to the innovations, and independence to $(Y_s)_{s<t}$ for the thinnings at time t, the model by Alzaid and Al-Osh [15] assumes a conditional multinomial distribution for the thinnings experienced by Y_t, in the sense that

$$(\alpha_1 \circ Y_t, \ldots, \alpha_p \circ Y_t) \ \sim \ \mathrm{Mult}(Y_t;\, \alpha_1, \ldots, \alpha_p).$$

Since only the INAR(p) model by Du and Li [16] leads to the well-known Yule-Walker equations for the autocorrelation function (ACF) of an ordinary AR(p) model, these are typically preferred in practice and also considered in the remaining article.

For the existence of the INAR(p) process by Du and Li [16], $\alpha_\bullet := \sum_{j=1}^{p} \alpha_j < 1$ has to be assumed. Then one obtains a pth-order Markov process, where the transition probabilities are computed as a convolution between p binomial distributions and the innovations' distribution. In particular, one obtains

$$P(Y_t = k \mid Y_{t-1} = l) \;=\; \sum_{j=0}^{\min\{k,l\}} \binom{l}{j}\, \alpha^j\, (1-\alpha)^{l-j} \cdot P(R_t = k-j) \tag{4}$$

for the INAR(1) model, and

$$\begin{aligned} P(Y_t = k \mid Y_{t-1} = l_1, Y_{t-2} = l_2) \;=\; & \sum_{j_1=0}^{\min\{k,l_1\}} \sum_{j_2=0}^{\min\{k-l_1,l_2\}} \binom{l_1}{j_1}\, \alpha_1^{j_1}\, (1-\alpha_1)^{l_1-j_1} \\ & \cdot \binom{l_2}{j_2}\, \alpha_2^{j_2}\, (1-\alpha_2)^{l_2-j_2} \cdot P(R_t = k-j_1-j_2) \end{aligned} \tag{5}$$

for the INAR(2) model, see, e.g., Weiß [4]. Furthermore, the stationary marginal mean is given by $E(Y_t) = \lambda/(1-\alpha_\bullet)$, and the variance satisfies

$$\mathrm{Var}(Y_t) \cdot \big(1 - \textstyle\sum_{i=1}^{p} \alpha_i\, \rho(i)\big) \;=\; E(Y_t)\, \textstyle\sum_{j=1}^{p} \alpha_j (1-\alpha_j) \;+ \nu\, \lambda, \tag{6}$$

also see Equation (2). The ACF $\rho(k) = \mathrm{Corr}(Y_t, Y_{t-k})$ is obtained from the ordinary AR(p)'s Yule-Walker equations

$$\rho(k) \;=\; \textstyle\sum_{i=1}^{p} \alpha_i\, \rho(|k-i|) \qquad \text{for } k \geq 1. \tag{7}$$

In particular, the INAR(1) model satisfies $\rho(k) = \alpha_1^k$.

2.2. Moment Properties of INARMA(1,1) Model

While the stochastic properties for INAR models as surveyed in Section 2.1 are well known among practitioners, such properties are not readily available for the INARMA(1,1) model. The INARMA(1,1) model is defined as

$$Y_t \;=\; \alpha_1 \circ Y_{t-1} + \beta_1 \circ R_{t-1} + R_t, \tag{8}$$

where the ARMA parameters $\alpha_1, \beta_1 \in [0,1]$ are used within the binomial thinning operations, and these are assumed to be executed independently of each other. Hence,

$$\mathrm{Cov}(\alpha_1 \circ W, \beta_1 \circ Z) \;=\; \alpha_1 \beta_1\, \mathrm{Cov}(W, Z). \tag{9}$$

This model is a special case of the GINARMA(p,q) model by Dion et al. [14]. For $\beta_1 = 0$, model (8) reduces to the INAR(1) model by McKenzie [2], Al-Osh and Alzaid [3], and for $\alpha_1 = 0$, it reduces to the INMA(1) model recently surveyed by Aleksandrov and Weiß [7]. But model (8) differs from the INARMA models considered by Dungey et al. [8], McKenzie [17], Bracher [18]. The INARMA(1,1)-like model by McKenzie [17] includes the AR-component only with lag 2 (i.e., with Y_{t-2} instead of Y_{t-1}), and the one by Dungey et al. [8] applies the β_1-thinning to R_t instead of R_{t-1}. In Bracher [18], the MA-component is constructed in a different way.

The relation between the counting series Y_t and the innovation terms R_t is summarized through the following lemma.

Lemma 1. *For the INARMA(1,1) model* (8)*,*
(a)

$$Cov(Y_t, R_{t+h}) = \begin{cases} Var(R_t) & h = 0, \\ (\alpha_1 + \beta_1) Var(R_t) & h = -1, \\ 0 & otherwise. \end{cases}$$

(b) $\quad Cov(\alpha_1 \circ Y_{t-1}, \beta_1 \circ R_{t-1}) \;=\; \alpha_1 \beta_1 Cov(Y_{t-1}, R_{t-1}) \;=\; \alpha_1 \beta_1 Var(R_t).$

Using Lemma 1, it is proven in Appendix A, under stationary moments condition, that

$$E(Y_t) \;=\; \frac{(1+\beta_1)\,\lambda}{1-\alpha_1}, \tag{10}$$

also see Corollary 2 in Dion et al. [14]. Analogously, we obtain

$$\mathrm{Var}(Y_t) \;=\; \frac{\big[\alpha_1(1+\beta_1)+\beta_1(1-\beta_1)\big]\,\lambda+\big[\beta_1(\beta_1+2\alpha_1)+1\big]\,\nu\lambda}{1-\alpha_1^2}. \tag{11}$$

The lag-h autocovariance with $h \geq 1$ is given by

$$\mathrm{Cov}(Y_t, Y_{t+h}) \;=\; \alpha_1^h\,\mathrm{Var}(Y_t)+\alpha_1^{h-1}\,\beta_1\,\nu\lambda, \tag{12}$$

see Appendix A for the derivation.

3. ML Estimation for INARMA(1,1) Models

ML estimation is easily implemented for the INAR models surveyed in Section 2.1, because these are Markov models such that the (conditional) likelihood function

$$L(\boldsymbol{\theta}) \;=\; P(Y_T=y_T,\ldots,Y_2=y_2 \mid Y_1=y_1)$$

factorizes as

$$L(\boldsymbol{\theta}) \;=\; \textstyle\prod_{t=p+1}^{T} P(Y_t=y_t \mid Y_{t-1}=y_{t-1},\ldots,Y_{t-p}=y_{t-p}),$$

see Appendix B.2.1 in Weiß [4]. The required transition probabilities are computed like in Equations (4) and (5). An estimate for the parameter vector $\boldsymbol{\theta}$ (which contains the thinning parameters $\alpha_1,\ldots,\alpha_p$ plus all parameters related to the innovations) is computed by numerically maximizing (the logarithm of) $L(\boldsymbol{\theta})$. For the non-Markovian INARMA(1,1) model, in contrast, an efficient likelihood computation is much more demanding.

3.1. *Efficient Implementation of ML Estimation*

According to (conditional) ML estimation, the estimate of the parameter vector $\boldsymbol{\theta}$ (which contains the two thinning parameters α_1, β_1 plus all parameters related to the innovations) is obtained as a maximizer of

$$L(\boldsymbol{\theta}) \;=\; P(Y_T=y_T,\ldots,Y_2=y_2 \mid Y_1=y_1).$$

For an efficient recursive computation of $L(\boldsymbol{\theta})$ (or the logarithm thereof, i.e., of $\ell(\boldsymbol{\theta})$), we adapt an approach known from ML estimation for Hidden-Markov models, see Zucchini et al. [19], Weiß [4]. Let us consider the probabilities

$$b_{kl}(t) \;=\; P(R_t=k, R_{t-1}=l;\; Y_t=y_t,\ldots,Y_2=y_2 \mid Y_1=y_1).$$

Then $L(\boldsymbol{\theta}) \;=\; \sum_{k,l} b_{kl}(T)$ holds. At a first glance, it seems that this sum has to be taken for $k,l=0,\ldots,\infty$. However, because of the model recursion $Y_t=\alpha_1\circ Y_{t-1}+\beta_1\circ R_{t-1}+R_t$, it is clear that $R_t \leq Y_t$ for all t. So if $M=\max\{y_T,\ldots,y_1\}$, then the $b_{kl}(t)$ only have to be computed for $k,l=0,\ldots,M$. Furthermore, $L(\boldsymbol{\theta}) \;=\; \sum_{k=0}^{y_T}\sum_{l=0}^{y_{T-1}} b_{kl}(T)$ holds.

The likelihood function is now computed recursively. We have

$$\begin{aligned}
b_{kl}(t+1) \;&=\; \sum_{i=0}^{y_{t-1}} P(R_{t+1}=k, R_t=l, R_{t-1}=i;\; Y_{t+1}=y_{t+1},\ldots \mid Y_1=y_1)\\
&= P(Y_{t+1}=y_{t+1} \mid R_{t+1}=k, R_t=l, Y_t=y_t)\, P(R_{t+1}=k)\\
&\quad \cdot \sum_{i=0}^{y_{t-1}} P(R_t=l, R_{t-1}=i;\; Y_t=y_t,\ldots \mid Y_1=y_1)\\
&= P(\alpha_1 \circ y_t + \beta_1 \circ l + k = y_{t+1})\, P(R_{t+1}=k) \sum_{i=0}^{y_{t-1}} b_{li}(t)\\
&= P(R_{t+1}=k) \left(\sum_{i=0}^{y_{t-1}} b_{li}(t)\right)\\
&\quad \cdot \sum_{y=0}^{y_{t+1}-k} \binom{y_t}{y} \alpha_1^y (1-\alpha_1)^{y_t-y} \binom{l}{y_{t+1}-k-y} \beta_1^{y_{t+1}-k-y} (1-\beta_1)^{l-y_{t+1}+k+y}.
\end{aligned}$$

For initialization, we approximate

$$\begin{aligned}
b_{kl}(2) \;&= P(R_2=k, R_1=l;\; Y_2=y_2 \mid Y_1=y_1)\\
&= P(Y_2=y_2 \mid R_2=k, R_1=l, Y_1=y_1)\, P(R_2=k)\, P(R_1=l \mid Y_1=y_1)\\
&\approx P(R_2=k)\, P(R_1=l \mid R_1 \le y_1)\\
&\quad \cdot \sum_{y=0}^{y_2-k} \binom{y_1}{y} \alpha_1^y (1-\alpha_1)^{y_1-y} \binom{l}{y_2-k-y} \beta_1^{y_2-k-y} (1-\beta_1)^{l-y_2+k+y}.
\end{aligned}$$

From these derivations, an algorithmic scheme for log-likelihood computation can be derived. First note that the update step $t \mapsto t+1$ only requires the sums $\sum_{i=0}^{y_{t-1}} b_{li}(t)$ but not the individual $b_{li}(t)$. Hence, it suffices to define the $(M+1)$-dimensional vectors $\boldsymbol{a}_t$ with entries $a_{t,l} = \sum_{i=0}^{y_{t-1}} b_{li}(t)$, which may have non-zero entries only for $l = 0,\ldots,y_t$. Next, the initialization step for $t=2$ can be included in the update step $t \mapsto t+1$ by defining $a_{1,l} = P(R_1 = l \mid R_1 \le y_1)$. Finally, computations can be simplified by defining the following matrices:

$$\mathbf{D} \;=\; \operatorname{diag}\big(P(R=0),\ldots,P(R=M)\big)$$

and $\mathbf{Q}_t \;=\; \big(q_{t,kl}\big)_{k,l=0,\ldots,M}$ for $t = 2,\ldots,T$ with

$$q_{t,kl} \;=\; \sum_{y=0}^{y_t-k} \binom{y_{t-1}}{y} \alpha_1^y (1-\alpha_1)^{y_{t-1}-y} \binom{l}{y_t-k-y} \beta_1^{y_t-k-y} (1-\beta_1)^{l-y_t+k+y}.$$

Then our algorithm becomes

$$\boldsymbol{a}_1, \quad \boldsymbol{a}_t \;=\; \mathbf{D}\,\mathbf{Q}_t\,\boldsymbol{a}_{t-1} \text{ for } t=2,\ldots,T, \quad L(\boldsymbol{\theta}) \;=\; \textstyle\sum_{k=0}^{y_T} a_{T,k} \;=\; \mathbf{1}^\top \boldsymbol{a}_T.$$

Finally, to make computations numerically more stable, also see Weiß [4], Zucchini et al. [19], we define

$$\begin{aligned}
&w_1 \;=\; \mathbf{1}^\top \boldsymbol{a}_1, \quad \boldsymbol{\phi}_1 \;=\; \boldsymbol{a}_1 / w_1;\\
&\boldsymbol{u}_t \;:=\; \mathbf{D}\,\mathbf{Q}_t\,\boldsymbol{\phi}_{t-1}, \qquad \tfrac{w_t}{w_{t-1}} \;=\; \mathbf{1}^\top \boldsymbol{u}_t, \quad \boldsymbol{\phi}_t \;=\; \boldsymbol{u}_t \,/\, \tfrac{w_t}{w_{t-1}} \text{ for } t=2,\ldots,T.
\end{aligned}$$

The log-likelihood function is obtained as

$$\ell(\boldsymbol{\theta}) \;=\; \ln w_T \;=\; \ln w_1 + \textstyle\sum_{t=2}^{T} \ln \frac{w_t}{w_{t-1}}.$$

The log-likelihood function $\ell(\boldsymbol{\theta})$ is numerically maximized by using a standard optimization routine.

Remark 1. *It should be noted that the same recursive scheme could also be used for implementing ML estimation for the INMA(1) model (i.e., where the AR-part is missing). Up to now, moment estimation or conditional least squares estimation are used for this model, see [5,7], and also the MCMC framework might be applied for this purpose, see [9,10,12]. If one wants to do ML estimation instead, the above scheme can be used together with one modification: the computation of $b_{kl}(t)$ and thus $\mathbf{Q}_t$ has to be simplified by setting $\alpha_1 = 0$. This implies that*

$$\begin{aligned} q_{t,kl} &= P(Y_t = y_t \mid R_t = k, R_{t-1} = l, Y_{t-1} = y_{t-1}) \\ &= P(\beta_1 \circ l + k = y_t) \;=\; \binom{l}{y_t - k} \beta_1^{y_t-k} (1-\beta_1)^{l-y_t+k}. \end{aligned}$$

The presented scheme is also easily modified to capture seasonality or trend. If these are incorporated by time-dependent thinning parameters or innovations parameters, respectively, as suggested by Freeland and McCabe [20] (p. 704), then one just has to modify $\mathbf{Q}_t$ or $\mathbf{D}_t = \operatorname{diag}\big(P(R_t = 0), \ldots, P(R_t = M)\big)$ (this diagonal matrix would then also depend on time t) accordingly.

Furthermore, it can also be adapted to fit higher-order INARMA models. If the AR-order p is > 1, we have to compute

$$\begin{aligned} q_{t,kl} &= P(Y_t = y_t \mid R_t = k, R_{t-1} = l, Y_{t-1} = y_{t-1}, \ldots, Y_{t-p} = y_{t-p}) \\ &= P(\alpha_1 \circ y_{t-1} + \ldots + \alpha_p \circ y_{t-p} + \beta_1 \circ l = y_t - k) \end{aligned}$$

as a convolution of $p+1$ binomial distributions. An MA-order $q > 1$ is more cumbersome as it goes along with an increase of dimensionality. We then have to define

$$b_{k_0 k_1 \ldots k_q}(t) \;=\; P(R_t = k_0, R_{t-1} = k_1, \ldots, R_{t-q} = k_q;\; Y_t = y_t, \ldots, Y_2 = y_2 \mid Y_1 = y_1)$$

with $q+1$ subscripts $k_0, k_1, \ldots, k_q = 0, \ldots, M$.

3.2. Performance of ML Estimation

We simulated INARMA processes of orders (1,0), (2,0) and (1,1), and with the innovations being either Poisson (Poi) or negative binomially (NB) distributed. In the latter case, the dispersion factor ν was chosen equal to 1.5. For the marginal means, we considered $\mu \in \{3,6\}$, and the dependence parameters were always chosen such that $\rho(1) \in \{0.35, 0.70\}$. We set $\alpha_2 = 0.25$ for the INAR(2) models, and $\beta_1 = 0.25$ for the INARMA(1,1) models. The considered sample sizes are $T \in \{100, 250, 500, 1000\}$. For each scenario, we simulated 1000 replications, and the model parameters were estimated with the ML approach described before (always choosing the appropriate model type).

Boxplots of the simulation results (and also tables with the means of the obtained estimates) are summarized in Appendix B.1. As can be seen from the boxplots, the respective ML estimates are certainly less biased with increasing T and also show decreasing dispersion for any of the considered models. However, the final sample properties differ a lot between different types of models and different types of parameters. One general observation is that the additional dispersion introduced by NB innovations deteriorates the estimation performance. We do not only observe increased bias for the innovations' mean λ and the dependence parameters $\alpha_1, \alpha_2, \beta_1$, also the estimation of ν itself suffers from rather large dispersion. Furthermore, the distribution of $\hat{\nu}$ is positively skewed and sometimes causes a strong overestimation of the true dispersion level ν. In addition, for the innovations' mean λ, we observe positive bias and skewness for small sample sizes, whereas the AR(1) parameter α_1 is negatively biased. The same holds for the AR(2) parameter α_2 (but with stronger bias than for α_1), whereas β_1 is positively biased for small T. Compared across models, we observe much more dispersion for the estimates of the dependence parameters in the INARMA(1,1) case than in the purely autoregressive cases. It is also worth noting that the dispersion of $\hat{\alpha}_1$ for INARMA(1,1) processes decreases with increasing mean μ.

4. Model Selection for INARMA Processes

In the simulations of Section 3.2, we always fitted the true model type to the given count time series. In practice, however, the true model behind the DGP is not known and therefore has to be identified based on the available data. A widely used approach is to apply information criteria (IC) for this purpose, and especially Akaike's and the Bayesian IC (AIC and BIC, respectively) are routinely used for this purpose see [21,22]. These criteria are computed together with the ML estimation of each candidate model, and that model is selected as the final one which minimizes the value of AIC or BIC, respectively. More details on these and further ICs can be found in Neath and Cavanaugh [21], Cavanaugh and Neath [22], Weiß and Feld [23]. In Weiß and Feld [23], the performance of these criteria was analyzed for count time series mainly generated by regression-type DGPs. They confirmed the consistency of the BIC in their study, but the actual rates of correct model identifications for smaller sample sizes T were often best for the AIC. Another related study is that of Zhu et al. [24]. They used AIC and BIC for selecting the order of the components of their mixture autoregressive Poisson regression model, but found out that these ICs "do not give a very satisfactory result" in this context.

Since model selection across INARMA models was not considered in Weiß and Feld [23], we also analyzed the AIC's and BIC's performance in our simulation study. For this purpose, we used all of the six considered models as possible candidate models for any of the simulated count time series. The obtained numbers of model selections (out of 1000 replications) are tabulated in Appendix B.2. Some of the conclusions found by Weiß and Feld [23] are confirmed also here. The BIC's ability for identifying the correct model always improves with increasing T, whereas the AIC for the smallest model, i.e., for the INAR(1) model in our comparison, stabilizes at a rate below 80%. On the other hand, it does most often better than the BIC for smaller sample sizes such as $T \leq 250$ (with exceptions only for INAR(1) DGPs). Besides looking at the correct identifications, it is also interesting to study the possible mis-identifications. It becomes clear that the purely autoregressive INAR models are mainly confused among themselves (by either choosing the wrong model order or the wrong distribution family), but we rarely observe an erroneous identification as INARMA(1,1). For an INARMA(1,1) DGP, in contrast, there is a very large risk of being mis-identified as an INAR(1) process. For $T = 100$ and also for larger T if $\rho(1) = 0.35$, such a mis-identification is even the most frequent decision, especially if using the BIC. This clearly differs from the INAR case, where the model order of the INAR(2) model is correctly identified in the majority of cases for any scenario. So not only the parameter estimation of an adequately chosen INARMA(1,1) model requires rather large sample sizes, these are also required for being able to correctly identify an INARMA(1,1) model at all. This shows that model selection should not solely rely on an information criterion, but further diagnostics should be done (e.g., a comparison of properties of the fitted models with the corresponding sample properties, see Section 6 below). At this point, it deems appropriate to recall the "two-units rule" for interpreting AIC and BIC, see Tables 1 in [21,22]. It says that if some candidate model's AIC (BIC) differs from the smallest AIC (BIC) by a difference ≤ 2, it should also be considered as a "viable candidate" [22] (p. 6).

5. On Properties of (Mis-)Fitted Models

The previous Section 4 showed that there is a considerable risk of choosing the wrong model for the actual count DGP. Such a (possibly falsely) chosen model is then used for interpreting the data, for forecasting future values, or for setting up a control chart for progressive process monitoring, see Weiß [4]. Hence, it is important to ask for the consequences of such a possible misfit of the DGP's true model. To answer this question (to some extent), we studied important stochastic properties of the fitted models and compared them to the respective properties of the true DGP.

More precisely, we focused on the marginal mean μ and dispersion ratio σ^2/μ as well as on the ACF values $\rho(1), \rho(2), \rho(3)$, see Section 2 for the required formulae. Appendix B.3 provides tables of the means of these properties.

The consequences of overfitting a Poi-INAR(1) DGP are quite moderate, a slight effect on $\rho(2)$ has to be noted. Things change if the INAR(1) process has NB innovations. If then falsely fitting a Poi model, we are not only unable to capture the apparent overdispersion $\sigma^2/\mu > 1$, also the ACF values are clearly undervalued. If erroneously fitting an INAR(1) or INARMA(1,1) model to a Poi-INAR(2) DGP, we severely underrate the actual ACF values, and this becomes even worse in the presence of overdispersion (i.e., NB-INAR(2) DGP). While analogous deviations are also observed for INARMA(1,1) DGPs, they appear to be less pronounced than in the INAR(2) case. Here, we usually underrate $\rho(1)$ but overrate $\rho(2), \rho(3)$. For the NB-INARMA(1,1) DGP, it is even worse to falsely choose a Poi-INARMA(1,1) model than an NB-INAR(1) or NB-INAR(2) model. So altogether, it appears that an inappropriate choice of the dispersion structure has the worst effect on the quality of the fitted model.

6. Real-Data Examples

Let us discuss two real-data examples, which both are time series $y_1, \ldots, y_T$ of monthly counts. The first time series consists of monthly counts of crime offense reports regarding burglaries from the 43th police car beat in Pittsburgh (1990–2001, so $T = 144$). It is available at the Forecasting Principles website, http://www.forecastingprinciples.com/index.php/crimedata. The second time series consists of monthly counts of claims caused by burn related injuries in the heavy manufacturing industry (1987–1994, so $T = 96$), see Example 2.5.1 in Freeland [25] for further details. Plots of both time series are provided by Figure 1.

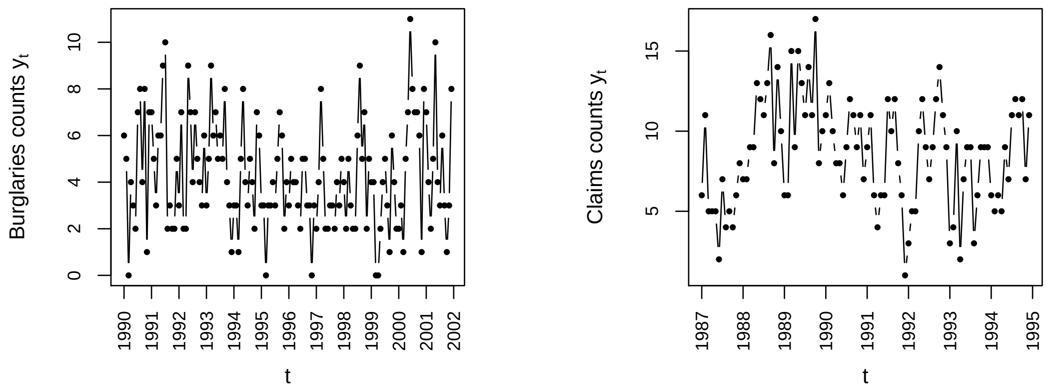


Figure 1. Time series of monthly counts of burglaries (**left**) and claims (**right**).

Both time series exhibit a quickly decaying ACF and a moderate extent of empirical overdispersion (about 30%), see the first row in Table 1, such that INAR(1), INAR(2) and INARMA(1,1) models with either Poisson or NB innovations are reasonable candidate models. These were fitted to both time series using ML estimation, as outlined in Section 3. Since both time series are rather short, we use the AIC for (initial) model selection; our analyses in Section 4 showed that the BIC most often has very low rates of correct model identification in such a case. The obtained AIC values as well as relevant stochastic properties of the fitted models (in analogy to Section 5) are summarized in the upper part of Table 1.

For the burglaries counts, the AIC selects the NB-INAR(1) model, but the AIC of the NB-INARMA(1,1) model is also quite low. Actually, also the corresponding Poi-counterparts have AICs that deviate from the lowest AIC by not more than two units, only the INAR(2) models violate this rule. If we now compare the tabulated properties of the fitted models with their corresponding sample counterparts (the latter are printed in italic font), the NB-INARMA(1,1) model shows the least deviations. The NB-INAR(1) model, in contrast, does not only show a smaller mean μ and dispersion ratio σ^2/μ than observed in the sample, also the ACF values show more deviations (and the Poi-models do even worse). So actually, the NB-INARMA(1,1) model provides the best fit to the burglaries counts data. As a further check for model adequacy, we computed an acceptance envelope in the spirit of Tsay [26] for the properties considered in Table 1. For this purpose, a parametric bootstrap with 10,000 replications for the fitted NB-INARMA(1,1) model was done. From the sample properties obtained for each simulation run, quartiles and standard errors for the sample statistics have been computed. It can be seen that the values for $\hat{\mu}$, $\hat{\sigma}^2/\hat{\mu}$, $\hat{\rho}(1)$, $\hat{\rho}(2)$, $\hat{\rho}(3)$ as computed from the time series (first line of Table 1) are always within the respective lower and upper quartile in the lower part of Table 1, so they do not contradict the fitted NB-INARMA(1,1) model.

Table 1. Upper part: Sample properties (italic font) and properties of maximum likelihood (ML)-fitted models for monthly counts of burglaries (left) and claims (right). Lower part: Results from parametric bootstrap experiment with 10,000 replications.

	Burglaries Counts:						**Claims Counts:**					
	AIC	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	**AIC**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
Sample	–	*4.319*	*1.271*	*0.255*	*0.014*	*0.040*	–	*8.604*	*1.320*	*0.452*	*0.362*	*0.213*
Poi-INAR(1)	643.7	4.311	1.000	0.210	0.044	0.009	490.5	8.667	1.000	0.396	0.157	0.062
NB-INAR(1)	641.9	4.312	1.264	0.238	0.057	0.013	490.4	8.670	1.236	0.426	0.181	0.077
Poi-INAR(2)	646.7	4.309	1.000	0.208	0.043	0.009	487.2	8.610	1.063	0.419	0.318	0.183
NB-INAR(2)	644.6	4.309	1.273	0.236	0.056	0.013	488.1	8.610	1.250	0.444	0.339	0.202
Poi-INARMA(1,1)	643.9	4.316	1.030	0.248	0.024	0.002	492.5	8.666	1.007	0.398	0.155	0.061
NB-INARMA(1,1)	642.7	4.319	1.273	0.266	0.019	0.001	492.4	8.682	1.264	0.418	0.175	0.073
Bootstrap for	**Fitted NB-INARMA(1,1):**						**Fitted NB-INAR(2):**					
lower quartile		4.153	1.149	0.200	−0.052	−0.070		8.188	1.048	0.331	0.218	0.069
median		4.313	1.256	0.254	0.006	−0.009		8.594	1.190	0.407	0.295	0.151
upper quartile		4.479	1.371	0.306	0.065	0.049		9.031	1.357	0.480	0.368	0.233
standard error		0.247	0.167	0.077	0.086	0.086		0.634	0.235	0.111	0.112	0.121

But why did the AIC select the NB-INAR(1) model? Here, it is important to recall the discussion in Section 4. If the true DGP would be NB-INARMA(1,1), then the AIC has a very low rate of correct model identification for small T. If we consider the scenario $\mu = 3$, $\rho(1) = 0.35$ and $T = 100$ in Appendix B.2, which is reasonably close to our data example, then the AIC falsely selects an NB-INAR(1) model in the majority of cases. So the AIC (and even more the BIC) is not a reliable tool for model selection for such small T. It gives rough indication in the sense that Poisson models do not work well, or that an INAR(2) model does not improve over an INAR(1) model. However, the decision for the final model should also consider further aspects like the stochastic properties of the fitted models.

An analogous conclusion applies to the claims counts data. This time, the AIC prefers the Poi-INAR(2) model against the remaining candidate models, but the NB-INAR(2) model also has an AIC satisfying the two-units rule. However, both the dispersion ratio and the ACF values are much smaller for the fitted Poi-INAR(2) model than for the sample itself, whereas the fitted NB-INAR(2) model gives a clearly better agreement (Table 1 also provides some results from a parametric bootstrap for the fitted NB-INAR(2) model with 10,000 replications, which do not contradict the adequacy of this model). Hence, it appears that the AIC was misleading also for this data example. Looking into Appendix B.2, DGP NB-INAR(2) with $\mu = 6$, $\rho(1) = 0.35$ and $T = 100$, we see a strong tendency of the AIC for falsely selecting a Poi-INAR(2) model (and this becomes even worse with increasing $\rho(1)$). So we conclude again that for short count time series, the AIC (and especially the BIC) should be used with caution. It gives a rough orientation regarding the correct model type (the choice of an INAR(2) model appears to be justified), but the final model selection should also use further criteria like a comparison of model properties.

7. Conclusions

We derived an efficient scheme for ML estimation of INARMA(1,1) models, and we also discussed possible extensions to higher-order models, or to INARMA models with seasonality or trend. Then we compared the INARMA(1,1) model to the INAR(1) and INAR(2) model (these three models constitute a reasonable set of candidate models for many applications) in several respects. The performance of ML estimation is generally rather good, but a small-sample bias might be observed especially in the presence of overdispersion. The BIC is consistent for model selection across these INARMA models, but it shows a worse performance than the AIC for small sample sizes (such as $T \leq 250$). In addition, the AIC has to be treated with caution for very short time series (such as $T = 100$): It gives a rough orientation but should always be complemented by further selection criteria. The consequences of fitting the wrong model to the actual DGP might be particularly severe in the presence of overdispersion, where a misfit of the dispersion behavior also causes a misfit of the ACF values.

Author Contributions: Conceptualization (C.W., N.K.); Methodology (C.W.); Software (C.W., M.F., N.K.); Writing—original draft preparation (C.W., M.F., Y.S.); Writing—review and editing (C.W.).

Funding: This research received no external funding.

Acknowledgments: The authors thank the three referees for their useful comments on an earlier draft of this article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Derivations

Using Equations (2) and (9) and Lemma 1, it follows that

$$\begin{aligned} E(Y_t) &= E(\alpha_1 \circ Y_{t-1} + \beta_1 \circ R_{t-1} + R_t) \\ &= \alpha_1 E(Y_{t-1}) + \beta_1 E(R_{t-1}) + E(R_t) \\ &= \alpha_1 E(Y_t) + \beta_1 E(R_t) + E(R_t), \end{aligned}$$

so

$$(1-\alpha_1)E(Y_t) = (1+\beta_1)E(R_t) = (1+\beta_1)\lambda.$$

Hence, $E(Y_t) = \frac{(1+\beta_1)\lambda}{(1-\alpha_1)}$ and the proof of (10) is complete.
Equation (11) is derived analogously, by solving

$$\begin{aligned} \mathrm{Var}(Y_t) &= \mathrm{Var}(\alpha_1 \circ Y_{t-1} + \beta_1 \circ R_{t-1} + R_t) \\ &= \mathrm{Var}(\alpha_1 \circ Y_{t-1}) + \mathrm{Var}(\beta_1 \circ R_{t-1}) + 2\mathrm{Cov}(\alpha_1 \circ Y_{t-1}, \beta_1 \circ R_{t-1}) + \mathrm{Var}(R_t) \\ &= \alpha_1(1-\alpha_1)E(Y_{t-1}) + \alpha_1^2\mathrm{Var}(Y_{t-1}) + \beta_1(1-\beta_1)E(R_{t-1}) + \beta_1^2\mathrm{Var}(R_{t-1}) \\ &\quad + 2\alpha_1\beta_1\mathrm{Var}(R_{t-1}) + \mathrm{Var}(R_t) \\ &= \alpha_1(1-\alpha_1)E(Y_t) + \alpha_1^2\mathrm{Var}(Y_t) + \beta_1(1-\beta_1)E(R_t) + \beta_1^2\mathrm{Var}(R_t) \\ &\quad + 2\alpha_1\beta_1\mathrm{Var}(R_t) + \mathrm{Var}(R_t), \end{aligned}$$

so

$$(1-\alpha_1^2)\mathrm{Var}(Y_t) = \alpha_1(1-\alpha_1)\,\frac{(1+\beta_1)\lambda}{(1-\alpha_1)} + \beta_1(1-\beta_1)\lambda + \beta_1^2\nu\lambda + 2\alpha_1\beta_1\nu\lambda + \nu\lambda.$$

Finally,

$$\begin{aligned}
\mathrm{Cov}(Y_t, Y_{t+h}) &= \mathrm{Cov}[Y_t, (\alpha_1 \circ Y_{t+h-1} + \beta_1 \circ R_{t+h-1} + R_{t+h})] \\
&= \alpha_1 \,\mathrm{Cov}[Y_t, Y_{t+h-1}] \\
&= \alpha_1 \,\mathrm{Cov}[Y_t, (\alpha_1 \circ Y_{t+h-2} + \beta_1 \circ R_{t+h-2} + R_{t+h-1})] \\
&= \alpha_1^2 \,\mathrm{Cov}[Y_t, Y_{t+h-2}] \\
&\ \ \vdots \\
&= \alpha_1^{h-1} \,\mathrm{Cov}[Y_t, (\alpha_1 \circ Y_t + \beta_1 \circ R_t + R_{t+1})] \\
&= \alpha_1^h \,\mathrm{Var}(Y_t) + \alpha_1^{h-1} \,\beta_1 \,\mathrm{Var}(R_t) \\
&= \alpha_1^h \,\mathrm{Var}(Y_t) + \alpha_1^{h-1} \,\beta_1 \,\nu\lambda,
\end{aligned}$$

which proves Equation (12).

Appendix B. Results from Simulation Study

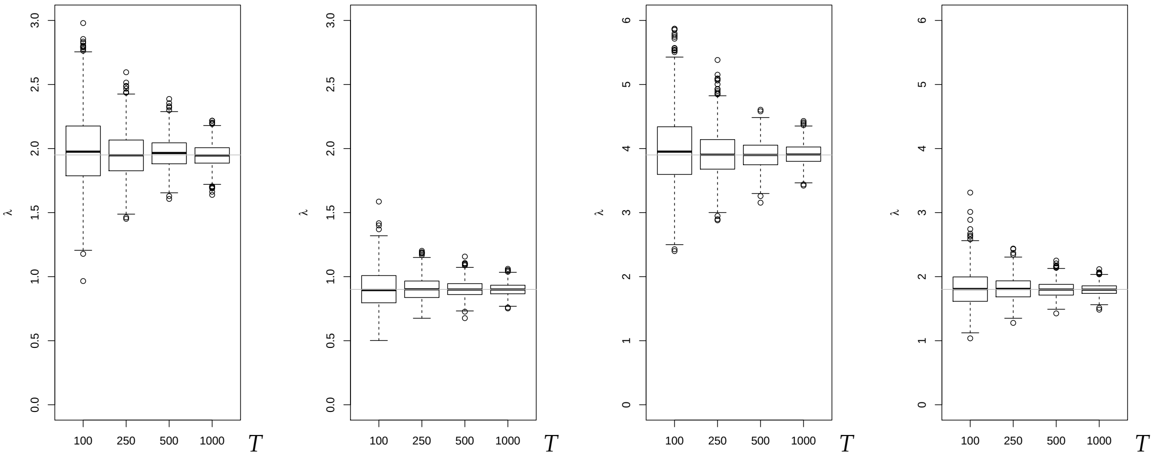
Appendix B.1. Parameter Estimation for Adequate Model

Boxplots and means of simulated ML estimates if the correct type of model is fitted to the simulated data (1000 replications).

Appendix B.1.1. ML-Estimates for DGP Poi-INAR(1)

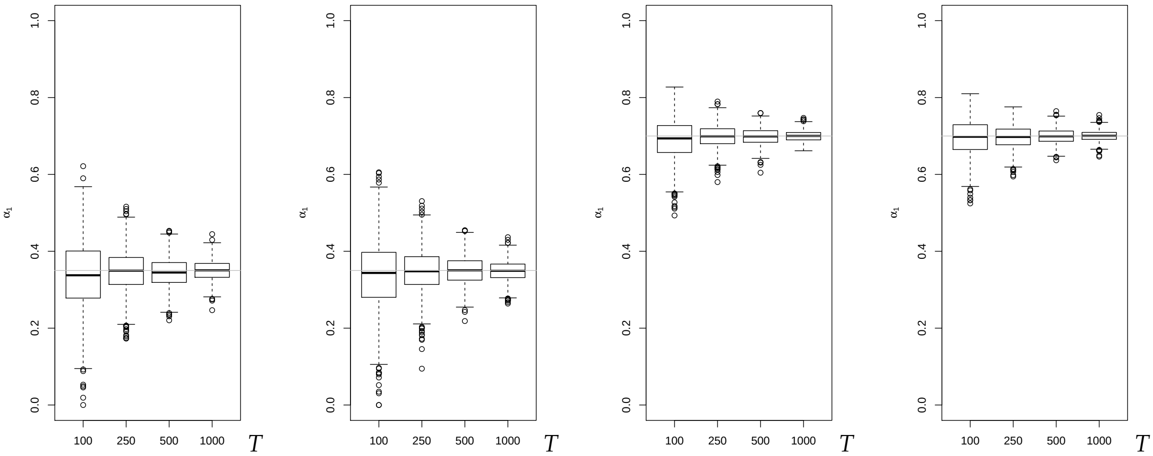
Estimation of λ:

$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$



Estimation of α_1:

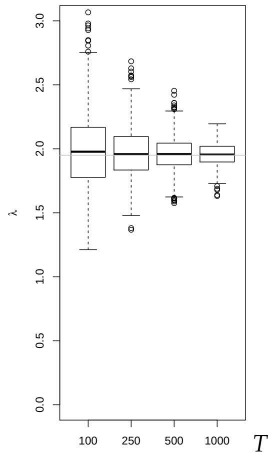
$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$



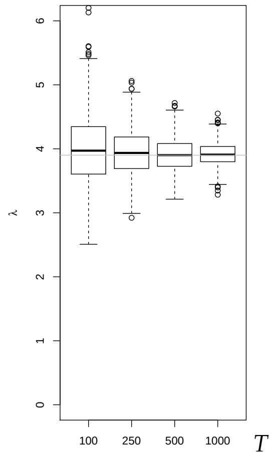
Appendix B.1.2. ML-Estimates for DGP NB-INAR(1)

Estimation of λ:

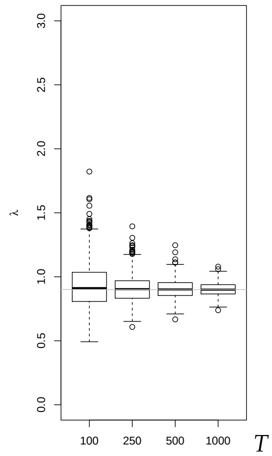
$\mu = 3, \rho(1) = 0.35$



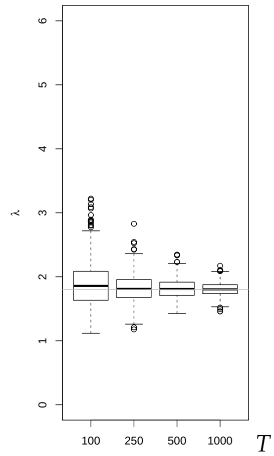
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

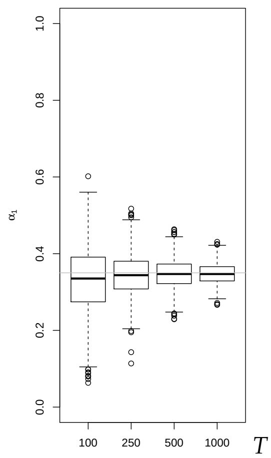


$\mu = 6, \rho(1) = 0.70$

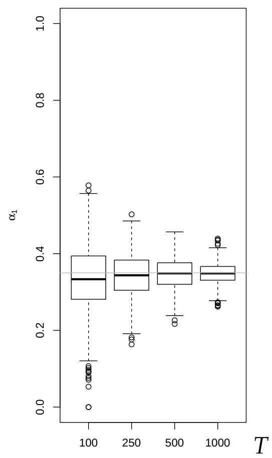


Estimation of α_1:

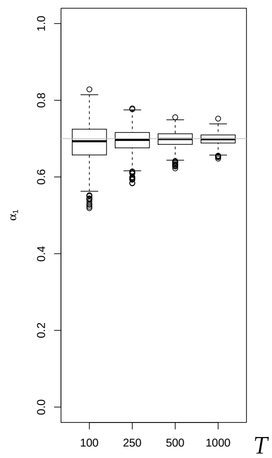
$\mu = 3, \rho(1) = 0.35$



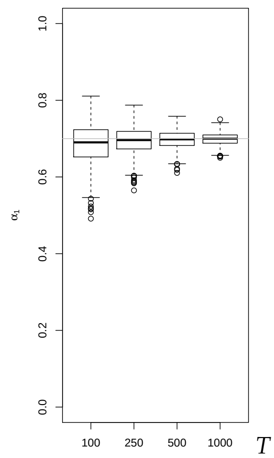
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

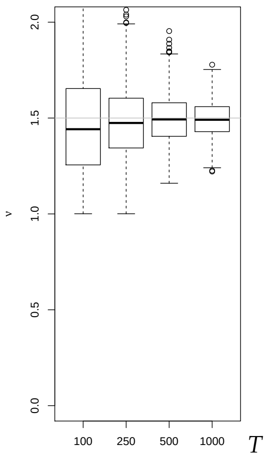


$\mu = 6, \rho(1) = 0.70$

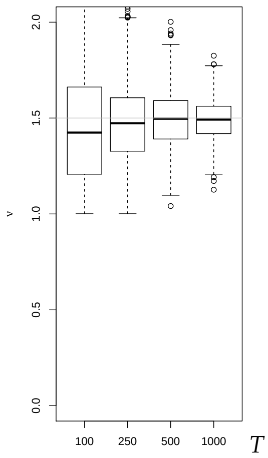


Estimation of ν:

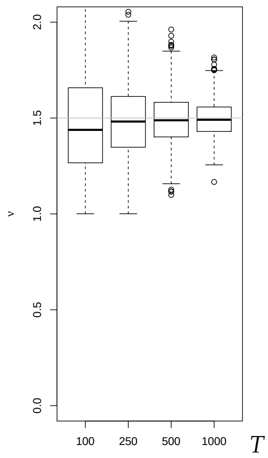
$\mu = 3, \rho(1) = 0.35$



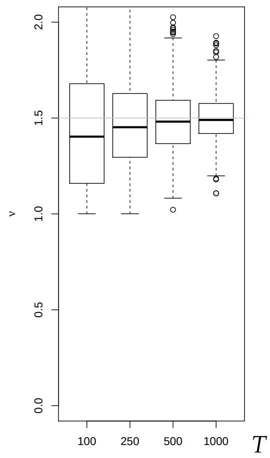
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$



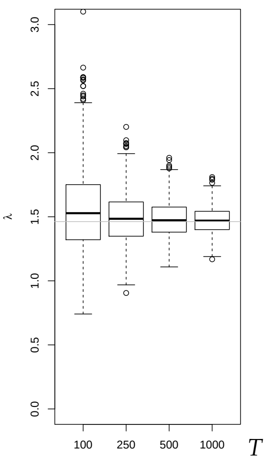
$\mu = 6, \rho(1) = 0.70$



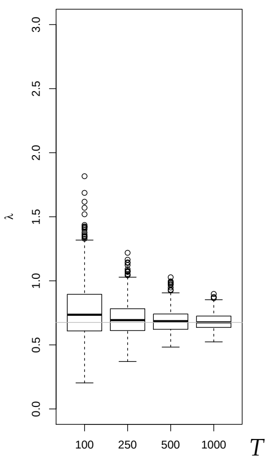
Appendix B.1.3. ML-Estimates for DGP Poi-INAR(2)

Estimation of λ:

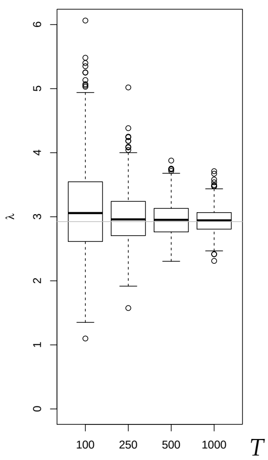
$\mu = 3, \rho(1) = 0.35$



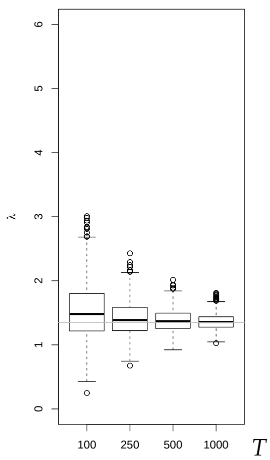
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

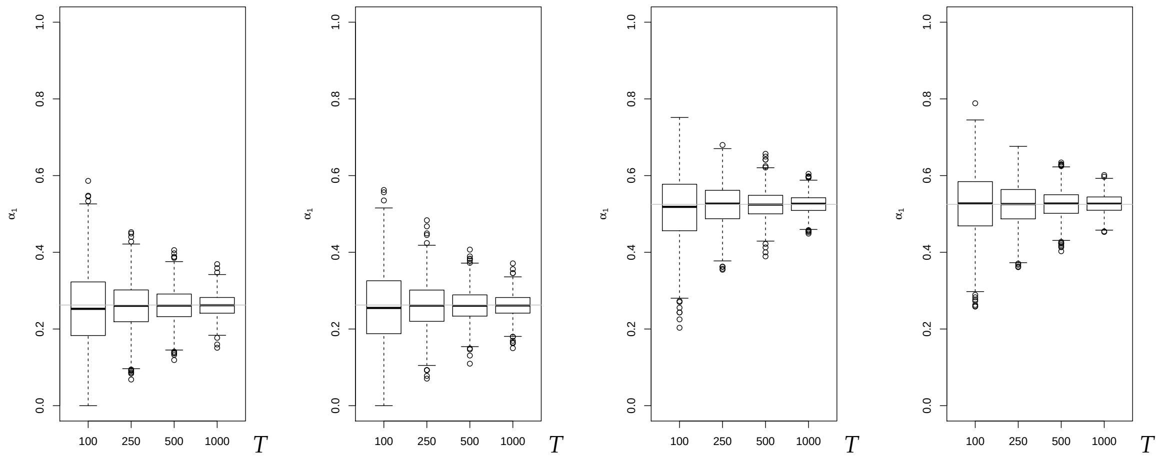


$\mu = 6, \rho(1) = 0.70$



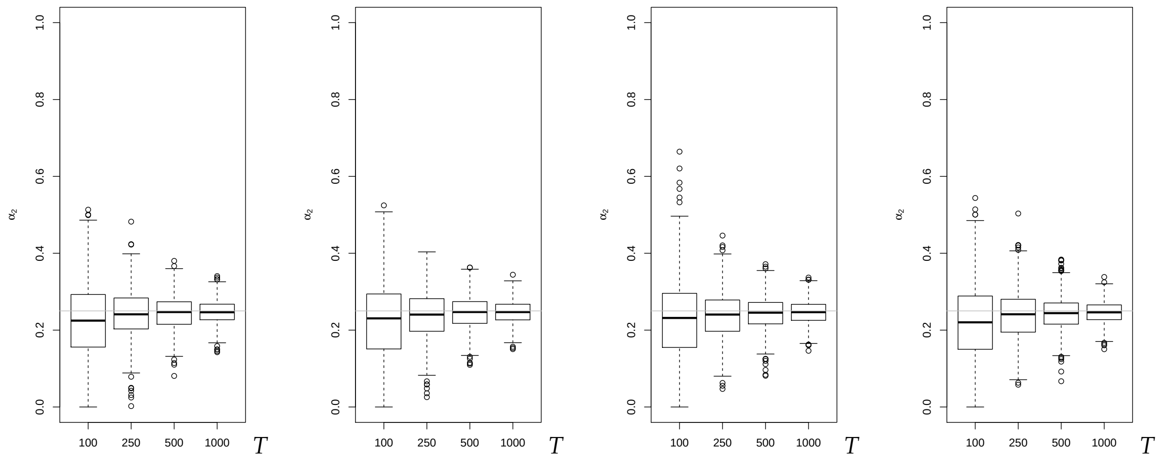
Estimation of α_1:

$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$



Estimation of α_2:

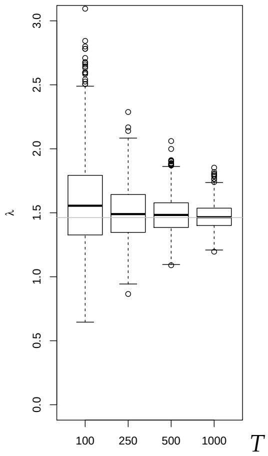
$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$



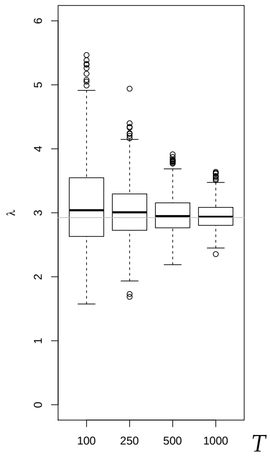
Appendix B.1.4. ML-Estimates for DGP NB-INAR(2)

Estimation of λ:

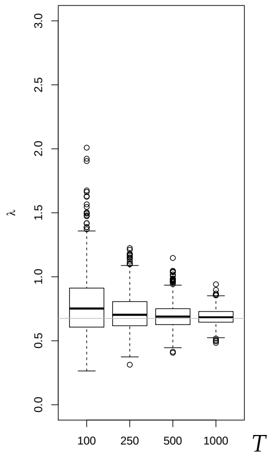
$\mu = 3, \rho(1) = 0.35$



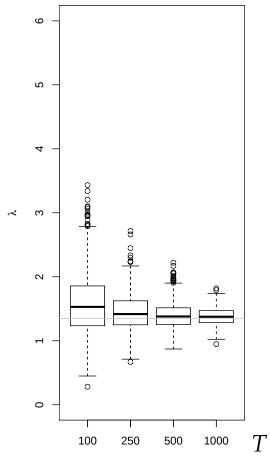
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

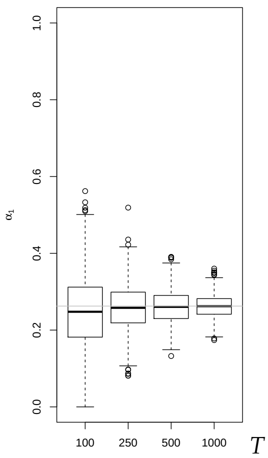


$\mu = 6, \rho(1) = 0.70$

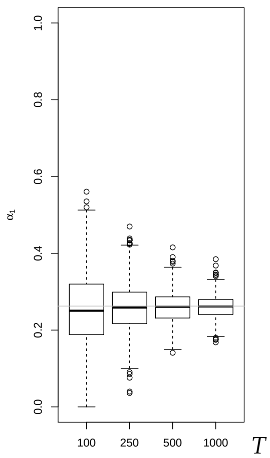


Estimation of α_1:

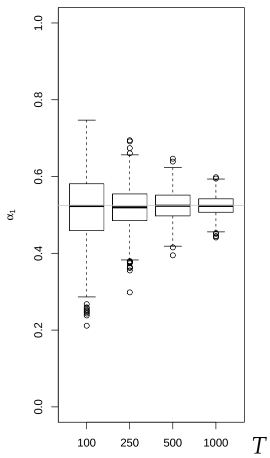
$\mu = 3, \rho(1) = 0.35$



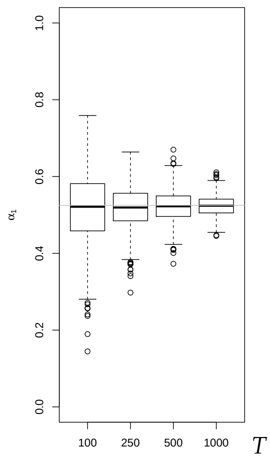
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

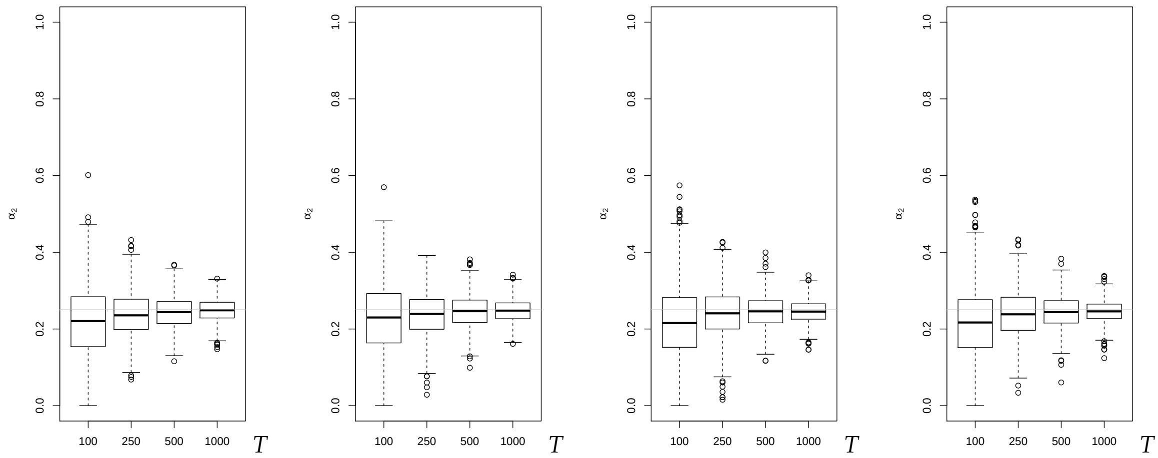


$\mu = 6, \rho(1) = 0.70$



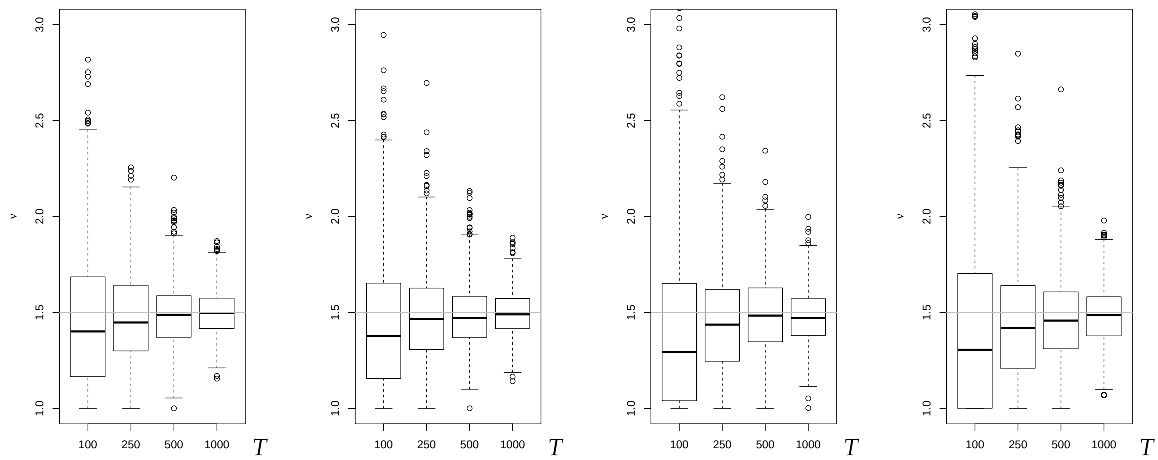
Estimation of α_2:

$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$



Estimation of ν:

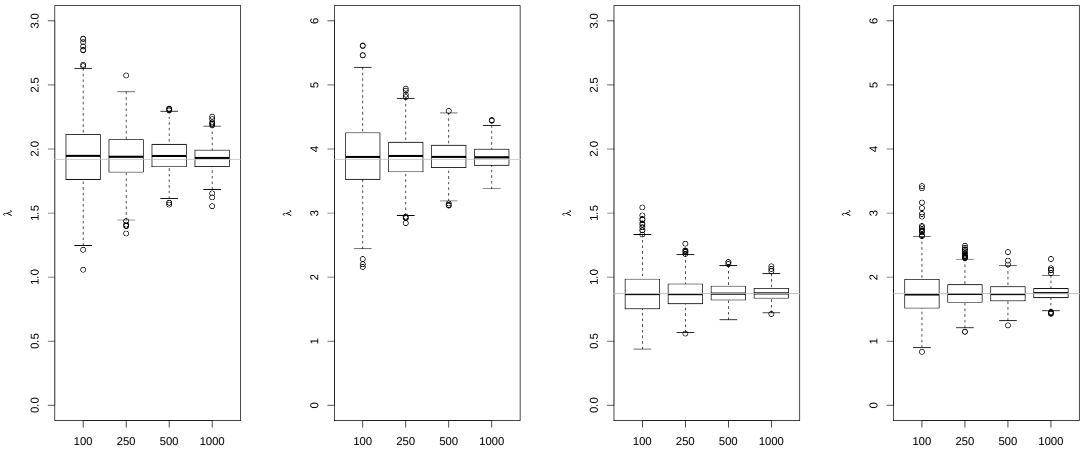
$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$



Appendix B.1.5. ML-Estimates for DGP Poi-INARMA(1,1)

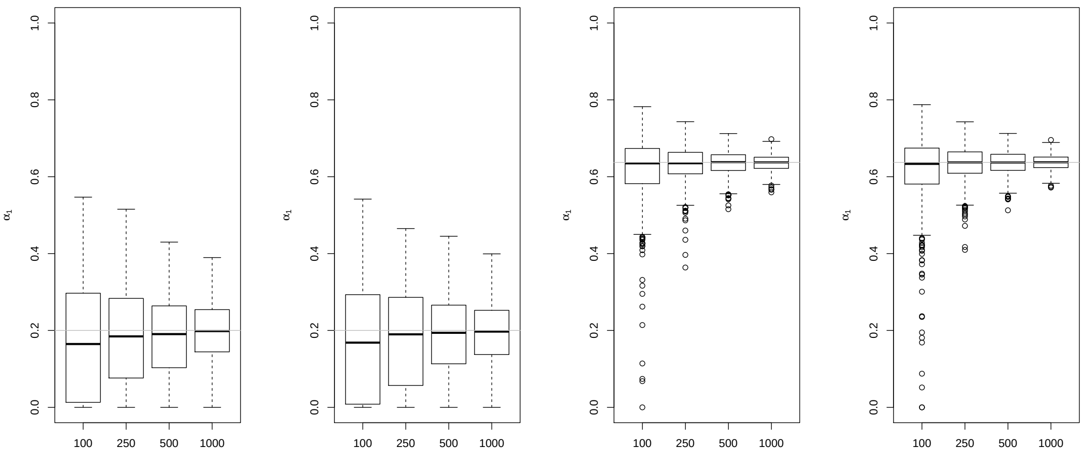
Estimation of λ:

$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$



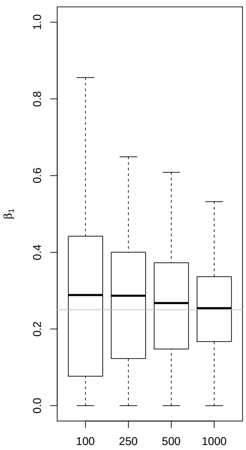
Estimation of α_1:

$\mu = 3, \rho(1) = 0.35$ $\mu = 3, \rho(1) = 0.70$ $\mu = 6, \rho(1) = 0.35$ $\mu = 6, \rho(1) = 0.70$

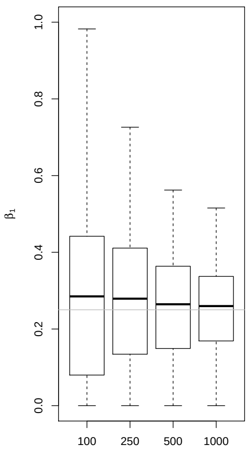


Estimation of β_1:

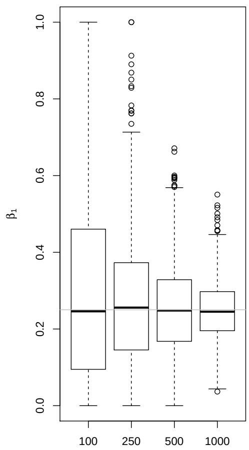
$\mu = 3, \rho(1) = 0.35$



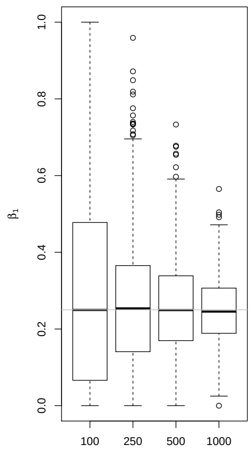
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$



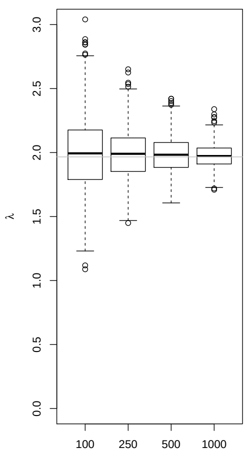
$\mu = 6, \rho(1) = 0.70$



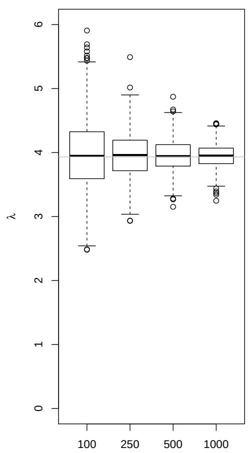
Appendix B.1.6. ML-Estimates for DGP NB-INARMA(1,1)

Estimation of λ:

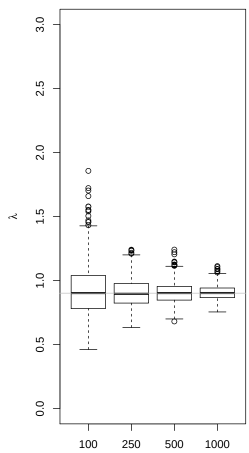
$\mu = 3, \rho(1) = 0.35$



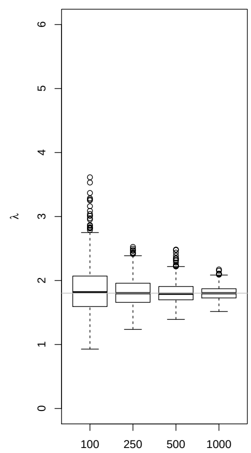
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

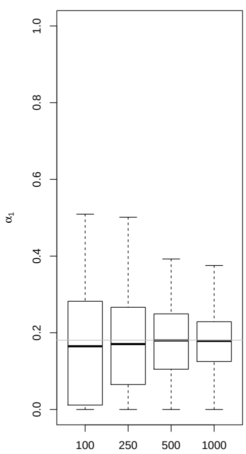


$\mu = 6, \rho(1) = 0.70$

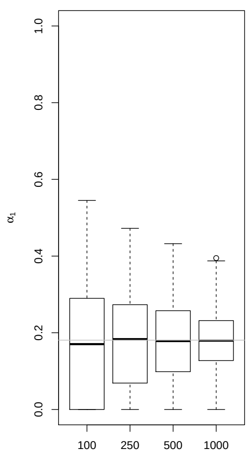


Estimation of α_1:

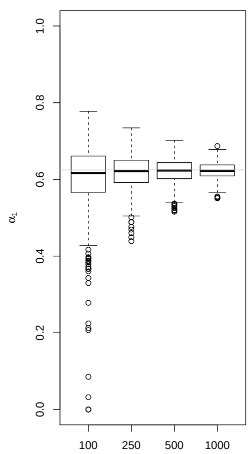
$\mu = 3, \rho(1) = 0.35$



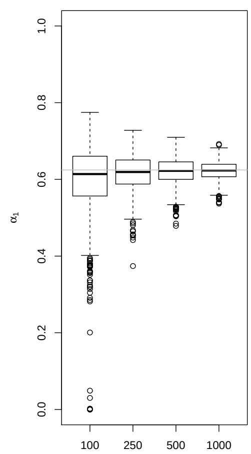
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

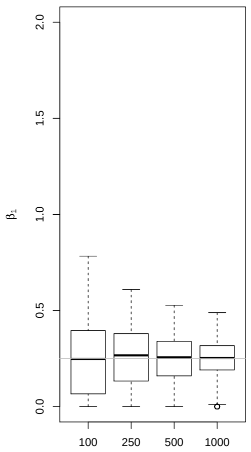


$\mu = 6, \rho(1) = 0.70$

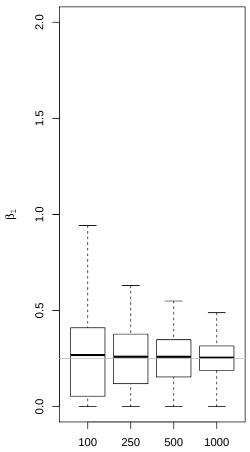


Estimation of β_1:

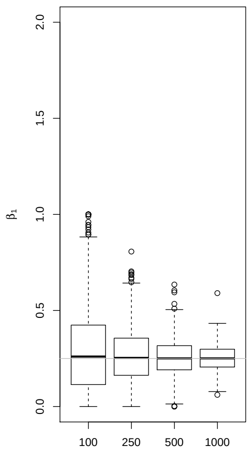
$\mu = 3, \rho(1) = 0.35$



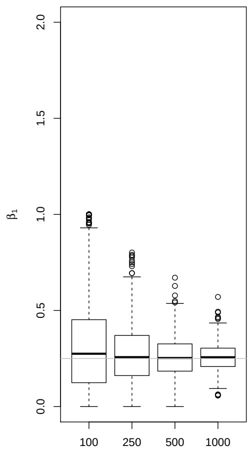
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$

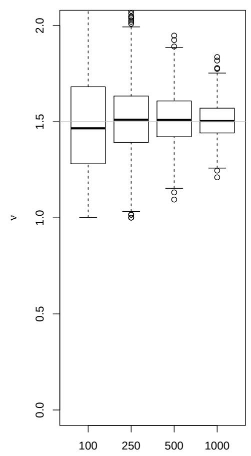


$\mu = 6, \rho(1) = 0.70$

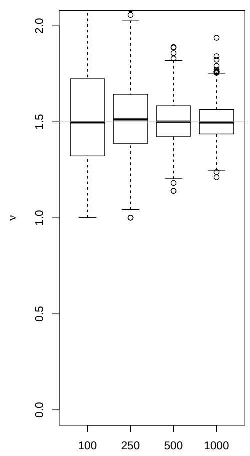


Estimation of ν:

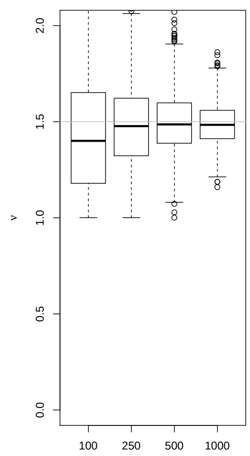
$\mu = 3, \rho(1) = 0.35$



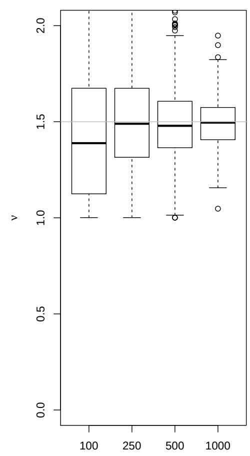
$\mu = 3, \rho(1) = 0.70$



$\mu = 6, \rho(1) = 0.35$



$\mu = 6, \rho(1) = 0.70$



Appendix B.1.7. Means of ML-Estimates

		$\mu = 3, \rho(1) = 0.35$				$\mu = 3, \rho(1) = 0.70$			
DGP	T	λ	α_1	α_2 **or** β_1	ν	λ	α_1	α_2 **or** β_1	ν
Poi-INAR(1)	100	1.982	0.338			0.908	0.691		
	250	1.952	0.348			0.903	0.699		
	500	1.964	0.345			0.903	0.698		
	1000	1.948	0.350			0.900	0.700		
NB-INAR(1)	100	1.989	0.332		1.470	0.929	0.689		1.466
$\nu = 1.5$	250	1.969	0.344		1.486	0.906	0.695		1.481
	500	1.960	0.347		1.494	0.905	0.698		1.498
	1000	1.956	0.347		1.495	0.902	0.699		1.492
Poi-INAR(2)	100	1.557	0.253	0.222		0.762	0.516	0.227	
$\alpha_2 = 0.25$	250	1.492	0.259	0.242		0.704	0.524	0.239	
	500	1.479	0.262	0.245		0.688	0.525	0.245	
	1000	1.472	0.262	0.247		0.680	0.526	0.246	
NB-INAR(2)	100	1.588	0.247	0.220	1.466	0.777	0.518	0.217	1.406
$\alpha_2 = 0.25$	250	1.504	0.258	0.238	1.473	0.720	0.519	0.239	1.456
$\nu = 1.5$	500	1.486	0.260	0.244	1.486	0.696	0.524	0.244	1.492
	1000	1.469	0.262	0.249	1.499	0.687	0.524	0.246	1.479
Poi-INARMA(1,1)	100	1.944	0.177	0.278		0.878	0.622	0.302	
$\beta_1 = 0.25$	250	1.944	0.180	0.265		0.872	0.633	0.272	
	500	1.947	0.183	0.258		0.876	0.636	0.252	
	1000	1.929	0.196	0.249		0.873	0.636	0.249	
NB-INARMA(1,1)	100	1.995	0.169	0.248	1.494	0.921	0.606	0.291	1.458
$\beta_1 = 0.25$	250	1.988	0.171	0.253	1.517	0.904	0.619	0.262	1.481
$\nu = 1.5$	500	1.985	0.176	0.247	1.516	0.906	0.621	0.255	1.499
	1000	1.976	0.175	0.252	1.507	0.905	0.623	0.251	1.487

		$\mu = 6, \rho(1) = 0.35$				$\mu = 6, \rho(1) = 0.70$			
DGP	T	λ	α_1	α_2 **or** β_1	ν	λ	α_1	α_2 **or** β_1	ν
Poi-INAR(1)	100	3.979	0.337			1.821	0.696		
	250	3.911	0.347			1.811	0.697		
	500	3.900	0.350			1.798	0.700		
	1000	3.909	0.349			1.800	0.700		
NB-INAR(1)	100	3.996	0.334		1.480	1.882	0.686		1.453
$\nu = 1.5$	250	3.945	0.344		1.485	1.826	0.694		1.466
	500	3.910	0.348		1.494	1.816	0.697		1.484
	1000	3.910	0.348		1.494	1.807	0.699		1.497
Poi-INAR(2)	100	3.114	0.255	0.224		1.524	0.525	0.220	
$\alpha_2 = 0.25$	250	2.988	0.261	0.239		1.413	0.524	0.239	
	500	2.955	0.261	0.246		1.381	0.526	0.243	
	1000	2.944	0.262	0.247		1.364	0.527	0.246	
NB-INAR(2)	100	3.118	0.252	0.229	1.438	1.587	0.516	0.216	1.442
$\alpha_2 = 0.25$	250	3.021	0.258	0.238	1.483	1.447	0.519	0.238	1.447
$\nu = 1.5$	500	2.966	0.259	0.246	1.485	1.397	0.522	0.244	1.471
	1000	2.948	0.261	0.247	1.497	1.379	0.523	0.246	1.482
Poi-INARMA(1,1)	100	3.896	0.173	0.281		1.754	0.622	0.310	
$\beta_1 = 0.25$	250	3.881	0.180	0.270		1.748	0.635	0.264	
	500	3.882	0.189	0.254		1.738	0.636	0.261	
	1000	3.871	0.193	0.252		1.751	0.637	0.249	
NB-INARMA(1,1)	100	3.955	0.175	0.257	1.534	1.853	0.598	0.318	1.447
$\beta_1 = 0.25$	250	3.957	0.176	0.248	1.523	1.813	0.617	0.271	1.491
$\nu = 1.5$	500	3.959	0.176	0.247	1.505	1.808	0.621	0.258	1.490
	1000	3.949	0.178	0.248	1.504	1.802	0.622	0.257	1.493

Appendix B.2. Model Identification

Numbers of selecting one of the candidate models Poi-INAR(1), Poi-INAR(2), Poi-INARMA(1,1), NB-INAR(1), NB-INAR(2), NB-INARMA(1,1) for a given type of data-generating process (DGP) out of 1000 replications, by using the information criteria AIC or BIC. Numbers of correct identifications highlighted in italic font.

				Model Identification by AIC						**Model Identification by BIC**					
DGP Poi-	μ	$\rho(1)$	T	**Poi-INAR(1)**	**NB-INAR(1)**	**Poi-INAR(2)**	**NB-INAR(2)**	**Poi-INARMA(1,1)**	**NB-INARMA(1,1)**	**Poi-INAR(1)**	**NB-INAR(1)**	**Poi-INAR(2)**	**NB-INAR(2)**	**Poi-INARMA(1,1)**	**NB-INARMA(1,1)**
INAR(1)	3	0.35	100	*733*	44	118	5	92	8	*945*	10	22	0	22	1
			250	*738*	45	123	5	87	2	*958*	11	19	0	12	0
			500	*758*	40	115	9	77	1	*970*	6	11	0	13	0
			1000	*716*	56	148	9	71	0	*987*	0	9	0	4	0
	3	0.70	100	*758*	44	125	7	66	0	*951*	10	22	1	16	0
			250	*783*	46	106	3	61	1	*967*	8	16	0	9	0
			500	*746*	56	131	6	60	1	*968*	3	19	1	9	0
			1000	*750*	52	124	7	65	2	*986*	1	12	0	1	0
	6	0.35	100	*717*	47	118	5	110	3	*933*	13	30	0	24	0
			250	*725*	35	122	6	109	3	*961*	4	18	0	17	0
			500	*765*	39	128	5	62	1	*977*	6	11	0	6	0
			1000	*730*	51	141	11	65	2	*982*	2	9	0	7	0
	6	0.70	100	*784*	35	110	1	67	3	*951*	10	25	0	14	0
			250	*725*	35	134	10	95	1	*961*	6	22	0	11	0
			500	*735*	53	133	5	74	0	*976*	7	7	0	10	0
			1000	*771*	39	118	4	68	0	*985*	5	7	0	3	0
INAR(2)	3	0.35	100	222	11	*727*	39	1	0	471	8	*515*	6	0	0
$\alpha_2 = 0.25$			250	12	0	*937*	51	0	0	79	0	*915*	6	0	0
			500	1	0	*947*	52	0	0	1	0	*996*	3	0	0
			1000	0	0	*935*	65	0	0	0	0	*994*	6	0	0
	3	0.70	100	169	9	*772*	44	5	1	350	7	*632*	10	1	0
			250	6	1	*927*	66	0	0	36	0	*958*	6	0	0
			500	0	0	*948*	52	0	0	2	0	*995*	3	0	0
			1000	0	0	*930*	70	0	0	0	0	*996*	4	0	0
	6	0.35	100	227	6	*732*	34	1	0	462	2	*531*	5	0	0
			250	16	0	*920*	64	0	0	92	0	*900*	8	0	0
			500	0	0	*928*	72	0	0	4	0	*991*	5	0	0
			1000	0	0	*947*	53	0	0	0	0	*997*	3	0	0
	6	0.70	100	183	18	*757*	40	1	1	378	7	*610*	4	1	0
			250	5	1	*949*	45	0	0	48	2	*945*	5	0	0
			500	0	0	*940*	60	0	0	2	0	*993*	5	0	0
			1000	0	0	*940*	60	0	0	0	0	*999*	1	0	0
INARMA(1,1)	3	0.35	100	589	61	68	8	*244*	30	890	23	15	1	*68*	3
$\beta_1 = 0.25$			250	425	61	66	6	*409*	33	841	16	14	0	*128*	1
			500	277	61	48	8	*554*	52	747	15	5	0	*231*	2
			1000	136	41	19	5	*738*	61	524	15	2	0	*456*	3
	3	0.70	100	498	45	75	8	*357*	17	778	19	24	0	*176*	3
			250	270	27	60	8	*607*	28	654	9	12	0	*323*	2
			500	114	13	26	3	*787*	57	456	4	4	0	*534*	2
			1000	10	3	6	1	*918*	62	144	5	1	0	*849*	1
	6	0.35	100	603	57	82	7	*222*	29	891	18	22	0	*68*	1
			250	437	90	60	8	*360*	45	840	25	13	0	*117*	5
			500	291	82	45	5	*531*	46	753	21	4	1	*219*	2
			1000	137	45	29	6	*710*	73	520	23	5	0	*451*	1
	6	0.70	100	514	43	77	5	*347*	14	811	13	19	0	*157*	0
			250	305	43	47	7	*565*	33	694	17	10	1	*278*	0
			500	119	29	25	5	*769*	53	472	12	9	0	*504*	3
			1000	19	4	7	4	*904*	62	190	5	2	0	*801*	2

				Model Identification by AIC						**Model Identification by BIC**					
DGP NB-	μ	$\rho(1)$	T	**Poi-INAR(1)**	**NB-INAR(1)**	**Poi-INAR(2)**	**NB-INAR(2)**	**Poi-INARMA(1,1)**	**NB-INARMA(1,1)**	**Poi-INAR(1)**	**NB-INAR(1)**	**Poi-INAR(2)**	**NB-INAR(2)**	**Poi-INARMA(1,1)**	**NB-INARMA(1,1)**
INAR(1)	3	0.35	100	225	*527*	42	91	45	70	505	*427*	27	10	23	8
$\nu = 1.5$			250	34	*723*	7	132	15	89	189	*774*	7	15	7	8
			500	0	*765*	0	139	0	96	16	*959*	0	17	1	7
			1000	0	*770*	0	157	0	73	0	*984*	0	12	0	4
	3	0.70	100	324	*471*	64	58	49	34	604	*327*	34	8	21	6
			250	82	*671*	22	123	27	75	279	*665*	22	13	14	7
			500	3	*794*	3	135	1	64	58	*914*	6	18	0	4
			1000	0	*786*	0	152	0	62	3	*981*	0	12	0	4
	6	0.35	100	238	*508*	49	96	35	74	506	*412*	37	14	19	12
			250	36	*720*	14	121	14	95	194	*775*	8	12	5	6
			500	2	*748*	0	161	2	87	20	*957*	0	17	2	4
			1000	0	*773*	0	153	0	74	1	*987*	0	11	0	1
	6	0.70	100	386	*376*	93	46	78	21	693	*241*	40	2	24	0
			250	158	*608*	51	106	32	45	447	*497*	26	8	17	5
			500	17	*746*	11	143	7	76	154	*808*	12	12	14	0
			1000	1	*769*	1	150	1	78	6	*961*	1	18	5	9
INAR(2)	3	0.35	100	104	101	349	*445*	1	0	294	138	358	*209*	1	0
$\alpha_2 = 0.25$			250	0	8	129	*863*	0	0	19	40	376	*565*	0	0
$\nu = 1.5$			500	0	0	22	*978*	0	0	0	0	101	*899*	0	0
			1000	0	0	0	*1000*	0	0	0	0	2	*998*	0	0
	3	0.70	100	110	65	526	*296*	2	1	250	85	546	*119*	0	0
			250	2	6	292	*699*	1	0	24	21	564	*391*	0	0
			500	0	0	82	*918*	0	0	0	0	267	*733*	0	0
			1000	0	0	4	*996*	0	0	0	0	52	*948*	0	0
	6	0.35	100	108	81	376	*435*	0	0	291	97	428	*184*	0	0
			250	4	4	138	*854*	0	0	31	46	349	*574*	0	0
			500	0	0	13	*987*	0	0	0	1	117	*882*	0	0
			1000	0	0	0	*1000*	0	0	0	0	6	*994*	0	0
	6	0.70	100	110	76	548	*263*	2	1	260	85	558	*97*	0	0
			250	0	9	403	*588*	0	0	21	24	686	*269*	0	0
			500	0	1	144	*855*	0	0	0	1	463	*536*	0	0
			1000	0	0	16	*984*	0	0	0	0	146	*854*	0	0
INARMA(1,1)	3	0.35	100	140	472	34	64	72	*218*	403	472	18	8	37	*62*
$\beta_1 = 0.25$			250	10	423	2	58	20	*487*	69	746	2	10	20	*153*
$\nu = 1.5$			500	0	280	0	39	1	*680*	3	707	0	8	4	*278*
			1000	0	88	0	23	0	*889*	0	469	0	3	0	*528*
	3	0.70	100	175	232	55	65	238	*235*	447	238	36	8	187	*84*
			250	14	206	5	50	110	*615*	125	394	5	14	191	*271*
			500	0	60	0	13	13	*914*	6	291	0	4	71	*628*
			1000	0	3	0	1	0	*996*	0	58	0	1	4	*937*
	6	0.35	100	122	506	23	59	60	*230*	357	528	14	9	32	*60*
			250	9	465	2	64	9	*451*	69	767	2	13	22	*127*
			500	0	314	0	52	0	*634*	4	738	0	8	1	*249*
			1000	0	126	0	18	0	*856*	0	543	0	1	0	*456*
	6	0.70	100	186	246	48	30	292	*198*	461	225	28	4	232	*50*
			250	22	210	8	52	166	*542*	135	402	4	7	273	*179*
			500	2	71	0	21	38	*868*	19	311	0	3	184	*483*
			1000	0	4	0	2	1	*993*	0	79	0	1	24	*896*

Appendix B.3. Properties of (Mis-)Fitted Models

Relevant stochastic properties of the true DGP (highlighted in italic font) are compared to the mean of the corresponding properties for the fitted models (fitted to each of the 1000 simulation runs).

		True DGP: Poi-INAR(1) with $\mu = 3$ and									
		$\rho(1) = 0.35$					$\rho(1) = 0.70$				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*3.000*	*1.000*	*0.350*	*0.123*	*0.043*	*3.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	3.001	1.000	0.338	0.122	0.047	2.966	1.000	0.691	0.480	0.335
	NB-INAR(1)	3.001	1.054	0.349	0.130	0.051	2.967	1.042	0.697	0.489	0.344
	Poi-INAR(2)	3.001	1.008	0.340	0.151	0.065	2.969	1.046	0.700	0.503	0.363
	NB-INAR(2)	3.001	1.062	0.351	0.158	0.069	2.970	1.086	0.707	0.511	0.372
	Poi-INARMA(1,1)	3.002	1.020	0.345	0.095	0.034	2.966	1.049	0.701	0.481	0.332
	NB-INARMA(1,1)	3.002	1.066	0.354	0.107	0.040	2.968	1.087	0.707	0.490	0.342
250	*DGP*	*3.000*	*1.000*	*0.350*	*0.123*	*0.043*	*3.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	2.998	1.000	0.348	0.124	0.045	3.009	1.000	0.699	0.490	0.343
	NB-INAR(1)	2.998	1.034	0.356	0.130	0.049	3.009	1.028	0.703	0.496	0.350
	Poi-INAR(2)	2.997	1.006	0.350	0.144	0.059	3.010	1.031	0.705	0.505	0.363
	NB-INAR(2)	2.997	1.038	0.357	0.149	0.061	3.010	1.056	0.709	0.511	0.368
	Poi-INARMA(1,1)	2.998	1.019	0.353	0.108	0.037	3.009	1.035	0.707	0.491	0.342
	NB-INARMA(1,1)	2.998	1.044	0.358	0.115	0.041	3.010	1.060	0.710	0.497	0.349
500	*DGP*	*3.000*	*1.000*	*0.350*	*0.123*	*0.043*	*3.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	2.999	1.000	0.345	0.120	0.043	2.997	1.000	0.698	0.488	0.341
	NB-INAR(1)	3.000	1.023	0.350	0.124	0.044	2.997	1.021	0.702	0.493	0.346
	Poi-INAR(2)	2.999	1.004	0.346	0.134	0.052	2.996	1.025	0.704	0.501	0.358
	NB-INAR(2)	3.000	1.025	0.351	0.137	0.053	2.997	1.042	0.706	0.505	0.361
	Poi-INARMA(1,1)	3.000	1.016	0.349	0.110	0.037	2.997	1.024	0.703	0.489	0.341
	NB-INARMA(1,1)	3.000	1.032	0.352	0.114	0.039	2.998	1.042	0.706	0.494	0.345
1000	*DGP*	*3.000*	*1.000*	*0.350*	*0.123*	*0.043*	*3.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	2.999	1.000	0.350	0.123	0.044	3.001	1.000	0.700	0.490	0.343
	NB-INAR(1)	2.999	1.019	0.354	0.126	0.045	3.000	1.015	0.702	0.493	0.347
	Poi-INAR(2)	2.998	1.003	0.351	0.134	0.051	3.000	1.016	0.703	0.499	0.354
	NB-INAR(2)	2.998	1.021	0.355	0.136	0.052	3.001	1.029	0.706	0.502	0.357
	Poi-INARMA(1,1)	2.999	1.014	0.354	0.117	0.040	3.001	1.018	0.704	0.491	0.343
	NB-INARMA(1,1)	2.999	1.026	0.356	0.121	0.042	3.001	1.031	0.706	0.494	0.346

		True DGP: Poi-INAR(1) with $\mu = 6$ and									
		$\rho(1) = 0.35$					$\rho(1) = 0.70$				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*6.000*	*1.000*	*0.350*	*0.123*	*0.043*	*6.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	6.015	1.000	0.337	0.122	0.046	6.009	1.000	0.696	0.486	0.341
	NB-INAR(1)	6.015	1.056	0.350	0.131	0.052	6.009	1.048	0.704	0.498	0.354
	Poi-INAR(2)	6.016	1.008	0.340	0.152	0.065	6.007	1.045	0.703	0.507	0.366
	NB-INAR(2)	6.015	1.061	0.348	0.157	0.068	6.002	1.085	0.709	0.514	0.374
	Poi-INARMA(1,1)	6.016	1.020	0.346	0.095	0.034	6.009	1.055	0.707	0.488	0.338
	NB-INARMA(1,1)	6.016	1.066	0.354	0.103	0.039	6.009	1.096	0.714	0.499	0.351
250	*DGP*	*6.000*	*1.000*	*0.350*	*0.123*	*0.043*	*6.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	5.995	1.000	0.347	0.124	0.045	5.991	1.000	0.697	0.487	0.341
	NB-INAR(1)	5.995	1.033	0.355	0.129	0.048	5.991	1.033	0.704	0.496	0.350
	Poi-INAR(2)	5.995	1.006	0.349	0.144	0.058	5.989	1.033	0.704	0.504	0.361
	NB-INAR(2)	5.995	1.037	0.353	0.146	0.059	5.987	1.057	0.708	0.508	0.366
	Poi-INARMA(1,1)	5.995	1.019	0.352	0.105	0.036	5.991	1.037	0.705	0.489	0.339
	NB-INARMA(1,1)	5.996	1.043	0.356	0.109	0.038	5.991	1.064	0.710	0.497	0.348
500	*DGP*	*6.000*	*1.000*	*0.350*	*0.123*	*0.043*	*6.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	5.999	1.000	0.350	0.124	0.044	5.989	1.000	0.700	0.490	0.343
	NB-INAR(1)	5.999	1.025	0.356	0.128	0.047	5.990	1.028	0.705	0.497	0.351
	Poi-INAR(2)	5.999	1.005	0.351	0.139	0.054	5.989	1.024	0.704	0.502	0.358
	NB-INAR(2)	5.999	1.027	0.354	0.140	0.055	5.985	1.044	0.707	0.505	0.361
	Poi-INARMA(1,1)	5.999	1.015	0.353	0.113	0.039	5.989	1.026	0.705	0.491	0.342
	NB-INARMA(1,1)	6.000	1.032	0.357	0.118	0.041	5.990	1.048	0.710	0.498	0.350
1000	*DGP*	*6.000*	*1.000*	*0.350*	*0.123*	*0.043*	*6.000*	*1.000*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	6.006	1.000	0.349	0.123	0.043	6.006	1.000	0.700	0.490	0.344
	NB-INAR(1)	6.006	1.018	0.353	0.126	0.045	6.006	1.017	0.704	0.495	0.349
	Poi-INAR(2)	6.006	1.003	0.350	0.133	0.051	6.005	1.016	0.703	0.499	0.354
	NB-INAR(2)	6.007	1.020	0.350	0.133	0.050	5.998	1.028	0.704	0.499	0.354
	Poi-INARMA(1,1)	6.006	1.013	0.352	0.116	0.040	6.006	1.018	0.704	0.491	0.343
	NB-INARMA(1,1)	6.006	1.024	0.355	0.120	0.041	6.006	1.031	0.707	0.496	0.348

		True DGP: Poi-INAR(1) with $\mu = 3, \nu = 1.5$ and									
		$\rho(1) = 0.35$					**$\rho(1) = 0.70$**				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*3.000*	*1.370*	*0.350*	*0.123*	*0.043*	*3.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	2.983	1.000	0.282	0.086	0.028	3.013	1.000	0.651	0.426	0.281
	NB-INAR(1)	2.983	1.348	0.332	0.118	0.044	3.014	1.274	0.689	0.478	0.333
	Poi-INAR(2)	2.983	1.006	0.283	0.116	0.044	3.014	1.053	0.661	0.456	0.316
	NB-INAR(2)	2.983	1.352	0.333	0.146	0.061	3.015	1.314	0.697	0.498	0.357
	Poi-INARMA(1,1)	2.984	1.036	0.302	0.063	0.016	3.013	1.066	0.663	0.423	0.272
	NB-INARMA(1,1)	2.984	1.359	0.342	0.093	0.033	3.015	1.327	0.700	0.478	0.328
250	*DGP*	*3.000*	*1.370*	*0.350*	*0.123*	*0.043*	*3.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	3.005	1.000	0.291	0.087	0.027	2.985	1.000	0.656	0.432	0.285
	NB-INAR(1)	3.005	1.360	0.344	0.122	0.044	2.985	1.283	0.695	0.484	0.338
	Poi-INAR(2)	3.005	1.005	0.292	0.110	0.040	2.985	1.049	0.666	0.460	0.318
	NB-INAR(2)	3.005	1.362	0.345	0.140	0.056	2.984	1.312	0.701	0.500	0.357
	Poi-INARMA(1,1)	3.006	1.042	0.308	0.065	0.015	2.985	1.059	0.668	0.430	0.278
	NB-INARMA(1,1)	3.005	1.368	0.349	0.104	0.035	2.985	1.321	0.703	0.485	0.336
500	*DGP*	*3.000*	*1.370*	*0.350*	*0.123*	*0.043*	*3.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	3.002	1.000	0.292	0.086	0.026	3.003	1.000	0.657	0.433	0.285
	NB-INAR(1)	3.002	1.366	0.347	0.122	0.043	3.004	1.293	0.698	0.488	0.341
	Poi-INAR(2)	3.002	1.004	0.292	0.106	0.037	3.004	1.048	0.667	0.462	0.319
	NB-INAR(2)	3.002	1.368	0.348	0.135	0.052	3.004	1.317	0.703	0.501	0.357
	Poi-INARMA(1,1)	3.002	1.044	0.309	0.065	0.014	3.003	1.053	0.668	0.432	0.279
	NB-INARMA(1,1)	3.002	1.373	0.351	0.111	0.037	3.004	1.317	0.703	0.489	0.340
1000	*DGP*	*3.000*	*1.370*	*0.350*	*0.123*	*0.043*	*3.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	2.997	1.000	0.291	0.085	0.025	2.994	1.000	0.658	0.433	0.286
	NB-INAR(1)	2.997	1.367	0.347	0.121	0.043	2.994	1.289	0.699	0.488	0.342
	Poi-INAR(2)	2.997	1.004	0.292	0.104	0.036	2.994	1.045	0.667	0.461	0.318
	NB-INAR(2)	2.997	1.368	0.348	0.132	0.050	2.994	1.306	0.702	0.498	0.353
	Poi-INARMA(1,1)	2.998	1.045	0.308	0.064	0.014	2.994	1.048	0.668	0.433	0.280
	NB-INARMA(1,1)	2.997	1.372	0.350	0.115	0.038	2.995	1.306	0.702	0.489	0.341

		True DGP: Poi-INAR(1) with $\mu = 6, \nu = 1.5$ and									
		$\rho(1) = 0.35$					**$\rho(1) = 0.70$**				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*6.000*	*1.370*	*0.350*	*0.123*	*0.043*	*6.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	6.002	1.000	0.278	0.083	0.026	6.018	1.000	0.640	0.412	0.267
	NB-INAR(1)	6.002	1.355	0.334	0.119	0.045	6.020	1.266	0.686	0.474	0.329
	Poi-INAR(2)	6.002	1.006	0.280	0.112	0.042	6.020	1.064	0.652	0.449	0.310
	NB-INAR(2)	6.002	1.357	0.335	0.145	0.061	6.021	1.309	0.694	0.496	0.357
	Poi-INARMA(1,1)	6.004	1.038	0.301	0.060	0.015	6.018	1.076	0.655	0.407	0.256
	NB-INARMA(1,1)	6.004	1.363	0.341	0.090	0.031	6.020	1.314	0.696	0.471	0.322
250	*DGP*	*6.000*	*1.370*	*0.350*	*0.123*	*0.043*	*6.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	6.012	1.000	0.286	0.084	0.025	5.982	1.000	0.646	0.418	0.272
	NB-INAR(1)	6.012	1.359	0.344	0.121	0.044	5.982	1.274	0.694	0.483	0.337
	Poi-INAR(2)	6.014	1.005	0.287	0.108	0.039	5.982	1.059	0.658	0.454	0.313
	NB-INAR(2)	6.014	1.359	0.343	0.139	0.056	5.981	1.300	0.699	0.499	0.357
	Poi-INARMA(1,1)	6.013	1.044	0.307	0.063	0.014	5.982	1.066	0.660	0.416	0.264
	NB-INARMA(1,1)	6.013	1.364	0.347	0.101	0.034	5.982	1.303	0.700	0.482	0.333
500	*DGP*	*6.000*	*1.370*	*0.350*	*0.123*	*0.043*	*6.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	5.999	1.000	0.289	0.085	0.025	6.003	1.000	0.647	0.419	0.271
	NB-INAR(1)	5.999	1.366	0.348	0.123	0.044	6.003	1.284	0.697	0.487	0.340
	Poi-INAR(2)	5.999	1.004	0.290	0.107	0.038	6.003	1.056	0.658	0.454	0.313
	NB-INAR(2)	5.999	1.366	0.348	0.138	0.054	6.002	1.304	0.701	0.499	0.355
	Poi-INARMA(1,1)	6.000	1.046	0.309	0.063	0.013	6.003	1.067	0.661	0.417	0.264
	NB-INARMA(1,1)	5.999	1.370	0.351	0.109	0.036	6.003	1.307	0.702	0.486	0.337
1000	*DGP*	*6.000*	*1.370*	*0.350*	*0.123*	*0.043*	*6.000*	*1.294*	*0.700*	*0.490*	*0.343*
	Poi-INAR(1)	6.000	1.000	0.288	0.084	0.024	6.000	1.000	0.647	0.419	0.271
	NB-INAR(1)	6.000	1.366	0.348	0.122	0.043	6.000	1.292	0.699	0.489	0.342
	Poi-INAR(2)	6.000	1.004	0.289	0.103	0.036	6.001	1.056	0.659	0.455	0.313
	NB-INAR(2)	6.000	1.366	0.348	0.133	0.050	6.001	1.307	0.702	0.498	0.353
	Poi-INARMA(1,1)	6.000	1.046	0.308	0.062	0.013	6.000	1.068	0.661	0.417	0.263
	NB-INARMA(1,1)	6.000	1.369	0.350	0.114	0.038	6.000	1.309	0.702	0.488	0.340

		True DGP: Poi-INAR(2) with $\mu = 3$, $\alpha_2 = 0.25$ and									
		$\rho(1) = 0.35$					**$\rho(1) = 0.70$**				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*3.000*	*1.056*	*0.350*	*0.342*	*0.177*	*3.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	2.989	1.000	0.310	0.109	0.041	3.018	1.000	0.602	0.368	0.227
	NB-INAR(1)	2.989	1.068	0.321	0.117	0.046	3.019	1.101	0.613	0.381	0.239
	Poi-INAR(2)	2.989	1.050	0.324	0.317	0.153	3.032	1.337	0.668	0.576	0.450
	NB-INAR(2)	2.989	1.101	0.333	0.327	0.161	3.033	1.394	0.675	0.583	0.459
	Poi-INARMA(1,1)	2.989	1.002	0.310	0.108	0.041	3.018	1.014	0.605	0.367	0.225
	NB-INARMA(1,1)	2.989	1.069	0.322	0.117	0.046	3.019	1.110	0.615	0.381	0.238
250	*DGP*	*3.000*	*1.056*	*0.350*	*0.342*	*0.177*	*3.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	2.995	1.000	0.326	0.111	0.039	2.998	1.000	0.617	0.383	0.239
	NB-INAR(1)	2.996	1.051	0.335	0.117	0.043	2.998	1.102	0.628	0.396	0.251
	Poi-INAR(2)	2.995	1.055	0.341	0.335	0.169	2.998	1.365	0.689	0.601	0.481
	NB-INAR(2)	2.996	1.086	0.347	0.340	0.175	2.999	1.410	0.694	0.608	0.488
	Poi-INARMA(1,1)	2.995	1.000	0.326	0.111	0.039	2.998	1.004	0.618	0.383	0.238
	NB-INARMA(1,1)	2.996	1.051	0.336	0.118	0.043	2.998	1.103	0.628	0.396	0.251
500	*DGP*	*3.000*	*1.056*	*0.350*	*0.342*	*0.177*	*3.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	3.002	1.000	0.331	0.112	0.039	3.000	1.000	0.621	0.387	0.241
	NB-INAR(1)	3.002	1.047	0.340	0.118	0.042	3.000	1.109	0.633	0.401	0.255
	Poi-INAR(2)	3.002	1.055	0.346	0.338	0.174	3.001	1.376	0.695	0.610	0.490
	NB-INAR(2)	3.002	1.079	0.351	0.343	0.178	3.001	1.406	0.699	0.614	0.495
	Poi-INARMA(1,1)	3.002	1.000	0.331	0.112	0.039	3.000	1.002	0.621	0.387	0.241
	NB-INARMA(1,1)	3.002	1.047	0.341	0.119	0.042	3.000	1.109	0.633	0.401	0.255
1000	*DGP*	*3.000*	*1.056*	*0.350*	*0.342*	*0.177*	*3.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	3.000	1.000	0.334	0.112	0.038	2.992	1.000	0.623	0.389	0.243
	NB-INAR(1)	3.000	1.043	0.342	0.118	0.041	2.993	1.113	0.636	0.405	0.258
	Poi-INAR(2)	3.000	1.056	0.348	0.339	0.175	2.993	1.379	0.698	0.613	0.495
	NB-INAR(2)	3.000	1.074	0.352	0.343	0.179	2.993	1.404	0.701	0.617	0.499
	Poi-INARMA(1,1)	3.000	1.000	0.334	0.112	0.038	2.992	1.000	0.623	0.389	0.243
	NB-INARMA(1,1)	3.000	1.044	0.343	0.119	0.042	2.993	1.113	0.636	0.405	0.258

		True DGP: Poi-INAR(2) with $\mu = 6$, $\alpha_2 = 0.25$ and									
		$\rho(1) = 0.35$					$\rho(1) = 0.70$				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*6.000*	*1.056*	*0.350*	*0.342*	*0.177*	*6.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	5.988	1.000	0.315	0.112	0.043	6.032	1.000	0.610	0.376	0.234
	NB-INAR(1)	5.989	1.070	0.329	0.122	0.049	6.033	1.126	0.627	0.398	0.255
	Poi-INAR(2)	5.987	1.053	0.328	0.320	0.157	6.035	1.339	0.674	0.577	0.453
	NB-INAR(2)	5.988	1.106	0.338	0.330	0.167	6.037	1.418	0.684	0.587	0.465
	Poi-INARMA(1,1)	5.988	1.001	0.315	0.111	0.042	6.032	1.008	0.611	0.375	0.233
	NB-INARMA(1,1)	5.988	1.071	0.331	0.123	0.049	6.033	1.130	0.629	0.399	0.256
250	*DGP*	*6.000*	*1.056*	*0.350*	*0.342*	*0.177*	*6.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	5.990	1.000	0.329	0.113	0.040	5.982	1.000	0.617	0.383	0.238
	NB-INAR(1)	5.990	1.064	0.342	0.122	0.045	5.982	1.145	0.638	0.409	0.263
	Poi-INAR(2)	5.990	1.055	0.343	0.333	0.170	5.983	1.364	0.689	0.601	0.480
	NB-INAR(2)	5.991	1.094	0.351	0.341	0.177	5.984	1.413	0.696	0.608	0.489
	Poi-INARMA(1,1)	5.991	1.000	0.329	0.113	0.040	5.982	1.003	0.618	0.382	0.238
	NB-INARMA(1,1)	5.991	1.065	0.343	0.123	0.046	5.982	1.145	0.638	0.409	0.264
500	*DGP*	*6.000*	*1.056*	*0.350*	*0.342*	*0.177*	*6.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	5.997	1.000	0.332	0.112	0.038	5.999	1.000	0.622	0.388	0.242
	NB-INAR(1)	5.997	1.054	0.343	0.120	0.043	5.999	1.154	0.644	0.416	0.269
	Poi-INAR(2)	5.997	1.055	0.346	0.338	0.173	5.999	1.374	0.695	0.610	0.490
	NB-INAR(2)	5.997	1.083	0.352	0.344	0.178	5.999	1.414	0.701	0.615	0.497
	Poi-INARMA(1,1)	5.997	1.000	0.332	0.112	0.039	5.999	1.001	0.622	0.388	0.242
	NB-INARMA(1,1)	5.996	1.055	0.345	0.121	0.043	5.999	1.154	0.645	0.416	0.270
1000	*DGP*	*6.000*	*1.056*	*0.350*	*0.342*	*0.177*	*6.000*	*1.384*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	5.996	1.000	0.333	0.112	0.038	6.013	1.000	0.624	0.390	0.244
	NB-INAR(1)	5.996	1.049	0.344	0.120	0.042	6.013	1.159	0.648	0.420	0.273
	Poi-INAR(2)	5.996	1.055	0.348	0.339	0.175	6.013	1.381	0.699	0.615	0.496
	NB-INAR(2)	5.996	1.075	0.352	0.343	0.178	6.013	1.408	0.703	0.618	0.501
	Poi-INARMA(1,1)	5.996	1.000	0.334	0.112	0.038	6.012	1.000	0.625	0.390	0.244
	NB-INARMA(1,1)	5.996	1.050	0.345	0.121	0.043	6.013	1.159	0.648	0.420	0.273

		True DGP: NB-INAR(2) with $\mu = 3$, $\alpha_2 = 0.25$, $\nu = 1.5$ and									
		$\rho(1) = 0.35$					$\rho(1) = 0.70$				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*3.000*	*1.352*	*0.350*	*0.342*	*0.177*	*3.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	2.986	1.000	0.272	0.084	0.028	3.006	1.000	0.584	0.346	0.208
	NB-INAR(1)	2.986	1.308	0.308	0.108	0.040	3.006	1.261	0.610	0.377	0.236
	Poi-INAR(2)	2.986	1.034	0.278	0.270	0.118	3.009	1.272	0.633	0.536	0.406
	NB-INAR(2)	2.987	1.331	0.316	0.309	0.148	3.011	1.517	0.663	0.566	0.440
	Poi-INARMA(1,1)	2.986	1.010	0.275	0.080	0.026	3.006	1.021	0.587	0.343	0.204
	NB-INARMA(1,1)	2.986	1.309	0.309	0.106	0.040	3.007	1.272	0.612	0.376	0.234
250	*DGP*	*3.000*	*1.352*	*0.350*	*0.342*	*0.177*	*3.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	2.991	1.000	0.288	0.087	0.027	3.013	1.000	0.594	0.355	0.213
	NB-INAR(1)	2.991	1.317	0.330	0.114	0.041	3.013	1.292	0.622	0.389	0.244
	Poi-INAR(2)	2.991	1.036	0.294	0.286	0.130	3.012	1.291	0.648	0.559	0.431
	NB-INAR(2)	2.991	1.336	0.339	0.330	0.167	3.013	1.573	0.683	0.596	0.474
	Poi-INARMA(1,1)	2.991	1.005	0.289	0.085	0.026	3.013	1.012	0.596	0.353	0.211
	NB-INARMA(1,1)	2.991	1.317	0.330	0.114	0.041	3.014	1.294	0.623	0.389	0.244
500	*DGP*	*3.000*	*1.352*	*0.350*	*0.342*	*0.177*	*3.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	2.998	1.000	0.290	0.086	0.026	3.009	1.000	0.600	0.361	0.218
	NB-INAR(1)	2.998	1.325	0.335	0.114	0.040	3.009	1.313	0.630	0.398	0.251
	Poi-INAR(2)	2.998	1.036	0.296	0.289	0.132	3.010	1.301	0.655	0.568	0.441
	NB-INAR(2)	2.997	1.344	0.343	0.335	0.171	3.011	1.605	0.693	0.608	0.488
	Poi-INARMA(1,1)	2.998	1.003	0.290	0.085	0.025	3.009	1.008	0.602	0.360	0.216
	NB-INARMA(1,1)	2.998	1.325	0.335	0.115	0.040	3.010	1.314	0.630	0.398	0.252
1000	*DGP*	*3.000*	*1.352*	*0.350*	*0.342*	*0.177*	*3.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	3.005	1.000	0.294	0.088	0.026	2.991	1.000	0.601	0.362	0.218
	NB-INAR(1)	3.005	1.333	0.340	0.117	0.040	2.991	1.311	0.631	0.398	0.252
	Poi-INAR(2)	3.005	1.037	0.301	0.294	0.136	2.990	1.302	0.657	0.571	0.444
	NB-INAR(2)	3.005	1.352	0.349	0.341	0.176	2.990	1.599	0.695	0.610	0.491
	Poi-INARMA(1,1)	3.005	1.002	0.295	0.087	0.026	2.991	1.004	0.602	0.361	0.217
	NB-INARMA(1,1)	3.005	1.333	0.341	0.117	0.041	2.991	1.311	0.631	0.399	0.252

		True DGP: NB-INAR(2) with $\mu = 6$, $\alpha_2 = 0.25$, $\nu = 1.5$ and									
		$\rho(1) = 0.35$					$\rho(1) = 0.70$				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*6.000*	*1.352*	*0.350*	*0.342*	*0.177*	*6.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	6.020	1.000	0.277	0.086	0.029	5.996	1.000	0.575	0.336	0.198
	NB-INAR(1)	6.020	1.299	0.319	0.114	0.044	5.996	1.299	0.615	0.384	0.242
	Poi-INAR(2)	6.021	1.036	0.286	0.282	0.126	5.995	1.253	0.622	0.527	0.395
	NB-INAR(2)	6.022	1.313	0.326	0.322	0.158	6.001	1.527	0.659	0.562	0.435
	Poi-INARMA(1,1)	6.021	1.007	0.279	0.084	0.027	5.996	1.019	0.579	0.334	0.195
	NB-INARMA(1,1)	6.021	1.301	0.321	0.115	0.044	5.997	1.305	0.617	0.384	0.242
250	*DGP*	*6.000*	*1.352*	*0.350*	*0.342*	*0.177*	*6.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	5.998	1.000	0.283	0.084	0.026	5.990	1.000	0.588	0.347	0.206
	NB-INAR(1)	5.998	1.326	0.332	0.115	0.042	5.990	1.336	0.632	0.402	0.256
	Poi-INAR(2)	5.999	1.035	0.291	0.283	0.127	5.991	1.287	0.642	0.556	0.427
	NB-INAR(2)	5.999	1.341	0.338	0.329	0.166	5.990	1.564	0.682	0.594	0.472
	Poi-INARMA(1,1)	5.998	1.005	0.284	0.082	0.025	5.990	1.010	0.590	0.346	0.204
	NB-INARMA(1,1)	5.999	1.327	0.333	0.116	0.042	5.991	1.337	0.633	0.402	0.256
500	*DGP*	*6.000*	*1.352*	*0.350*	*0.342*	*0.177*	*6.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	5.997	1.000	0.288	0.084	0.025	5.999	1.000	0.593	0.353	0.211
	NB-INAR(1)	5.997	1.330	0.338	0.117	0.041	6.000	1.353	0.639	0.410	0.263
	Poi-INAR(2)	5.997	1.035	0.295	0.289	0.132	6.000	1.297	0.650	0.566	0.438
	NB-INAR(2)	5.997	1.342	0.344	0.337	0.172	6.000	1.591	0.691	0.606	0.486
	Poi-INARMA(1,1)	5.997	1.003	0.288	0.083	0.025	6.000	1.006	0.595	0.353	0.210
	NB-INARMA(1,1)	5.997	1.331	0.339	0.117	0.041	6.000	1.353	0.640	0.410	0.263
1000	*DGP*	*6.000*	*1.352*	*0.350*	*0.342*	*0.177*	*6.000*	*1.620*	*0.700*	*0.618*	*0.499*
	Poi-INAR(1)	5.995	1.000	0.289	0.084	0.025	5.997	1.000	0.595	0.354	0.211
	NB-INAR(1)	5.995	1.339	0.341	0.117	0.041	5.997	1.359	0.641	0.411	0.264
	Poi-INAR(2)	5.995	1.035	0.296	0.290	0.132	5.998	1.298	0.651	0.569	0.441
	NB-INAR(2)	5.995	1.350	0.346	0.339	0.174	5.998	1.601	0.695	0.610	0.491
	Poi-INARMA(1,1)	5.995	1.002	0.289	0.084	0.024	5.997	1.003	0.595	0.354	0.211
	NB-INARMA(1,1)	5.995	1.339	0.341	0.118	0.041	5.997	1.359	0.641	0.412	0.265

		True DGP: Poi-INARMA(1,1) with $\mu = 3$, $\beta_1 = 0.25$ and									
		$\rho(1) = 0.35$					**$\rho(1) = 0.70$**				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*3.000*	*1.067*	*0.350*	*0.070*	*0.014*	*3.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	3.007	1.000	0.330	0.115	0.042	2.995	1.000	0.659	0.437	0.291
	NB-INAR(1)	3.007	1.081	0.343	0.125	0.047	2.995	1.073	0.669	0.450	0.304
	Poi-INAR(2)	3.007	1.004	0.331	0.128	0.050	2.995	1.025	0.664	0.451	0.308
	NB-INAR(2)	3.008	1.083	0.344	0.136	0.055	2.996	1.093	0.674	0.463	0.320
	Poi-INARMA(1,1)	3.009	1.033	0.347	0.067	0.021	2.996	1.151	0.690	0.432	0.274
	NB-INARMA(1,1)	3.009	1.099	0.356	0.075	0.026	2.996	1.211	0.698	0.445	0.287
250	*DGP*	*3.000*	*1.067*	*0.350*	*0.070*	*0.014*	*3.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	2.997	1.000	0.331	0.112	0.039	3.004	1.000	0.665	0.443	0.296
	NB-INAR(1)	2.997	1.060	0.342	0.119	0.043	3.004	1.050	0.672	0.452	0.305
	Poi-INAR(2)	2.996	1.001	0.332	0.118	0.042	3.004	1.014	0.668	0.451	0.306
	NB-INAR(2)	2.997	1.061	0.342	0.124	0.046	3.005	1.060	0.675	0.459	0.313
	Poi-INARMA(1,1)	2.997	1.040	0.344	0.065	0.018	3.004	1.153	0.697	0.442	0.282
	NB-INARMA(1,1)	2.997	1.080	0.350	0.071	0.021	3.004	1.184	0.701	0.449	0.289
500	*DGP*	*3.000*	*1.067*	*0.350*	*0.070*	*0.014*	*3.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	2.999	1.000	0.331	0.111	0.037	3.002	1.000	0.665	0.443	0.296
	NB-INAR(1)	2.999	1.053	0.340	0.117	0.041	3.002	1.046	0.672	0.452	0.304
	Poi-INAR(2)	2.998	1.001	0.331	0.113	0.039	3.003	1.009	0.667	0.449	0.302
	NB-INAR(2)	2.999	1.054	0.341	0.119	0.042	3.003	1.052	0.673	0.456	0.309
	Poi-INARMA(1,1)	2.999	1.046	0.343	0.065	0.016	3.003	1.150	0.697	0.444	0.283
	NB-INARMA(1,1)	2.999	1.076	0.348	0.069	0.018	3.003	1.175	0.701	0.449	0.289
1000	*DGP*	*3.000*	*1.067*	*0.350*	*0.070*	*0.014*	*3.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	2.996	1.000	0.334	0.112	0.038	2.993	1.000	0.666	0.443	0.295
	NB-INAR(1)	2.996	1.052	0.343	0.118	0.041	2.993	1.039	0.671	0.451	0.303
	Poi-INAR(2)	2.996	1.000	0.334	0.113	0.038	2.993	1.004	0.667	0.446	0.298
	NB-INAR(2)	2.996	1.052	0.343	0.119	0.041	2.994	1.041	0.672	0.452	0.305
	Poi-INARMA(1,1)	2.997	1.055	0.348	0.069	0.016	2.993	1.152	0.698	0.444	0.283
	NB-INARMA(1,1)	2.997	1.076	0.351	0.072	0.017	2.993	1.168	0.701	0.448	0.287

		True DGP: Poi-INARMA(1,1) with $\mu = 6$, $\beta_1 = 0.25$ and									
		$\rho(1) = 0.35$					**$\rho(1) = 0.70$**				
T	**Properties for**	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*6.000*	*1.067*	*0.350*	*0.070*	*0.014*	*6.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	5.999	1.000	0.327	0.114	0.042	5.982	1.000	0.660	0.438	0.292
	NB-INAR(1)	6.000	1.088	0.344	0.126	0.049	5.982	1.080	0.673	0.456	0.310
	Poi-INAR(2)	5.999	1.003	0.328	0.126	0.049	5.985	1.026	0.665	0.453	0.310
	NB-INAR(2)	5.998	1.090	0.344	0.137	0.056	5.987	1.099	0.678	0.468	0.325
	Poi-INARMA(1,1)	6.001	1.033	0.345	0.065	0.020	5.982	1.151	0.690	0.433	0.276
	NB-INARMA(1,1)	6.002	1.102	0.354	0.072	0.025	5.983	1.204	0.699	0.447	0.291
250	*DGP*	*6.000*	*1.067*	*0.350*	*0.070*	*0.014*	*6.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	6.003	1.000	0.334	0.114	0.040	5.998	1.000	0.665	0.443	0.296
	NB-INAR(1)	6.003	1.070	0.348	0.124	0.045	5.999	1.070	0.677	0.459	0.312
	Poi-INAR(2)	6.003	1.001	0.335	0.119	0.043	5.997	1.014	0.668	0.451	0.305
	NB-INAR(2)	6.002	1.070	0.347	0.127	0.047	5.998	1.078	0.679	0.465	0.319
	Poi-INARMA(1,1)	6.004	1.039	0.347	0.066	0.018	5.999	1.149	0.696	0.443	0.283
	NB-INARMA(1,1)	6.004	1.086	0.354	0.073	0.022	5.999	1.186	0.702	0.453	0.294
500	*DGP*	*6.000*	*1.067*	*0.350*	*0.070*	*0.014*	*6.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	6.000	1.000	0.332	0.112	0.038	5.991	1.000	0.666	0.444	0.296
	NB-INAR(1)	6.000	1.061	0.345	0.120	0.043	5.991	1.066	0.678	0.460	0.312
	Poi-INAR(2)	6.000	1.001	0.333	0.114	0.039	5.992	1.008	0.668	0.449	0.302
	NB-INAR(2)	6.000	1.061	0.344	0.122	0.043	5.992	1.070	0.679	0.463	0.316
	Poi-INARMA(1,1)	6.000	1.046	0.345	0.067	0.017	5.991	1.154	0.699	0.445	0.284
	NB-INARMA(1,1)	6.000	1.079	0.350	0.071	0.019	5.991	1.182	0.703	0.453	0.293
1000	*DGP*	*6.000*	*1.067*	*0.350*	*0.070*	*0.014*	*6.000*	*1.156*	*0.700*	*0.446*	*0.284*
	Poi-INAR(1)	6.001	1.000	0.332	0.111	0.037	6.016	1.000	0.666	0.444	0.296
	NB-INAR(1)	6.001	1.059	0.344	0.119	0.042	6.016	1.061	0.677	0.459	0.311
	Poi-INAR(2)	6.001	1.000	0.332	0.112	0.038	6.016	1.003	0.667	0.446	0.299
	NB-INAR(2)	6.001	1.059	0.343	0.119	0.042	6.017	1.063	0.678	0.461	0.313
	Poi-INARMA(1,1)	6.001	1.053	0.347	0.068	0.016	6.016	1.152	0.699	0.445	0.284
	NB-INARMA(1,1)	6.001	1.077	0.349	0.070	0.017	6.016	1.171	0.702	0.451	0.291

T	Properties for	True DGP: NB-INARMA(1,1) with $\mu = 3$, $\beta_1 = 0.25$, $\nu = 1.5$ and									
		$\rho(1) = 0.35$					$\rho(1) = 0.70$				
		μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*3.000*	*1.452*	*0.350*	*0.063*	*0.011*	*3.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	2.984	1.000	0.269	0.078	0.024	3.007	1.000	0.618	0.385	0.241
	NB-INAR(1)	2.984	1.398	0.313	0.105	0.037	3.008	1.298	0.652	0.428	0.282
	Poi-INAR(2)	2.984	1.003	0.270	0.092	0.031	3.005	1.034	0.625	0.407	0.266
	NB-INAR(2)	2.984	1.398	0.313	0.116	0.043	3.006	1.322	0.657	0.442	0.299
	Poi-INARMA(1,1)	2.986	1.041	0.305	0.048	0.010	3.009	1.164	0.650	0.366	0.211
	NB-INARMA(1,1)	2.986	1.419	0.337	0.062	0.018	3.010	1.462	0.686	0.419	0.261
250	*DGP*	*3.000*	*1.452*	*0.350*	*0.063*	*0.011*	*3.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	3.006	1.000	0.279	0.080	0.023	2.983	1.000	0.621	0.387	0.242
	NB-INAR(1)	3.006	1.421	0.328	0.110	0.038	2.983	1.314	0.658	0.434	0.287
	Poi-INAR(2)	3.007	1.001	0.280	0.087	0.027	2.983	1.024	0.626	0.404	0.261
	NB-INAR(2)	3.007	1.421	0.328	0.114	0.040	2.983	1.323	0.660	0.440	0.294
	Poi-INARMA(1,1)	3.007	1.050	0.312	0.050	0.009	2.984	1.173	0.655	0.372	0.213
	NB-INARMA(1,1)	3.007	1.443	0.345	0.062	0.016	2.984	1.477	0.694	0.430	0.268
500	*DGP*	*3.000*	*1.452*	*0.350*	*0.063*	*0.011*	*3.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	3.003	1.000	0.280	0.079	0.023	2.998	1.000	0.622	0.388	0.242
	NB-INAR(1)	3.003	1.426	0.330	0.110	0.037	2.999	1.323	0.660	0.436	0.288
	Poi-INAR(2)	3.004	1.001	0.280	0.083	0.025	2.998	1.018	0.626	0.400	0.257
	NB-INAR(2)	3.004	1.426	0.330	0.111	0.038	2.998	1.327	0.661	0.439	0.291
	Poi-INARMA(1,1)	3.004	1.052	0.313	0.050	0.009	2.999	1.176	0.657	0.372	0.212
	NB-INARMA(1,1)	3.004	1.449	0.346	0.062	0.014	2.999	1.490	0.697	0.433	0.270
1000	*DGP*	*3.000*	*1.452*	*0.350*	*0.063*	*0.011*	*3.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	2.999	1.000	0.282	0.080	0.023	2.997	1.000	0.623	0.388	0.242
	NB-INAR(1)	2.999	1.424	0.332	0.111	0.037	2.997	1.317	0.661	0.437	0.289
	Poi-INAR(2)	2.999	1.000	0.282	0.082	0.024	2.997	1.015	0.626	0.399	0.255
	NB-INAR(2)	2.999	1.424	0.331	0.111	0.037	2.997	1.318	0.661	0.438	0.290
	Poi-INARMA(1,1)	2.999	1.053	0.315	0.051	0.009	2.997	1.177	0.658	0.374	0.213
	NB-INARMA(1,1)	2.999	1.448	0.348	0.062	0.013	2.997	1.482	0.697	0.434	0.271

T	Properties for	True DGP: NB-INARMA(1,1) with $\mu = 6$, $\beta_1 = 0.25$, $\nu = 1.5$ and $\rho(1) = 0.35$					$\rho(1) = 0.70$				
		μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$	μ	σ^2/μ	$\rho(1)$	$\rho(2)$	$\rho(3)$
100	*DGP*	*6.000*	*1.452*	*0.350*	*0.063*	*0.011*	*6.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	6.008	1.000	0.272	0.079	0.025	6.025	1.000	0.607	0.372	0.229
	NB-INAR(1)	6.008	1.431	0.328	0.115	0.042	6.027	1.315	0.654	0.430	0.285
	Poi-INAR(2)	6.009	1.003	0.274	0.093	0.032	6.028	1.032	0.614	0.393	0.253
	NB-INAR(2)	6.010	1.433	0.329	0.125	0.049	6.030	1.328	0.657	0.440	0.297
	Poi-INARMA(1,1)	6.010	1.046	0.311	0.051	0.011	6.026	1.180	0.644	0.350	0.196
	NB-INARMA(1,1)	6.011	1.448	0.348	0.067	0.021	6.028	1.460	0.683	0.413	0.255
250	*DGP*	*6.000*	*1.452*	*0.350*	*0.063*	*0.011*	*6.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	5.996	1.000	0.277	0.079	0.023	5.987	1.000	0.612	0.375	0.230
	NB-INAR(1)	5.996	1.431	0.334	0.114	0.040	5.987	1.345	0.663	0.441	0.293
	Poi-INAR(2)	5.996	1.001	0.277	0.085	0.027	5.987	1.025	0.617	0.392	0.250
	NB-INAR(2)	5.997	1.430	0.333	0.117	0.042	5.987	1.351	0.665	0.445	0.299
	Poi-INARMA(1,1)	5.997	1.051	0.314	0.051	0.009	5.988	1.187	0.651	0.358	0.199
	NB-INARMA(1,1)	5.997	1.445	0.347	0.064	0.017	5.988	1.486	0.694	0.429	0.267
500	*DGP*	*6.000*	*1.452*	*0.350*	*0.063*	*0.011*	*6.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	5.999	1.000	0.278	0.078	0.022	5.992	1.000	0.614	0.378	0.233
	NB-INAR(1)	5.999	1.423	0.336	0.114	0.039	5.992	1.342	0.666	0.444	0.296
	Poi-INAR(2)	5.998	1.001	0.278	0.082	0.024	5.993	1.021	0.619	0.393	0.250
	NB-INAR(2)	5.999	1.423	0.335	0.115	0.040	5.994	1.346	0.667	0.447	0.299
	Poi-INARMA(1,1)	5.999	1.052	0.314	0.051	0.009	5.993	1.189	0.653	0.362	0.201
	NB-INARMA(1,1)	5.999	1.437	0.346	0.063	0.015	5.993	1.483	0.697	0.433	0.270
1000	*DGP*	*6.000*	*1.452*	*0.350*	*0.063*	*0.011*	*6.000*	*1.492*	*0.700*	*0.437*	*0.273*
	Poi-INAR(1)	6.001	1.000	0.279	0.078	0.022	5.996	1.000	0.615	0.378	0.233
	NB-INAR(1)	6.001	1.429	0.337	0.114	0.039	5.996	1.346	0.667	0.445	0.297
	Poi-INAR(2)	6.001	1.000	0.279	0.080	0.023	5.996	1.016	0.618	0.390	0.246
	NB-INAR(2)	6.002	1.429	0.337	0.114	0.039	5.999	1.348	0.668	0.447	0.299
	Poi-INARMA(1,1)	6.001	1.053	0.314	0.051	0.009	5.997	1.192	0.655	0.362	0.201
	NB-INARMA(1,1)	6.001	1.444	0.348	0.063	0.014	5.996	1.490	0.699	0.435	0.271

References

1. Holan, S.; Lund, R.; Davis, G. The ARMA alphabet soup: A tour of ARMA model variants. *Stat. Surv.* **2010**, *4*, 232–274. [CrossRef]
2. McKenzie, E. Some simple models for discrete variate time series. *Water Resour. Bull.* **1985**, *21*, 645–650. [CrossRef]
3. Al-Osh, M.; Alzaid, A. First-order integer-valued autoregressive process. *J. Time Ser. Anal.* **1987**, *8*, 261–275. [CrossRef]
4. Weiß, C. *An Introduction to Discrete-Valued Time Series*; John Wiley & Sons, Inc.: Chichester, UK, 2018.
5. Brännäs, K.; Hall, A. Estimation Integer-valued moving average models. *Appl. Stoch. Model. Bus. Ind.* **2001**, *17*, 277–291. [CrossRef]
6. Brännäs, K.; Quoreshi, A. Integer-valued moving average modelling of the number of transactions in stocks. *Appl. Financ. Econ.* **2010**, *20*, 1429–1440. [CrossRef]
7. Aleksandrov, B.; Weiß, C. Parameter estimation and diagnostic tests for INMA(1) processes. *TEST* **2019**, forthcoming. [CrossRef]
8. Dungey, M.; Martin, V.; Tang, C.; Tremayne, A. A threshold mixed count time series model: Estimation and application. *Stud. Nonlinear Dyn. Econom.* **2019**, forthcoming. [CrossRef]

9. Neal, P.; Subba Rao, T. MCMC for integer-valued ARMA processes. *J. Time Ser. Anal.* **2007**, *28*, 92–100. [CrossRef]
10. Enciso-Mora, V.; Neal, P.; Subba Rao, T. Efficient order selection algorithms for integer-valued ARMA processes. *J. Time Ser. Anal.* **2009**, *30*, 1–18. [CrossRef]
11. Alzahrani, N.; Neal, P.; Spencer, S.E.; McKinley, T.J.; Touloupou, P. Model selection for time series of count data. *Comput. Stat. Data Anal.* **2018**, *122*, 33–44. [CrossRef]
12. Wheatley, S.; Schatz, M.; Sornette, D. The ARMA point process and its estimation. *arXiv* **2018**, arXiv:1806.09948.
13. Steutel, F.; Van Harn, K. Discrete analogues of self-decomposability and stability. *Ann. Probab.* **1979**, *7*, 3893–3899. [CrossRef]
14. Dion, J.; Gauthier, G.; Latour, A. Branching processes with immigration and integer-valued time series. *Serdica Math.* **1995**, *21*, 123–136.
15. Alzaid, A.; Al-Osh, M. An integer-valued pth order autoregressive structure (INAR(p)) process. *J. Appl. Probab.* **1990**, *27*, 314–324. [CrossRef]
16. Du, J.; Li, Y. The integer-valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.* **1991**, *12*, 129–142.
17. McKenzie, E. Some ARMA models for dependent sequences of Poisson counts. *Adv. Appl. Probab.* **1988**, *20*, 822–835. [CrossRef]
18. Bracher, J. A new INARMA(1,1) model with Poisson marginals. In *14th Workshop on Stochastic Models, Statistics and Their Application (SMSA 2019)*; Steland, A., Okhrin, O., Rafajłowicz, E., Eds.; Springer International Publishing: Cham, Switzerland, 2019.
19. Zucchini, W.; MacDonald, I.; Langrock, R. *Hidden Markov Models for Time Series: An Introduction Using R*, 2nd ed.; Chapman & Hall/CRC: London, UK, 2016.
20. Freeland, R.; McCabe, B. Analysis of low count time series data by Poisson autoregression. *J. Time Ser. Anal.* **2004**, *25*, 701–722. [CrossRef]
21. Neath, A.; Cavanaugh, J. The Bayesian information criterion: Background, derivation, and applications. *WIREs Comput. Stat.* **2012**, *4*, 199–203. [CrossRef]
22. Cavanaugh, J.; Neath, A. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *WIREs Comput. Stat.* **2019**, forthcoming. [CrossRef]
23. Weiß, C.; Feld, H. On the performance of information criteria for model identification of count time series. *Stud. Nonlinear Dyn. Econom.* **2019**, forthcoming. [CrossRef]
24. Zhu, F.; Li, Q.; Wang, D. A mixture integer-valued ARCH model. *J. Stat. Plan. Inference* **2010**, *140*, 2025–2036. [CrossRef]
25. Freeland, R. Statistical Analysis of Discrete Time Series With Application to The Analysis of Workers' Compensation Claims Data. Ph.D. Thesis, University of British Columbia, Vancouver, BC, Canada, 1998.
26. Tsay, R.S. Model checking via parametric bootstraps in time deries analysis. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1992**, *41*, 1–15.