Testing the Compounding Structure of the CP-INARCH Model

Christian H. Weiß ·
Esmeralda Gonçalves ·
Nazaré Mendes Lopes

Received: date / Accepted: date

Abstract A statistical test to distinguish between a Poisson INARCH model and a Compound Poisson INARCH model is proposed, based on the form of the probability generating function of the compounding distribution of the conditional law of the model.

For first-order autoregression, the normality of the test statistics' asymptotic distribution is established, either in the case where the model parameters are specified, or when such parameters are consistently estimated. As the test statistics' law involves the moments of inverse conditional means of the Compound Poisson INARCH process, the analysis of their existence and calculation is performed by two approaches. For higher-order autoregressions, we use a bootstrap implementation of the test.

A simulation study illustrating the finite-sample performance of this test methodology in what concerns its size and power concludes the paper.

Keywords Count-data time series · compound Poisson distribution · INGARCH model · diagnostic tests · inverse moments

Mathematics Subject Classification (2000) 60J10 · 62M02 · 62M10

C.H. Weiß
Department of Mathematics and Statistics, Helmut Schmidt University, 22008 Hamburg, Germany.
E-mail: weissc@hsu-hh.de

E. Gonçalves
CMUC, Departamento de Matemática, Universidade de Coimbra, Portugal.
E-mail: esmerald@mat.uc.pt

N. Mendes Lopes
CMUC, Departamento de Matemática, Universidade de Coimbra, Portugal.
E-mail: nazare@mat.uc.pt

1 Introduction

The *INGARCH models*, which constitute an *in*teger-valued counterpart to the conventional *g*eneralized *a*uto*r*egressive *c*onditional *h*eteroskedasticity models, were introduced by Heinen (2003); Ferland et al. (2006). Instead of considering the conditional variances as in the conventional GARCH model, they assume the conditional means $M_t := E[X_t \mid X_{t-1}, \ldots]$ to satisfy a linear recursion,

$$M_t \;=\; \alpha_0 \;+\; \textstyle\sum_{i=1}^{p} \alpha_i\, X_{t-i} \;+\; \sum_{j=1}^{q} \beta_j\, M_{t-j}, \tag{1}$$

where $\alpha_0 > 0$ and $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \geq 0$. Having specified the conditional mean, the most common type of conditional distribution is the Poisson one, i. e., $X_t \sim \text{Poi}(M_t)$, leading to the *Poisson* INGARCH model, where existence and strict stationarity with finite first and second order moments can be shown under the condition $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ (Ferland et al., 2006). The Poisson INGARCH model was further investigated by several authors including Fokianos et al. (2009); Weiß (2009); Neumann (2011). But also different choices for the conditional distribution have been considered in the literature, see, e. g., Xu et al. (2012); Zhu (2012); Gonçalves et al. (2015a,b) and the discussion below. The INGARCH models exhibit an ARMA-like autocorrelation structure, and they are particularly well-suited for time series of counts showing overdispersion, i. e., which have a variance larger than the mean. In particular, the case $q = 0$, referred to as an INARCH(p) model, has the same autocorrelation structure as a usual AR(p) model. So the INARCH model, which is the main focus of the present work, constitutes a count-data type of autoregressive model.

The standard INARCH model has a conditional Poisson distribution and is therefore conditionally equidispersed. Its unconditional distribution, however, exhibits overdispersion, where the degree of overdispersion depends on the dependence parameters $\alpha_1, \ldots, \alpha_p$. To overcome this limitation, Xu et al. (2012) proposed the family of *dispersed INARCH models* (DINARCH), which again assume a linear relationship for the conditional mean, but with an additional scaling factor $\theta > 0$ for the conditional variance:

$$M_t \;=\; \alpha_0 + \textstyle\sum_{i=1}^{p} \alpha_i X_{t-i}, \qquad V[X_t \mid X_{t-1}, \ldots] \;=\; \theta\, M_t. \tag{2}$$

So the standard Poisson INARCH model is an instance of the DINARCH model with $\theta = 1$. A more comprehensive instance of the DINARCH model is obtained from a family of INGARCH models that was recently developed by Gonçalves et al. (2015a), who proposed to use a conditional compound Poisson (CP) distribution (Johnson et al., 2005). The CP-INARCH model to be considered in the sequel is defined by the conditional probability generating function (pgf)

$$\text{pgf}_{X_t|X_{t-1},\ldots}(z) \;=\; \exp\Big(\tfrac{M_t}{H'(1)}\,\big(H(z)-1\big)\Big) \quad \text{with } M_t \text{ according to (2)}, \tag{3}$$

where $H(z)$ denotes the pgf of the compounding distribution (assumed to be normalized to $H(0) = 0$ for uniqueness). From Theorem 5 in Gonçalves et al. (2015a), we know that the above condition $\sum_{i=1}^{p} \alpha_i < 1$ again guarantees the existence of a strictly stationary and ergodic solution to the CP-INARCH model (3), and this solution has finite first and second order moments. The CP-INARCH model constitutes an instance of the DINARCH model, where

$$V[X_t \mid X_{t-1}, \ldots] \;=\; M_t\, \underbrace{\big(1 + H''(1)/H'(1)\big)}_{=\theta}. \tag{4}$$

Example 1 (Special CP-INARCH Models) Choosing $H(z) = z$, we obtain the standard Poisson INARCH model. But also the NB-INARCH(p) model (*negative binomial*) proposed by Xu et al. (2012) is a special type of CP-INARCH(p) model, where the compounding distribution is a log-series distribution (Johnson et al., 2005),

$$H(z,\theta) \;=\; 1 - \frac{\ln\big(\theta + (1-\theta)z\big)}{\ln\theta} \quad \text{with } H'(1,\theta) \;=\; -\frac{1-\theta}{\ln\theta}, \quad H''(1,\theta) \;=\; \frac{(1-\theta)^2}{\ln\theta}. \tag{5}$$

Hence, we simply have $1 + H''(1,\theta)/H'(1,\theta) \;=\; \theta$. Further examples include the INARCH model proposed by Zhu (2012) having a conditional generalized Poisson (GP) distribution, and the one by Gonçalves et al. (2015b) having a conditional *Neyman type-A* (NTA) distribution. The latter has a Poisson compounding structure: the NTA($\mu/\phi,\ \phi$)-distribution is defined by the pgf (Johnson et al., 2005)

$$\text{pgf}(z) \;=\; \exp\left(\frac{\mu}{\phi}\left(e^{\phi(z-1)} - 1\right)\right) \;=\; \exp\left(\mu\, \frac{1-e^{-\phi}}{\phi} \left(\frac{e^{\phi z} - 1}{e^{\phi} - 1} - 1\right)\right), \tag{6}$$

and for the NTA-INARCH model, the mean parameter μ is replaced by M_t. The compounding pgf, $H(z,\phi) \;=\; (e^{\phi z} - 1)/(e^{\phi} - 1)$ with $H^{(k)}(z,\phi) \;=\; \phi^k\, e^{\phi z}/(e^{\phi} - 1)$, is the one from the zero-truncated Poisson distribution and therefore satisfies the normalization constraint $H(0,\phi) = 0$. In particular, we have $1 + H''(1,\phi)/H'(1,\phi) \;=\; 1 + \phi$.

In the sequel, we shall consider the problem of distinguishing between the simple Poisson INARCH model and true CP-INARCH model, i. e., we are confronted with the following hypotheses:

$$\begin{aligned} &H_0 : (X_t)_{\mathbb{Z}} \text{ is a Poisson INARCH process} \quad (\text{i. e., } H(z) = z);\\ &H_1 : (X_t)_{\mathbb{Z}} \text{ is a true CP-INARCH process} \quad (\text{i. e., } H(z) \neq z). \end{aligned} \tag{7}$$

Note that hypotheses (7) refer to the *conditional* process distribution (given the past). Therefore, such tests as proposed by Lee et al. (2017); Weiß et al. (2017), which test for *marginal* overdispersion or zero inflation (with the null being a *marginal* Poisson distribution), are not reasonable in our setup.

In Section 2, we develop a general approach for analyzing the conditional compounding structure of a CP-INARCH model. This approach is then used in Section 3 to develop a test procedure for the INARCH model, where the test statistic involves the factorial moment of order r of X_t. For first-order autoregression, the normality of the test statistics' asymptotic distribution under the null hypothesis (7) is established either in the case of specified parameters, or in that one, important in practice, where such parameters are consistently estimated. As the test statistics' law involves the moments of inverse conditional means of the Compound Poisson INGARCH process, the analysis of their existence and calculation is performed by two approaches. For higher-order autoregressions, a bootstrap implementation is presented. In Section 4, a simulation study is presented illustrating the finite-sample performance of this test methodology in what concerns its size and power for different values of r. Also a real-data example is provided. Section 5 concludes, and Appendix A includes the detailed derivations.

2 Analyzing the Compounding Structure of CP-INARCH Models

Given the past observations $X_{t-1},\ldots$, the conditional CP model in (3) implies that first a stopping count N_t is generated according to $\text{Poi}\big(M_t/H'(1)\big)$, and then (independently) the N_t i.i.d. counts $Y_{t,1},\ldots,Y_{t,N_t}$ according to the compounding model having the pgf $H(z)$, also see Johnson et al. (2005). The next observation is obtained as $X_t = Y_{t,1}+\ldots+Y_{t,N_t}$.

To distinguish between the null hypothesis H_0 and the alternative hypothesis H_1 according to (7), information about $H(z)$ is required, the unique pgf of the $Y_{t,i}$. In fact, it suffices to check if the mean $H'(1)$ of the compounding distribution is equal to 1 (H_0) or larger than 1 (H_1). Hence, the mean statistic

$$\frac{1}{T}\sum_{t=1}^{T}\frac{Y_{t,1}+\ldots+Y_{t,N_t}}{N_t} \;=\; \frac{1}{T}\sum_{t=1}^{T}\frac{X_t}{N_t}$$

would be a reasonable candidate to infer $H'(1)$. But we do not observe N_t in practice, we only know that it has mean $M_t/H'(1)$. Therefore, we may consider a slightly modified version,

$$\frac{1}{T}\sum_{t=1}^{T}\frac{X_t}{M_t},$$

which we expect to give values close to 1. Note that the summands X_t/M_t are just the residuals ε_t as defined in Zhu & Wang (2010). To be more precise, for an underlying INARCH(p) model structure, the statistic

$$\widehat{C}_p \;:=\; \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{X_t}{M_t} \;=\; \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{X_t}{\alpha_0+\sum_{i=1}^{p}\alpha_i X_{t-i}} \tag{8}$$

could be computed from the available data $X_1,\ldots,X_T$ and from the parameters $\alpha_0,\ldots,\alpha_p$ of the null model. Does this statistic allow to distinguish between H_0 and H_1?

Since the conditional mean $E[X_t \mid X_{t-1},\ldots] = M_t$ for *any* CP-INARCH process according to (2), we necessarily have

$$E\left[\frac{X_t}{M_t}\right] \;=\; 1, \qquad Cov\left[\frac{X_t}{M_t},\frac{X_{t-k}}{M_{t-k}}\right] \;=\; 0 \quad \text{for } k \geq 1,$$

which immediately follows by applying the laws of total expectation and covariance. For the variance, we obtain

$$\begin{aligned} V\left[\frac{X_t}{M_t}\right] &= V\left[\frac{E[X_t \mid X_{t-1},\ldots]}{M_t}\right] \;+\; E\left[\frac{V[X_t \mid X_{t-1},\ldots]}{M_t^2}\right] \\ &= 0 \;+\; \left(1+\frac{H''(1)}{H'(1)}\right)\, E\left[\frac{1}{M_0}\right] \end{aligned}$$

because of (4) and because of stationarity. Here, $E\big[M_0^{-1}\big]$ is an inverse moment with $0 < M_0^{-1} \leq 1/\alpha_0$. Altogether, the summands in (8) are always uncorrelated such that we finally obtain:

$$E\big[\widehat{C}_p\big] \;=\; 1, \qquad V\big[\widehat{C}_p\big] \;=\; \frac{1}{T-p}\,\left(1+\frac{H''(1)}{H'(1)}\right)\, E\left[\frac{1}{M_0}\right]. \tag{9}$$

(9) implies that the variance of $\widehat{C}_p$ is inflated by $1+H''(1)/H'(1)$ (compared to the null model with $H''(1) = 0$). But the mean of $\widehat{C}_p$ is always 1, independent of the type of CP-INARCH(p) model.

Therefore, we consider a higher-order extension of the test statistic $\widehat{C}_p$ from (8) such that also its mean is affected if violating H_0. Considering that the r^{th} factorial moment ($r \in \mathbb{N}$) of the Poisson distribution Poi(μ) just equals μ^r (Johnson et al., 2005), it follows that

$$E\big[(X_t)_{(r)} \mid X_{t-1},\ldots\big] \;=\; M_t^r,$$

where $x_{(r)} = x\cdots(x-r+1)$ denotes the falling factorial. So we define

$$\widehat{C}_{p;\,r} \;:=\; \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{(X_t)_{(r)}}{M_t^r} \;=\; \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{X_t(X_t-1)\cdots(X_t-r+1)}{\big(\alpha_0+\sum_{i=1}^{p}\alpha_i X_{t-i}\big)^r}, \tag{10}$$

where $\widehat{C}_p = \widehat{C}_{p;\,1}$. If $(X_t)_{\mathbb{Z}}$ is Poisson INARCH(p) with given parameter values for $\alpha_0,\alpha_1,\ldots,\alpha_p$ (i. e., if H_0 holds), we obtain with analogous computations as for (9) that

$$E\left[\frac{(X_t)_{(r)}}{M_t^r}\right] \;=\; 1, \qquad Cov\left[\frac{(X_t)_{(r)}}{M_t^r},\frac{(X_{t-k})_{(r)}}{M_{t-k}^r}\right] \;=\; 0 \quad \text{for } k \geq 1.$$

To compute the variance, we need the following identity for falling factorials:

$$x_{(r)}^2 \;=\; \textstyle\sum_{k=0}^{r}\binom{r}{k}^2 k!\, x_{(2r-k)}.$$

Then we obtain

$$\begin{aligned}
V\left[\frac{(X_t)_{(r)}}{M_t^r}\right] &= V\left[\frac{E\big[(X_t)_{(r)} \mid X_{t-1},\ldots\big]}{M_t^r}\right] \;+\; E\left[\frac{V\big[(X_t)_{(r)} \mid X_{t-1},\ldots\big]}{M_t^{2r}}\right] \\
&= 0 \;+\; E\left[\frac{E\big[(X_t)_{(r)}^2 \mid X_{t-1},\ldots\big]}{M_t^{2r}} - 1\right] \\
&= \left(\sum_{k=0}^{r} \binom{r}{k}^2 k!\, E\big[M_t^{-k}\big]\right) - 1 \;=\; \sum_{k=1}^{r} \binom{r}{k}^2 k!\, E\big[M_t^{-k}\big]. \qquad (11)
\end{aligned}$$

Overall, under H_0, i. e., for a Poisson INARCH(p) model, we obtain that

$$E\big[\widehat{C}_{p;\,r}\big] \;=\; 1, \qquad V\big[\widehat{C}_{p;\,r}\big] \;=\; \frac{1}{T-p}\sum_{k=1}^{r} \binom{r}{k}^2 k!\, E\big[M_t^{-k}\big]. \qquad (12)$$

The following example considers the case of the alternative H_1.

Example 2 (Second Order Statistic) Let us consider the second order statistic $\widehat{C}_{p;\,2}$, i. e., the case $r = 2$. Under H_0, (12) implies

$$E\big[\widehat{C}_{p;\,2}\big] \;=\; 1, \qquad V\big[\widehat{C}_{p;\,2}\big] \;=\; \frac{1}{T-p}\Big(4\,E\big[M_0^{-1}\big] + 2\,E\big[M_0^{-2}\big]\Big).$$

If, in contrast, the Poisson assumption is violated (H_1), then also the mean becomes sensitive to such a violation. For an underlying CP-INARCH(p) model, we have

$$\begin{aligned}
&E\big[(X_t)_{(2)} \mid X_{t-1},\ldots\big] \\
&\quad = \; V[X_t \mid X_{t-1},\ldots] + E[X_t \mid X_{t-1},\ldots]^2 - E[X_t \mid X_{t-1},\ldots] \\
&\quad \overset{(4)}{=} \; M_t \underbrace{\left(1 + \frac{H''(1)}{H'(1)}\right)}_{=\theta} + M_t^2 - M_t \;=\; M_t^2 + (\theta - 1)\, M_t,
\end{aligned}$$

such that

$$E\big[\widehat{C}_{p;\,2}\big] \;=\; E\left[\frac{(X_t)_{(2)}}{M_t^2}\right] \;=\; 1 + (\theta - 1)\, E\big[M_0^{-1}\big].$$

Therefore, $\widehat{C}_{p;\,2}$ might be a useful statistic to distinguish between H_0 and H_1 in practice.

3 Testing the CP-INARCH's Compounding Structure

In the following, we use the statistic $\widehat{C}_{p;\,r}$ from (10) to test hypotheses (7). Note that its computation requires to specify the parameter values for $\alpha_0, \alpha_1, \ldots, \alpha_p$; so if no such values are available, they need to be estimated from the same time series data that is also used for computing the test statistic. To be able to execute the test (in either scenario), knowledge about the distribution of $\widehat{C}_{p;\,r}$ under the null is required. In Sections 3.1 and 3.2, we show that in the first-order autoregressive case ($p = 1$), even a closed-form analytic solution for $\widehat{C}_{1;\,r}$'s asymptotic distribution can be derived, provided that the asymptotics of the used estimators are available. In this context, we shall alos discuss certain inverse moments of a Poisson INARCH(1) process, see Section 3.3. In general, i. e., for $p > 1$ or for different estimators, a bootstrap implementation is required, which is discussed in Section 3.4.

3.1 Case of Specified Parameters

In Sections 3.1 to 3.3, we concentrate on the case of first-order autoregression, i. e., on the case $p = 1$. According to (7), H_0 assumes the two-parametric *Poisson INARCH(1) model* given by

$$X_t \;\big|\; X_{t-1}, X_{t-2}, \ldots \quad \sim \; \text{Poi}(\alpha_0 + \alpha_1 \cdot X_{t-1}). \tag{13}$$

Though being a rather simple model, it has already found a number of real applications, e. g., to monthly claims counts (Weiß, 2009), to download counts (Zhu & Wang, 2010), to counts of iceberg orders (Jung & Tremayne, 2011), and to monthly strike counts data (Weiß, 2010). A Poisson INARCH(1) process is a stationary, ergodic Markov chain (Ferland et al., 2006; Zhu & Wang, 2011) with simple Poisson probabilities as the transition probabilities. According to Neumann (2011), it is β-mixing (and hence also α-mixing) with exponentially decreasing weights. All moments of a Poisson INARCH(1) process exist (Ferland et al., 2006), and they can be determined according to the recursive scheme provided by Weiß (2009, 2010), see equation (22) below.

For the first-order version of the DINARCH model (2), unconditonal mean and variance are given by (Xu et al., 2012, 4.3)

$$\mu \;=\; \frac{\alpha_0}{1-\alpha_1} \qquad \text{and} \qquad \sigma^2 \;=\; \frac{\theta}{1-\alpha_1^2} \cdot \frac{\alpha_0}{1-\alpha_1}. \tag{14}$$

So θ allows to control the degree of overdispersion independently of α_1.

Let us investigate the distribution of the statistics $\widehat{C}_{1;\,r}$ introduced in the previous Section 2 under H_0, in the case of the Poisson INARCH(1) model with specified parameter values for α_0, α_1. We denote the (inverse) moments

$$q_{k,l} := q_{k,l}(\alpha_0, \alpha_1) \;:=\; E\Big[\frac{X_0^k}{(\alpha_0 + \alpha_1\, X_0)^l}\Big] \qquad \text{for } k, l \geq 0. \tag{15}$$

The moments $q_{k,l}$ from (15) are just the stationary marginal moments for $l = 0$, and for $l > 0$, they are easily computed numerically from the stationary marginal distribution of the Poisson INARCH(1) process $(X_t)_{\mathbb{Z}}$, see Section 3.3 below. The $q_{k,l}$ allow us to rewrite (12) as

$$E\big[\widehat{C}_{1;\,r}\big] \;=\; 1, \qquad V\big[\widehat{C}_{1;\,r}\big] \;=\; \tfrac{1}{T-1}\,\textstyle\sum_{k=1}^{r}\,\binom{r}{k}^2\,k!\,q_{0,k}. \tag{16}$$

As stated above, we know that the null model, the Poisson INARCH(1) model, is α-mixing with exponentially decreasing weights and has existing moments up to any order. So we apply the central limit theorem of Ibragimov (1962) to obtain that the statistics $\widehat{C}_{1;\,r}$ are even asymptotically normally distributed. Hence, one could test the null of a Poisson INARCH(1) model against the alternative of a true CP-INARCH(1) model based on the resulting approximate normal distribution for $\widehat{C}_{1;\,r}$.

These asymptotics, however, only hold for the case of specified H_0 parameters, since these are required to compute the statistics $\widehat{C}_{1;\,r}$. In practice, however, one usually has to estimate these parameters. Plugging-in these estimators into the definition of $\widehat{C}_{1;\,r}$, we obtain a statistic with a different asymptotic distribution than the one mentioned before. So to make the test applicable in practice, further investigations are required.

3.2 Case of Estimated Parameters

To derive an asymptotic approximation to the distribution of $\widehat{C}_{1;\,r}$ under H_0 *but in the presence of estimated parameters*, say, $\hat{\alpha}_0$ and $\hat{\alpha}_1$, we shall look at the first-order Taylor approximation of

$$\widehat{C}_{1;\,r}(\alpha_0,\alpha_1) \;=\; \frac{1}{T-1}\,\sum_{t=2}^{T}\frac{(X_t)_{(r)}}{(\alpha_0+\alpha_1\,X_{t-1})^r},$$

which has the partial derivatives

$$\frac{\partial}{\partial\alpha_0}\,\widehat{C}_{1;\,r} = \frac{1}{T-1}\,\sum_{t=2}^{T}\frac{-r\,(X_t)_{(r)}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}},$$

$$\frac{\partial}{\partial\alpha_1}\,\widehat{C}_{1;\,r} = \frac{1}{T-1}\,\sum_{t=2}^{T}\frac{-r\,(X_t)_{(r)}\,X_{t-1}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}}. \tag{17}$$

By conditioning, it follows that

$$E\Big[\frac{(X_t)_{(r)}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}}\Big] \;=\; q_{0,1}, \qquad E\Big[\frac{(X_t)_{(r)}\,X_{t-1}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}}\Big] \;=\; q_{1,1}, \tag{18}$$

where we used the abbreviation from (15). So we approximate $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$ by

$$\widetilde{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1) \;:=\; \widehat{C}_{1;\,r}(\alpha_0,\alpha_1) \;-\; r\,q_{0,1}\,\big(\hat{\alpha}_0-\alpha_0\big) \;-\; r\,q_{1,1}\,\big(\hat{\alpha}_1-\alpha_1\big), \tag{19}$$

and an approximation of the distribution of $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$ is obtained by deriving the distribution of $\widetilde{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$.

To obtain a closed-form analytic solution, we shall use the usual moment estimators $\hat{\alpha}_0 := \bar{X}\big(1-\hat{\rho}(1)\big)$ and $\hat{\alpha}_1 := \hat{\rho}(1)$, the asymptotic distribution of which is studied in Weiß & Schweer (2016). These estimators are also robust with respect to violating H_0, as they do not rely on a conditional Poisson distribution (which would be the case for maximum likelihood estimators). Using the bias approximations for $\hat{\alpha}_0, \hat{\alpha}_1$ given there, it immediately follows that

$$\begin{aligned} E\big[\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)\big] &\approx E\big[\widehat{C}_{1;\,r}(\alpha_0,\alpha_1)\big] - r\,q_{0,1}\,E\big[\hat{\alpha}_0-\alpha_0\big] - r\,q_{1,1}\,E\big[\hat{\alpha}_1-\alpha_1\big] \\ &\approx 1 \;-\; r\,\frac{q_{0,1}}{T-1}\,\left(\frac{1+3\alpha_1}{1-\alpha_1}\,\alpha_0 + \frac{2\alpha_1^2(1+2\alpha_1^2)}{1-\alpha_1^3}\right) \\ &\quad + r\,\frac{q_{1,1}}{T-1}\,\left(1+3\alpha_1+\frac{\alpha_1}{\alpha_0}\left(1+\frac{2\alpha_1(1+2\alpha_1^2)}{1+\alpha_1+\alpha_1^2}\right)\right). \end{aligned} \tag{20}$$

The derivation of the asymptotic variance of the approximate quantity (19), however, is more demanding, see Appendix A.1 for the details. We finally obtain the approximate variance $\sigma_{1;\,r}^2/(T-1)$ with

$$\begin{aligned} \sigma_{1;\,r}^2 = \sum_{k=1}^{r} \binom{r}{k}^2 k!\,q_{0,k} \;+\; r^2\,q_{0,1}^2\,\frac{\alpha_0}{1-\alpha_1}\Big(\alpha_0(1+\alpha_1)+\frac{1+2\alpha_1^4}{1+\alpha_1+\alpha_1^2}\Big) \\ -\,2r^2\,q_{0,1} \;+\; r^2\,q_{1,1}^2\,(1-\alpha_1^2)\Big(1+\frac{\alpha_1(1+2\alpha_1^2)}{\alpha_0(1+\alpha_1+\alpha_1^2)}\Big) \\ -\,2r^2\,q_{0,1}q_{1,1}\,\Big(\alpha_0(1+\alpha_1)+\frac{(1+2\alpha_1)\alpha_1^3}{1+\alpha_1+\alpha_1^2}\Big). \end{aligned} \tag{21}$$

So the test statistics $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$ can now be applied in practice by choosing the critical values from a normal distribution with mean and variance given according to (20) and (21), respectively.

3.3 Inverse Moments

Before investigating the finite-sample performance of the proposed test, some background on the (numerical) computation of the Poisson INARCH(1)'s inverse moments is required. Equation (15) defines the moments

$$q_{k,l} \;=\; E\Big[\frac{X_0^k}{(\alpha_0+\alpha_1\,X_0)^l}\Big] \qquad \text{for } k,l\geq 0,$$

which are just the stationary marginal moments μ_k for $l=0$. These can be computed exactly in two steps. First, the marginal cumulants κ_k are calculated according to the scheme provided by Weiß (2009, 2010). Denoting the Stirling numbers of the first kind (Douglas, 1980, Appendix 9.1) by $s_{k,j}$, it holds that

$$\kappa_1 \;=\; \tfrac{\alpha_0}{1-\alpha_1}, \qquad \kappa_k \;=\; -(1-\alpha_1^k)^{-1}\cdot \textstyle\sum_{j=1}^{k-1} s_{k,j}\cdot\kappa_j \quad \text{for } k\geq 2. \tag{22}$$

In the second step, these cumulants are transformed into the moments μ_k via (Smith, 1995)

$$\mu_k \;=\; \textstyle\sum_{j=0}^{k-1} \binom{k-1}{j}\,\kappa_{k-j}\,\mu_j \qquad \text{for } k\geq 1. \tag{23}$$

So it remains to consider the case $l>0$. Applying the binomial sum formula to $X_0^k \;=\; \alpha_1^{-k}\,\big((\alpha_0+\alpha_1 X_0)-\alpha_0\big)^k$, we obtain

$$q_{k,l} = \sum_{j=0}^{k} \binom{k}{j}\,(-1)^{k-j}\,\frac{\alpha_0^{k-j}}{\alpha_1^k}\cdot E\big[(\alpha_0+\alpha_1\,X_0)^{j-l}\big], \tag{24}$$

$$\text{where}\quad E\big[(\alpha_0+\alpha_1\,X_0)^{j-l}\big] \;=\; \begin{cases} q_{0,l-j} & \text{if } j<l,\\ \sum_{i=0}^{j-l}\binom{j-l}{i}\,\alpha_0^{j-l-i}\alpha_1^i\,\mu_i & \text{if } j\geq l,\end{cases}$$

where the last expression again follows from the binomial sum formula. So equation (24) implies that $q_{k,l}$ can be traced back to either the usual moments μ_k or to purely inverse moments of the form $q_{0,l}$. So it suffices to discuss how to obtain the $q_{0,l} \;=\; E\big[(\alpha_0+\alpha_1\,X_0)^{-l}\big]$ for $l\geq 1$, the value of which is obviously bounded by $0<q_{0,l}<\alpha_0^{-l}$.

If only being interested in the numerical computation of $q_{0,l}$ (as required for applying the proposed $\widehat{C}_{1;\,r}$-test), the Markov chain approximation (Weiß, 2010) can be used: we compute the Poisson INARCH(1)'s transition probabilities

$$p_{r|s} := P(X_t=r \mid X_{t-1}=s) \;=\; \exp\left(-\alpha_0-\alpha_1\,s\right)(\alpha_0+\alpha_1\,s)^r\,/\,r!$$

for all $0\leq r,s\leq M$ (with M sufficiently large), define the matrix $\mathbf{P}_M := (p_{r|s})_{r,s=0,\ldots,M}$, and numerically solve the eigenvalue problem $\mathbf{P}_M\,\boldsymbol{p}=\boldsymbol{p}$ (invariance equation) in $\boldsymbol{p}$. The normalized eigenvector $\boldsymbol{p}$ (i. e., with non-negative entries summing up to one) is used as an approximation for the marginal probabilities $\big(P(X_t=0),\ldots,P(X_t=M)\big)^\top$, and $q_{0,l}$ is approximated by the sum

$$q_{0,l} \;\approx\; \sum_{r=0}^{M} \frac{1}{(\alpha_0+\alpha_1\,r)^l}\cdot p_r. \tag{25}$$

The calculation of $q_{0,l}$ may also be performed following the method provided in Adell et al. (1996) to calculate negative moments of nonnegative random variables, and taking into account that the distribution of X_t given all the past is Poisson with mean $M_t=\alpha_0+\alpha_1 X_{t-1}$. Let us begin by stating a result relating the radius of convergence of the moment generating function of M_1 with the values of the coefficient α_1.

Lemma 1 *If the moment generating function of* M_1, $\mathrm{mgf}_{M_1}(u)=E\big[\exp(uM_1)\big]$, *is defined for every* $u\in(u_1;u_2)$, *where* $u_1<0<u_2$ *with* $\min\{-u_1,u_2\}=b$, *then* $\alpha_1<\frac{\ln(b+1)}{u}$ *for all* $0<u<b$.

Proof For $u \in (-b; b)$, we have

$$\begin{aligned} \mathrm{mgf}_{M_1}(u) = E\big[\exp(uM_1)\big] \;&=\; E\Big[E\big[\exp\big(u\,(\alpha_0+\alpha_1 X_0)\big)\mid X_{-1}\big]\Big] \\ &= \exp(u\alpha_0)\,E\big[\exp\big(M_0\,(\exp(u\alpha_1)-1)\big)\big] \\ &= \exp(u\alpha_0)\,\mathrm{mgf}_{M_0}\big(\exp(u\alpha_1)-1\big). \end{aligned}$$

Then

$$-b < \exp(u\alpha_1) - 1 < b \quad \Leftrightarrow \quad -\infty < u\alpha_1 < \ln(b+1),$$

and for all $0 < u < b$, we obtain that $\alpha_1 < \frac{\ln(b+1)}{u}$.

To find $q_{0,l} = E[M_t^{-l}]$ for $l \geq 1$, with $M_t = \alpha_0 + \alpha_1 X_{t-1} = E[X_t \mid X_{t-1}]$, we note that

$$\begin{aligned} q_{0,l} = E\left[\frac{1}{M_t^l}\right] &= \frac{1}{\alpha_0^l}\,E\left[E\left[\left(\frac{\frac{\alpha_0}{\alpha_1}}{\frac{\alpha_0}{\alpha_1}+X_{t-1}}\right)^l \mid X_{t-2}\right]\right] \\ &= \frac{1}{\alpha_1^l}\,E\left[\sum_{n=0}^{+\infty}\frac{(-1)^n\,M_{t-1}^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l}\right]. \end{aligned}$$

Let us now consider that the moment generating function of M_1, $\mathrm{mgf}_{M_1}(u) = E\big[\exp(uM_1)\big]$, is defined for every $u \in (u_1; u_2)$ where $u_1 < 0 < u_2$ such that the radius of convergence satisfies $\min\{-u_1, u_2\} = b > 2$ (also see Lemma 1). With these conditions, we have (see Appendix A.2)

$$q_{0,l} = E\left[\frac{1}{M_t^l}\right] \;=\; \frac{1}{\alpha_1^l}\sum_{n=0}^{+\infty}\frac{(-1)^n}{n!}\,E[M_{t-1}^n]\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l}, \tag{26}$$

that is, the change between the expectation and the infinite sum is allowed. So according to the previous Lemma 1, $\alpha_1 < \frac{\ln(b+1)}{u}$ for all $0 < u < b$. Thus, if $\alpha_1 \geq \frac{\ln(b+1)}{u} > \frac{\ln(b+1)}{b}$ with $b > 2$, the equality (26) may not be true. In the Figure 1, we plot b according to the equation $\ln(b+1)/b = \alpha_1$ (lower bound for the radius of convergence) against α_1. This value b decreases with increasing α_1 and falls below 2 for $\alpha_1 = \ln(3)/2 \approx 0.549$; the dashed line refers to the above condition $b > 2$.

In Table 1, we present the values for $q_{0,l}$ with $l = 1, 2, 3, 4$ obtained with the two approaches (25) and (26). In the latter case, the summation in n was stopped if the difference between successive summands felt below 10^{-8}, or if 100 summands were reached. In the left block, the marginal mean is 2.5, in the right, it is 5.0. We note the non-convergence of the approach (26) only for $\alpha_1 > 0.6 > \frac{\ln(3)}{2}$.

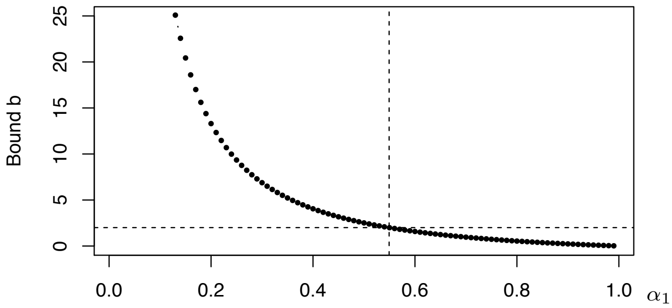

Fig. 1 Solution b of equation $\ln(b+1)/b = \alpha_1$ against α_1.

α_0	α_1	l	$q_{0,l}$ by (25)	n. s.	$q_{0,l}$ by (26)	α_0	α_1	l	$q_{0,l}$ by (25)	n. s.	$q_{0,l}$ by (26)
2	0.2	1	0.4064081	11	0.4064081	4	0.2	1	0.2016350	11	0.2016350
		2	0.1676993	11	0.1676993			2	0.0409823	12	0.0409823
		3	0.0702093	12	0.0702093			3	0.0083949	12	0.0083949
		4	0.0298009	12	0.0298009			4	0.0017328	12	0.0017328
1.5	0.4	1	0.4299554	18	0.4299554	3	0.4	1	0.2075920	20	0.2075920
		2	0.1980567	19	0.1980567			2	0.0447126	21	0.0447126
		3	0.0972296	20	0.0972296			3	0.0099853	22	0.0099853
		4	0.0505194	20	0.0505194			4	0.0023098	22	0.0023098
1	0.6	1	0.4973967	48	0.4973967	2	0.6	1	0.2238847	56	0.2238847
		2	0.3046319	52	0.3046320			2	0.0563205	60	0.0563205
		3	0.2212899	55	0.2212899			3	0.0159104	62	0.0159104
		4	0.1815225	57	0.1815225			4	0.0050165	64	0.0050165
0.5	0.8	1	0.8060558	100	$2.247 \cdot 10^{26}$	1	0.8	1	0.2940770	100	$2.091 \cdot 10^{27}$
		2	1.1167550	100	$1.706 \cdot 10^{27}$			2	0.1322086	100	$1.269 \cdot 10^{28}$
		3	1.9693380	100	$7.069 \cdot 10^{27}$			3	0.0845151	100	$4.043 \cdot 10^{28}$
		4	3.7735840	100	$2.155 \cdot 10^{28}$			4	0.0672318	100	$9.050 \cdot 10^{28}$

Table 1 Approximations for $q_{0,l}$ with $l = 1,2,3,4$ with approaches (25) and (26), where "n. s." is the number of summands used for (26).

3.4 Bootstrap Implementation

In cases where a closed-form analytic solution for the test statistic's asymptotic distribution is not available, it is recommended to use a parametric bootstrap implementation. So more generally than in Sections 3.1 to 3.3, let us consider the p^{th}-order autoregressive case together with the statistic $\widehat{C}_{p;\,r}$ from (10) (and appropriate parameter estimators) to test the hypotheses (7). Let B denote the number of bootstrap replications.

Solution 1 Let $x_1, \ldots, x_T$ be the available time series stemming from an INARCH(p) process.

1. Assuming that the null holds, i. e., that the data generating mechanism is a *Poisson* INARCH(p) process, estimate the parameters $\alpha_0, \alpha_1, \ldots, \alpha_p$ as $\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_p$.
   Compute the test statistic $\widehat{C}_{p;\,r}(\hat{\alpha}_0, \ldots, \hat{\alpha}_p)$.
2. Using a Poisson INARCH(p) model with parameter values $\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_p$ as the data generating mechanism,

   - generate B bootstrap replicates $x^*_{b,1},\ldots,x^*_{b,T}$ of the time series, $b = 1,\ldots,B$,
   - and compute the respective parameter estimates $\hat{\alpha}^*_{b,0},\ldots,\hat{\alpha}^*_{b,p}$ as well as test statistics $\widehat{C}^*_{b;\,p;\,r}(\hat{\alpha}^*_{b,0},\ldots,\hat{\alpha}^*_{b,p})$.
3. Determine the critical value(s) from $\widehat{C}^*_{1;\,p;\,r},\ldots,\widehat{C}^*_{B;\,p;\,r}$, e. g., as the $1-\alpha$-quantile in the case of an upper-sided test.
   Apply this decision rule to the test statistic $\widehat{C}_{p;\,r}(\hat{\alpha}_0,\ldots,\hat{\alpha}_p)$ computed in Step 1.

In the subsequent simulation study, we shall investigate the finite-sample performance of both the asymptotic implementation and the bootstrap implementation of our proposed test.

4 Simulation Study and Data Application

To analyze the quality of the approximate distribution (20), (21) of the statistics $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$ as well as the finite-sample performance of the proposed tests, a simulation study has been done with 10 000 replications per scenario. The results are discussed in Sections 4.1 and 4.2. The real-data example presented in Section 4.3 exemplifies the application of the test in practice.

4.1 Performance of Asymptotic Approximation

We first analyze the quality of the approximate distribution (20), (21) of the statistics $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$. The results shown in Table 2 refer to simulated Poisson INARCH(1) processes (13) (upper half: $\mu = 2.5$; lower half: $\mu = 5.0$). They show mean and standard deviation as computed according to the approximate formulae (20), (21), and compare these values with the corresponding sample counterparts obtained from simulations. The simulated means are below the theoretical value $C_{1;\,r} = 1$ under the null, but the approximate formula (20) accounts for the negative bias to some degree. The approximation (21) of the standard deviation works rather well especially for the second-order statistic $\widehat{C}_{1;\,2}$; for higher orders $r = 3, 4$, the quality of approximation deteriorates with increasing α.

4.2 Performance of $\widehat{C}_{p;\,r}$-Test

The most important criterion for the practitioner are the true rejection rates (size, power) if applying the proposed test. First, let us continue with the asymptotic implementation if using $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$ as a test statistic. From each simulated time series, upper-sided tests on the nominal level 5 % were designed and executed: the null was rejected if $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$ exceeds the critical value $\hat{\mu}_{1;\,r} + z_{0.95}\,\hat{\sigma}_{1;\,r}$, where $z_{0.95}$ denotes the 95 %-quantile of the standard normal

			$E[\widehat{C}_{1;\,2}(\cdot)]$		$\sqrt{V}[\widehat{C}_{1;\,2}(\cdot)]$		$E[\widehat{C}_{1;\,3}(\cdot)]$		$\sqrt{V}[\widehat{C}_{1;\,3}(\cdot)]$		$E[\widehat{C}_{1;\,4}(\cdot)]$		$\sqrt{V}[\widehat{C}_{1;\,4}(\cdot)]$	
α_0	α_1	T	appr	simul	appr	simul	appr	simul	appr	simul	appr	simul	appr	simul
2	0.2	100	0.999	0.992	0.058	0.060	0.999	0.976	0.186	0.184	0.998	0.953	0.444	0.426
		250	1.000	0.997	0.037	0.037	1.000	0.993	0.118	0.115	0.999	0.985	0.280	0.273
		500	1.000	0.999	0.026	0.026	1.000	0.995	0.083	0.082	1.000	0.991	0.198	0.195
		1000	1.000	0.999	0.018	0.018	1.000	0.997	0.059	0.058	1.000	0.995	0.140	0.139
1.5	0.4	100	0.996	0.990	0.064	0.066	0.994	0.981	0.206	0.207	0.992	0.964	0.501	0.504
		250	0.998	0.997	0.041	0.041	0.997	0.992	0.130	0.130	0.997	0.985	0.316	0.313
		500	0.999	0.998	0.029	0.029	0.999	0.995	0.092	0.093	0.998	0.990	0.223	0.225
		1000	1.000	0.999	0.020	0.020	0.999	0.998	0.065	0.066	0.999	0.998	0.158	0.161
1	0.6	100	0.982	0.987	0.088	0.091	0.973	0.972	0.269	0.277	0.964	0.946	0.698	0.685
		250	0.993	0.996	0.056	0.057	0.989	0.994	0.170	0.174	0.986	0.994	0.440	0.464
		500	0.996	0.998	0.039	0.040	0.995	0.994	0.120	0.120	0.993	0.989	0.311	0.310
		1000	0.998	0.999	0.028	0.028	0.997	0.997	0.085	0.086	0.996	0.996	0.220	0.234
0.5	0.8	100	0.876	0.958	0.219	0.180	0.813	0.942	0.616	0.840	0.751	0.926	1.933	3.401
		250	0.951	0.984	0.138	0.136	0.926	0.977	0.388	0.418	0.901	0.967	1.219	1.472
		500	0.975	0.992	0.098	0.098	0.963	0.995	0.274	0.320	0.951	1.010	0.861	1.311
		1000	0.988	0.995	0.069	0.069	0.982	0.994	0.194	0.197	0.975	0.996	0.609	0.643
4	0.2	100	1.000	0.995	0.029	0.030	0.999	0.988	0.089	0.088	0.999	0.977	0.196	0.192
		250	1.000	0.999	0.018	0.019	1.000	0.997	0.056	0.057	1.000	0.994	0.124	0.124
		500	1.000	0.999	0.013	0.013	1.000	0.998	0.040	0.040	1.000	0.996	0.087	0.086
		1000	1.000	1.000	0.009	0.009	1.000	0.998	0.028	0.028	1.000	0.996	0.062	0.062
3	0.4	100	0.998	0.996	0.030	0.032	0.997	0.986	0.094	0.093	0.996	0.973	0.207	0.199
		250	0.999	0.998	0.019	0.020	0.999	0.995	0.059	0.059	0.999	0.990	0.131	0.128
		500	1.000	0.999	0.014	0.014	0.999	0.997	0.042	0.042	0.999	0.995	0.092	0.092
		1000	1.000	1.000	0.010	0.010	1.000	0.999	0.030	0.029	1.000	0.998	0.065	0.064
2	0.6	100	0.993	0.994	0.037	0.039	0.990	0.987	0.108	0.114	0.986	0.977	0.242	0.276
		250	0.997	0.997	0.023	0.023	0.996	0.995	0.068	0.068	0.994	0.991	0.152	0.151
		500	0.999	0.998	0.016	0.016	0.998	0.997	0.048	0.048	0.997	0.995	0.108	0.108
		1000	0.999	0.999	0.012	0.012	0.999	0.998	0.034	0.034	0.999	0.998	0.076	0.076
1	0.8	100	0.960	0.984	0.077	0.070	0.940	0.978	0.191	0.215	0.919	0.971	0.455	0.531
		250	0.984	0.994	0.048	0.047	0.976	0.992	0.120	0.154	0.968	0.997	0.287	0.772
		500	0.992	0.996	0.034	0.034	0.988	0.995	0.085	0.089	0.984	0.996	0.203	0.238
		1000	0.996	0.999	0.024	0.025	0.994	0.998	0.060	0.063	0.992	0.999	0.143	0.169

Table 2 Mean and standard deviation of $\widehat{C}_{1;\,r}(\hat{\alpha}_0, \hat{\alpha}_1)$: approximation by (20), (21) vs. simulated values.

distribution, and where $\hat{\mu}_{1;\,r}, \hat{\sigma}_{1;\,r}$ were computed according to (20), (21) by plugging-in the obtained parameter estimates. The fraction of rejections among the replications was computed for each scenario, which expresses the empirical size under the null of the Poi-INARCH(1) model, and the empirical power otherwise. As the alternative model, the NB-INARCH(1) model by Xu et al. (2012) with different levels of the dispersion parameter $\theta > 1$ was used, see (5) in Example 1, i. e., the counts were generated according to the recursive scheme

$$X_t \;\big|\; X_{t-1}, X_{t-2}, \ldots \quad \sim \; \text{NB}\left(\frac{\alpha_0 + \alpha_1\, X_{t-1}}{\theta - 1},\ \frac{1}{\theta}\right) \quad \text{with } \theta > 1. \tag{27}$$

The obtained results are summarized in Table 3.

If we look at the size values (highlighted in gray) in Table 3, we see that the empirical size usually agrees quite well with the nominal level 0.05. An exception is the fourth-order statistic for large α and small T, where the empirical size values are visibly smaller than 0.05. So up to now, there is not much difference between the orders $r = 2, 3, 4$ under the null (except for large α). Hence, the crucial question is about the power of these tests with respect to the alternative (27). From Table 3, it can be seen that the power values quickly increase with increasing T, and the power is generally better for lower values of the dependence parameter α_1. It can also be seen that the respective power

			$\widehat{C}_{1;2}(\hat\alpha_0,\hat\alpha_1)$; $\theta =$				$\widehat{C}_{1;3}(\hat\alpha_0,\hat\alpha_1)$; $\theta =$				$\widehat{C}_{1;4}(\hat\alpha_0,\hat\alpha_1)$; $\theta =$			
α_0	α_1	T	1	1.2	1.4	1.6	1	1.2	1.4	1.6	1	1.2	1.4	1.6
2	0.2	100	0.051	0.354	0.720	0.901	0.051	0.328	0.667	0.874	0.051	0.272	0.561	0.786
		250	0.049	0.636	0.966	0.999	0.053	0.581	0.947	0.997	0.056	0.478	0.878	0.985
		500	0.051	0.874	0.999	1.000	0.052	0.829	0.998	1.000	0.057	0.717	0.988	1.000
		1000	0.049	0.989	1.000	1.000	0.055	0.975	1.000	1.000	0.056	0.927	1.000	1.000
1.5	0.4	100	0.053	0.337	0.691	0.893	0.054	0.305	0.644	0.855	0.049	0.244	0.532	0.755
		250	0.053	0.608	0.956	0.999	0.055	0.561	0.929	0.996	0.053	0.448	0.848	0.979
		500	0.051	0.848	0.999	1.000	0.058	0.805	0.997	1.000	0.060	0.680	0.983	1.000
		1000	0.052	0.984	1.000	1.000	0.060	0.969	1.000	1.000	0.061	0.905	1.000	1.000
1	0.6	100	0.063	0.305	0.617	0.830	0.050	0.266	0.579	0.805	0.036	0.193	0.448	0.668
		250	0.061	0.522	0.910	0.991	0.060	0.487	0.888	0.987	0.047	0.353	0.760	0.942
		500	0.059	0.748	0.993	1.000	0.057	0.722	0.989	1.000	0.048	0.555	0.944	0.997
		1000	0.056	0.944	1.000	1.000	0.060	0.932	1.000	1.000	0.058	0.803	0.999	1.000
0.5	0.8	100	0.054	0.206	0.443	0.642	0.040	0.189	0.412	0.617	0.019	0.111	0.277	0.456
		250	0.056	0.336	0.695	0.891	0.052	0.325	0.696	0.894	0.025	0.186	0.493	0.745
		500	0.057	0.522	0.891	0.985	0.062	0.503	0.896	0.989	0.032	0.294	0.696	0.923
		1000	0.056	0.729	0.990	1.000	0.061	0.735	0.993	1.000	0.042	0.457	0.917	0.996
4	0.2	100	0.049	0.362	0.737	0.918	0.048	0.344	0.707	0.905	0.047	0.307	0.647	0.866
		250	0.053	0.647	0.972	0.999	0.053	0.621	0.964	0.998	0.057	0.557	0.934	0.995
		500	0.050	0.886	1.000	1.000	0.052	0.858	0.999	1.000	0.053	0.801	0.997	1.000
		1000	0.048	0.990	1.000	1.000	0.050	0.984	1.000	1.000	0.054	0.967	1.000	1.000
3	0.4	100	0.059	0.358	0.712	0.914	0.050	0.338	0.693	0.901	0.049	0.296	0.627	0.850
		250	0.056	0.632	0.968	0.999	0.051	0.610	0.960	0.999	0.054	0.539	0.927	0.995
		500	0.053	0.867	0.999	1.000	0.049	0.846	0.999	1.000	0.055	0.788	0.997	1.000
		1000	0.053	0.989	1.000	1.000	0.050	0.981	1.000	1.000	0.053	0.960	1.000	1.000
2	0.6	100	0.067	0.334	0.661	0.870	0.054	0.317	0.657	0.867	0.048	0.269	0.574	0.811
		250	0.056	0.568	0.942	0.996	0.058	0.565	0.942	0.997	0.053	0.488	0.897	0.991
		500	0.051	0.813	0.998	1.000	0.056	0.809	0.998	1.000	0.058	0.730	0.992	1.000
		1000	0.058	0.969	1.000	1.000	0.056	0.973	1.000	1.000	0.057	0.941	1.000	1.000
1	0.8	100	0.065	0.239	0.486	0.708	0.054	0.242	0.520	0.760	0.037	0.187	0.421	0.667
		250	0.062	0.364	0.747	0.930	0.058	0.410	0.826	0.969	0.043	0.318	0.724	0.927
		500	0.056	0.540	0.934	0.996	0.059	0.634	0.974	0.999	0.048	0.508	0.927	0.996
		1000	0.060	0.763	0.995	1.000	0.063	0.863	1.000	1.000	0.055	0.745	0.997	1.000

Table 3 Simulated rejection rates for upper-sided test $\widehat{C}_{1;r}(\hat\alpha_0,\hat\alpha_1)$, nominal level 5 %, under H_0: Poi-INARCH(1) model ($\theta = 1$), and H_1: NB-INARCH(1) model ($\theta > 1$).

values are larger in the lower half of the table, where we have a larger marginal mean. Comparing the power among the different orders $r = 2, 3, 4$, Table 3 shows a rather clear picture. The fourth-order test is always worse than the second-order test, and with very few exceptions ($\alpha_0 = 1$, $\alpha_1 = 0.8$), the same conclusion also holds between the third- and second-order test. This desirable increase in the rejection rates with increasing θ is caused by increases in both the mean and the standard deviation of $\widehat{C}_{1;r}(\hat\alpha_0,\hat\alpha_1)$ (the actual values are omitted in Table 3). Taking these power results together with the described properties under the null, it appears to be preferable to use the second-order test $\widehat{C}_{1;2}(\hat\alpha_0,\hat\alpha_1)$ in practice.

Next, we investigate the bootstrap implementation of the test, as described in Section 3.4. This approach is computationally much more demanding than the above asymptotic implementation. So to be able to still manage 10 000 Monte-Carlo replicates, we used the warp-speed method by Giacomini et al. (2013) to perform the simulation experiments.

As a first experiment, we considered again the first-order autoregressive case together with moment estimators, but now using the bootstrap implementation. So while the critical value was computed before as $\hat\mu_{1;r} + z_{0.95}\,\hat\sigma_{1;r}$ by utilizing normality, we now compute it as the 95 %-sample quantile from the

		$\widehat{C}_{1;2}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$				$\widehat{C}_{1;3}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$				$\widehat{C}_{1;4}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$			
α_0 α_1	T	1	1.2	1.4	1.6	1	1.2	1.4	1.6	1	1.2	1.4	1.6
1.5 0.4	100	0.047	0.308	0.660	0.883	0.046	0.292	0.630	0.846	0.046	0.250	0.541	0.753
	250	0.050	0.579	0.959	0.998	0.050	0.534	0.927	0.994	0.051	0.424	0.846	0.976
	500	0.049	0.842	0.999	1.000	0.047	0.794	0.998	1.000	0.045	0.660	0.980	1.000
	1000	0.045	0.982	1.000	1.000	0.044	0.965	1.000	1.000	0.040	0.891	1.000	1.000
0.5 0.8	100	0.063	0.192	0.371	0.540	0.062	0.206	0.414	0.613	0.048	0.201	0.408	0.606
	250	0.063	0.296	0.607	0.827	0.057	0.302	0.649	0.854	0.055	0.279	0.610	0.824
	500	0.058	0.394	0.808	0.963	0.064	0.415	0.847	0.979	0.062	0.352	0.766	0.954
	1000	0.052	0.615	0.969	1.000	0.048	0.641	0.981	1.000	0.048	0.517	0.918	0.998

Table 4 Simulated rejection rates for bootstrap implementation of upper-sided test $\widehat{C}_{1;r}(\hat{\alpha}_0, \hat{\alpha}_1)$, nominal level 5 %, under H_0: Poi-INARCH(1) model ($\theta = 1$), and H_1: NB-INARCH(1) model ($\theta > 1$).

		$\widehat{C}_{1;2}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$				$\widehat{C}_{1;3}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$				$\widehat{C}_{1;4}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$			
α_0 α_1	T	1	1.2	1.4	1.6	1	1.2	1.4	1.6	1	1.2	1.4	1.6
1.5 0.4	100	0.051	0.320	0.708	0.904	0.051	0.290	0.644	0.851	0.049	0.248	0.559	0.763
	250	0.052	0.601	0.959	0.999	0.051	0.538	0.918	0.996	0.052	0.429	0.841	0.978
	500	0.060	0.857	0.999	1.000	0.053	0.797	0.997	1.000	0.051	0.660	0.982	1.000
	1000	0.049	0.986	1.000	1.000	0.047	0.970	1.000	1.000	0.049	0.900	1.000	1.000
0.5 0.8	100	0.047	0.216	0.473	0.712	0.051	0.214	0.445	0.661	0.053	0.213	0.419	0.616
	250	0.049	0.401	0.820	0.969	0.051	0.327	0.687	0.912	0.050	0.290	0.609	0.836
	500	0.047	0.659	0.983	1.000	0.044	0.495	0.917	0.992	0.042	0.379	0.823	0.962
	1000	0.048	0.895	1.000	1.000	0.048	0.739	0.995	1.000	0.048	0.542	0.954	0.998

Table 5 Simulated rejection rates for bootstrap implementation of upper-sided test $\widehat{C}_{1;r}(\hat{\alpha}_{0;\mathrm{ML}}, \hat{\alpha}_{1;\mathrm{ML}})$, nominal level 5 %, under H_0: Poi-INARCH(1) model ($\theta = 1$), and H_1: NB-INARCH(1) model ($\theta > 1$).

bootstrap replicates of the test statistic. Results are shown in Table 4; to save some space, we now only display the results of $\mu = 2.5$ and $\alpha_1 = 0.4, 0.8$. Comparing with the respective rejection rates in in Table 3 for the asymptotic implementation, we see evry similar results for $\alpha_1 = 0.4$. For the large $\alpha_1 = 0.8$, we see slightly increased sizes (remember that the critical values are computed differently), which is welcome only for the fourth-order statistic. But altogether, both implementations work similarly well.

For the remaining scenarios to be discussed, an asymptotic implementation is not available, so the bootstrap implementation is the only option. We start by still considering the first-order autoregressive case, but now together with conditional maximim likelihood (ML) estimation. Comparing the results of Table 5 with those of Tables 3 and 4, we see a similar performance for $\alpha_1 = 0.4$, but an even improved performance for the large $\alpha_1 = 0.8$. In particular, while the conclusions were a bit diffuse with respect to Table 4, if using ML estimation, the use of the second-order statistic ($r = 2$) is preferable throughout.

At this point, let us discuss the problem of extreme autocorrelation in some more detail. We have seen that increasing autocorrelation level α_1 leeds to reduced power and sometimes even to worse size values (moment estimation, especially $r > 2$). Although not being a particularly realistic scenario, let us increase α_1 even beyond 0.8. The size values in Table 6 show that the higher-order statistics ($r = 3, 4$) together with moment estimation are even stronger affected if $\alpha_1 = 0.9$, whereas the second-order statistic or ML-based statistics still work well for such a strong dependence level.

			$\widehat{C}_{1;\,2}(\hat{\alpha}_0, \hat{\alpha}_1)$			$\widehat{C}_{1;\,3}(\hat{\alpha}_0, \hat{\alpha}_1)$			$\widehat{C}_{1;\,4}(\hat{\alpha}_0, \hat{\alpha}_1)$		
			MMa	MMb	MLb	MMa	MMb	MLb	MMa	MMb	MLb
0.5	0.8	100	0.054	0.063	0.047	0.040	0.062	0.051	0.019	0.048	0.053
		250	0.056	0.063	0.049	0.052	0.057	0.051	0.025	0.055	0.050
		500	0.057	0.058	0.047	0.062	0.064	0.044	0.032	0.062	0.042
		1000	0.056	0.052	0.048	0.061	0.048	0.048	0.042	0.048	0.048
0.25	0.9	100	0.057	0.043	0.043	0.029	0.041	0.052	0.006	0.033	0.046
		250	0.058	0.058	0.047	0.042	0.057	0.052	0.011	0.029	0.047
		500	0.053	0.061	0.050	0.046	0.058	0.046	0.021	0.035	0.050
		1000	0.053	0.055	0.047	0.057	0.054	0.053	0.038	0.049	0.054

Table 6 Simulated sizes for upper-sided test $\widehat{C}_{1;\,r}(\hat{\alpha}_0, \hat{\alpha}_1)$, nominal level 5 %, if using moment estimation together with asymptotic implementation ("MMa") or bootstrap implementation ("MMb"), or if using ML estimation together with bootstrap implementation ("MLb").

				$\widehat{C}_{1;\,2}$	$\widehat{C}_{2;\,2}(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)$; $\theta =$				$\widehat{C}_{1;\,3}$	$\widehat{C}_{2;\,3}(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)$; $\theta =$				$\widehat{C}_{1;\,4}$	$\widehat{C}_{2;\,4}(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)$; $\theta =$			
α_0	α_1	α_2	T	1	1	1.2	1.4	1.6	1	1	1.2	1.4	1.6	1	1	1.2	1.4	1.6
1.125	0.3	0.25	100	0.105	0.047	0.309	0.663	0.889	0.105	0.052	0.270	0.582	0.832	0.095	0.052	0.237	0.503	0.745
			250	0.164	0.050	0.589	0.956	0.999	0.152	0.049	0.522	0.909	0.993	0.131	0.049	0.410	0.813	0.972
			500	0.264	0.048	0.847	1.000	1.000	0.246	0.056	0.772	0.995	1.000	0.197	0.053	0.632	0.970	1.000
			1000	0.416	0.048	0.982	1.000	1.000	0.372	0.053	0.959	1.000	1.000	0.294	0.049	0.867	0.999	1.000
0.375	0.6	0.25	100	0.093	0.047	0.201	0.423	0.639	0.098	0.053	0.221	0.446	0.610	0.092	0.047	0.214	0.419	0.579
			250	0.158	0.050	0.377	0.772	0.944	0.147	0.051	0.307	0.647	0.876	0.143	0.053	0.282	0.596	0.819
			500	0.235	0.050	0.604	0.966	0.999	0.197	0.050	0.460	0.874	0.984	0.170	0.050	0.352	0.750	0.936
			1000	0.389	0.050	0.865	1.000	1.000	0.309	0.053	0.676	0.986	1.000	0.236	0.048	0.485	0.892	0.991

Table 7 Simulated rejection rates for bootstrap implementation of upper-sided test $\widehat{C}_{1;\,r}(\hat{\alpha}_{0;\,\mathrm{ML}}, \hat{\alpha}_{1;\,\mathrm{ML}})$, nominal level 5 %, under H_0: Poi-INARCH(2) model ($\theta = 1$), and H_1: NB-INARCH(1) model ($\theta > 1$). Columns $\widehat{C}_{1;\,r}$ present sizes if falsely assuming Poi-INARCH(1) model.

Finally, let us consider higher-order autoregressions. More precisely, we consider a bootstrap implementation of the INARCH(2) case together with ML estimation, see the results in Table 7. The model parametrizations are chosen such that still $\mu = 2.5$ and $\rho(1) = 0.4, 0.8$, but $\alpha_2 = 0.25 > 0$. First, we check the effect of a model misspecification: although being concerned with second-order autoregression, a Poisson INARCH(1) model is fitted to the simulated data and statistic $\widehat{C}_{1;\,r}$ is computed. Table 7 shows that such a model misspecification leads to strongly increased sizes (especially for large T), so it is important to carefully identify the correct model order (which, in turn, is also more reliably done for large T). If correctly assuming an INARCH(2) model, size and power values are similar to those in Table 5, although the power is slightly better in the latter case. Table 7 again indicates that using the second-order statistic is to be recommended for practice.

4.3 Real-Data Example

Let us conclude our empirical investigations with a real-data example. For this purpose, we consider the earthquakes counts discussed in Section 5.1 of Zhu (2012), which is a time series of length $T = 107$ providing the annual counts of major earthquakes (magnitude ≥ 7) for the years 1900–2006. As shown by Zhu (2012), the data exhibit a rather strong autocorrelation $\hat{\rho}(1) \approx 0.570$ and

Model	Estimation	$\widehat{C}_{\cdot;2}$	crit^{a}	crit^{b}	$\widehat{C}_{\cdot;3}$	crit^{a}	crit^{b}	$\widehat{C}_{\cdot;4}$	crit^{a}	crit^{b}
INARCH(1)	MM	1.036	1.011	1.014	1.117	1.035	1.036	1.257	1.081	1.073
	ML	1.037		1.012	1.119		1.036	1.260		1.076
INGARCH(1, 1)	ML	1.034		1.087	1.109		1.140	1.238		1.214

Table 8 Earthquakes counts: test statistics and critical values (crit^{a}: asymptotic implementation; crit^{b}: bootstrap implementation) using moment estimation (MM) or ML estimation.

also strong overdispersion ($\hat{\mu} \approx 19.36$, $\hat{\sigma}^2 \approx 51.09$). Zhu (2012) concluded that either INARCH(1) or INGARCH(1, 1) models might be appropriate.

Assuming first an underlying INARCH(1) structure, we apply all the previously discussed implementations of the $\widehat{C}_{1;r}$-test with $r = 2, 3, 4$. Ignoring the fact that we are doing multiple testing, all tests are designed on a 5 % level as before. For the bootstrap implementations, we use $B = 1\,000$ replicates. The obtained values of the test statistics as well as of the critical values are shown in the upper part of Table 8, leading to a clear result: the null of a Poisson INARCH(1) model has to be rejected. With different arguments, Zhu (2012) obtained the same conclusion, and he preferred to use the GP-INARCH(1) model from Example 1 instead.

But Zhu (2012) also presented INGARCH(1, 1) models as a further alternative for describing the data, which is based on the equation $M_t = \alpha_0 + \alpha_1\, X_{t-1} + \beta_1\, M_{t-1}$ for the conditional means. Our statistic $\widehat{C}_{p;r}$ from (10) cannot be applied in this case, since it is limited to pure autoregression. In fact, the computation of the conditional mean M_t, as required for the denominator of the statistic, requires the complete past of the process, not only the last p observations. As a feasible solution, we propose the following modification: when fitting a Poisson INGARCH(1, 1) model to the data (by using ML estimation), follow the suggestion by Ferland et al. (2006) and treat the initial conditional mean as a further parameter, leading to the estimate $\hat{m}_1$. Compute the remaining conditional means recursively from $\hat{m}_t = \hat{\alpha}_0 + \hat{\alpha}_1\, x_{t-1} + \hat{\beta}_1\, \hat{m}_{t-1}$, and finally the modified test statistics

$$\widehat{C}_{1,1;r} \;:=\; \frac{1}{T-1} \sum_{t=2}^{T} \frac{(x_t)_{(r)}}{\hat{m}_t^r}.$$

Here, we obtain the estimates $\hat{\alpha}_0 \approx 2.699$, $\hat{\alpha}_1 \approx 0.392$, $\hat{\beta}_1 \approx 0.470$ and $\hat{m}_1 \approx 9.123$. Critical values are again computed from a parametric bootstrap with $B = 1\,000$ replications. The results in Table 8 do not give a unique picture: $\widehat{C}_{1,1;2}, \widehat{C}_{1,1;3}$ do not lead to a rejection, whereas $\widehat{C}_{1,1;4}$ is slightly larger than its critical value. So it appears that a Poisson INGARCH(1, 1) model is much better able to handle both the autocorrelation and overdispersion in the data than a Poisson INARCH(1) model does, but the $\widehat{C}_{1,1;4}$-test indicates that a model with additional conditional dispersion might do even better. It is interesting to point out that Zhu (2012) considered a GP-INGARCH(1, 1) model as being best suited to describe these data. On the other hand, one should be careful with the interpretation in view of the multiple testing.

5 Conclusions

The INGARCH models have known, since their introduction by Heinen (2003); Ferland et al. (2006), great extension and development namely through the assumption of new conditional distributions in alternative to the Poisson one, initially considered by those authors. Recently, Gonçalves et al. (2015a) introduced a wide class of this type of models, the CP-INGARCH with compound Poisson conditional distribution, which includes the main INGARCH models present in literature and, particularly, the simple Poisson INGARCH ones.

In order to contribute to the distinction between a simple Poisson INARCH model and a true CP-INARCH, we proposed in this paper a test for such hypotheses based on the form of the probability generating function of the compounding distribution related to the model conditional law. The normality of the test statistics' asymptotic distribution, for the particular case of a INARCH(1) process, was established either in the case, where the model parameters are specified, or when such parameters are consistently estimated. This involves the moments of inverse conditional means of CP-INARCH process, the analysis of their existence and calculation was conducted using two methods. For higher-order models, a bootstrap implementation of the proposed test was presented.

Acknowledgements

The authors thank the referees for useful comments on an earlier draft of this article. This work was partially supported by the Centre for Mathematics of the University of Coimbra – UID/MAT/00324/2013, funded by the Portuguese Government through FCT/MEC and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020.

References

Adell, J.A., Cal, J., Pérez-Palomares, D. (1996) On the Cheney and Sharma operator. *J. Math. Anal. Appl.* 200, 663–679.

Douglas, J.B. (1980) *Analysis with standard contagious distributions.* Int. Cooperative Publishing House, Fairland, Maryland USA.

Ferland, R., Latour, A., Oraichi, D. (2006) Integer-valued GARCH processes. *J. Time Ser. Anal.* 27(6), 923–942.

Fokianos, K., Rahbek, A., Tjøstheim, D. (2009) Poisson autoregression. *J. Amer. Stat. Assoc.* 10(4), 1430–1439.

Giacomini, R., Politis, D.N., White, H. (2013) A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory* 29(3), 567–589.

Gonçalves, E., Mendes-Lopes, N., Silva, F. (2015a) Infinitely divisible distributions in integer-valued GARCH models. *J. Time Ser. Anal.* 36(4), 503–527.

Gonçalves, E., Mendes-Lopes, N., Silva, F. (2015b) A new approach to integer-valued time series modeling: the Neyman type-A INGARCH model. *Lith. Math. J.* 55(2), 231–242.

Gradshteyn, I.S., Ryzhic, I.M. (2007) *Table of Integrals, Series and Products.* 7th edition, (Jeffrey, A., Zwillinger, D., eds.) Academic Press, New York.

Heinen, A. (2003) Modelling time series count data: an autoregressive conditional Poisson model. *CORE Discussion Paper* 2003-63, University of Louvain, Belgium.

Ibragimov, I. (1962) Some limit theorems for stationary processes. *Theory Probab. Appl.* 7(4), 349–382.

Johnson, N.L., Kemp, A.W., Kotz, S. (2005) *Univariate discrete distributions.* 3rd edition, John Wiley & Sons, Inc., Hoboken, New Jersey.

Jung, R.C., Tremayne, A.R. (2011) Useful models for time series of counts or simply wrong ones? *Adv. Stat. Anal.* 95(1), 59–91.

Lee, S., Park, S., Chen, C.W.S. (2017) On Fisher's dispersion test for integer-valued autoregressive Poisson models with applications. *Commun. Statist. Theory Meth.*, forthcoming.

Neumann, M.H. (2011) Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* 17(4), 1268–1284.

Smith, P.J. (1995) A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *Amer. Stat.* 49(2), 217–218.

Weiß, C.H. (2009) Modelling time series of counts with overdispersion. *Stat. Meth. Appl.* 18(4), 507–519.

Weiß, C.H. (2010) The INARCH(1) model for overdispersed time series of counts. *Commun. Statist. Simul. Comp.* 39(6), 1269–1291

Weiß, C.H., Schweer, S. (2016) Bias corrections for moment estimators in Poisson INAR(1) and INARCH(1) processes. *Statist. Probab. Lett.* 112, 124–130.

Weiß, C.H., Homburg, A., Puig, P. (2017) Testing for zero inflation and overdispersion in INAR(1) models. *Statist. Papers*, forthcoming.

Xu, H.-Y., Xie, M., Goh, T.N., Fu, X. (2012) A model for integer-valued time series with conditional overdispersion. *Comp. Stat. Data Anal.* 56(12), 4229–4242.

Zhu, F., Wang, D. (2010) Diagnostic checking integer-valued ARCH(p) models using conditional residual autocorrelations. *Comp. Stat. Data Anal.* 54(2), 496–508.

Zhu, F., Wang, D. (2011) Estimation and testing for a Poisson autoregressive model. *Metrika* 73(2), 211–230.

Zhu, F. (2012) Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *J. Math. Anal. Appl.* 389(1), 58–71.

A Derivations

A.1 Derivation of Formula (21)

To obtain the asymptotic variance of the approximate quantity $\widetilde{C}_{1;r}(\hat{\alpha}_0, \hat{\alpha}_1)$ from (19), we start by defining the vectors

$$\boldsymbol{Y}_t^{(r)} := \Big(\frac{(X_t)_{(r)}}{(\alpha_0 + \alpha_1 X_{t-1})^r} - 1,\ X_t - f_1,\ X_t^2 - f_2 - f_1^2,\ X_t X_{t-1} - \alpha_1 f_2 - f_1^2\Big)^\top \quad \text{(A.1)}$$

with mean $\mathbf{0}$, and by deriving a central limit theorem for $(\boldsymbol{Y}_t^{(r)})_{\mathbb{Z}}$.

Lemma 2 *Let $(X_t)_{\mathbb{Z}}$ be a stationary INARCH(1) process, define $\boldsymbol{Y}_t^{(r)}$ as in formula (A.1). Denote $f_k := \alpha_0 / \prod_{i=1}^k (1-\alpha_1^i)$ such that $\mu = f_1$ and $\sigma^2 = f_2$: Then*

$$\begin{aligned} &\tfrac{1}{\sqrt{T}} \textstyle\sum_{t=1}^T \boldsymbol{Y}_t^{(r)} \xrightarrow{\mathcal{D}} \mathrm{N}\big(\mathbf{0}, \boldsymbol{\Sigma}^{(r)}\big) \qquad \text{with } \boldsymbol{\Sigma}^{(r)} = \big(\sigma_{ij}^{(r)}\big) \text{ given by} \\ &\sigma_{ij}^{(r)} \ = \ E\big[Y_{0,i}^{(r)} Y_{0,j}^{(r)}\big] + \textstyle\sum_{k=1}^\infty \Big(E\big[Y_{0,i}^{(r)} Y_{k,j}^{(r)}\big] + E\big[Y_{k,i}^{(r)} Y_{0,j}^{(r)}\big]\Big), \end{aligned} \quad \text{(A.2)}$$

where $Y_{k,i}^{(r)}$ denotes the i-th entry of $\boldsymbol{Y}_k^{(r)}$, and where the entries $\sigma_{ij}^{(r)}$ of the symmetric matrix $\boldsymbol{\Sigma}^{(r)}$ are given as follows:

$$\sigma_{11}^{(r)} \ = \ \textstyle\sum_{k=1}^r \binom{r}{k}^2 k!\, q_{0,k} \qquad \textit{(remember (15))}, \quad \sigma_{12}^{(r)} \ = \ \frac{r}{1-\alpha_1},$$

$$\sigma_{13}^{(r)} \ = \ \frac{2r\, f_1}{1-\alpha_1} + \frac{r^2}{1-\alpha_1^2} + \frac{r\,\alpha_1}{(1-\alpha_1)(1-\alpha_1^2)}, \quad \sigma_{14}^{(r)} \ = \ \frac{2\,r\,f_1}{1-\alpha_1} + \frac{r^2\,\alpha_1}{1-\alpha_1^2} + \frac{r\,\alpha_1^2}{(1-\alpha_1)(1-\alpha_1^2)},$$

and $\quad \sigma_{22}^{(r)} \ = \ \dfrac{f_1}{(1-\alpha_1)^2},$

$$\begin{aligned} \sigma_{23}^{(r)} \ &= \ \frac{1+\alpha_1+2\alpha_1^2}{(1-\alpha_1)(1-\alpha_1^2)}\, f_2 + \frac{2\,f_1^2}{(1-\alpha_1)^2}, \qquad \sigma_{24}^{(r)} \ = \ \frac{\alpha_1(2+\alpha_1+\alpha_1^2)}{(1-\alpha_1)(1-\alpha_1^2)}\, f_2 + \frac{2\,f_1^2}{(1-\alpha_1)^2}, \\ \sigma_{33}^{(r)} \ &= \ \frac{1+2\alpha_1+8\alpha_1^2+9\alpha_1^3+4\alpha_1^4+6\alpha_1^5}{(1-\alpha_1^2)^2}\, f_3 + \frac{2(3+4\alpha_1+7\alpha_1^2+4\alpha_1^3)}{1-\alpha_1^2}\, f_2^2 + \frac{4\,f_1^3}{(1-\alpha_1)^2}, \\ \sigma_{34}^{(r)} \ &= \ \frac{\alpha_1(2+5\alpha_1+8\alpha_1^2+10\alpha_1^3+3\alpha_1^4+2\alpha_1^5)}{(1-\alpha_1^2)^2}\, f_3 + \frac{2(1+6\alpha_1+6\alpha_1^2+4\alpha_1^3+\alpha_1^4)}{1-\alpha_1^2}\, f_2^2 + \frac{4\,f_1^3}{(1-\alpha_1)^2}, \\ \sigma_{44}^{(r)} \ &= \ \frac{\alpha_1(1+3\alpha_1+8\alpha_1^2+8\alpha_1^3+8\alpha_1^4+2\alpha_1^5)}{(1-\alpha_1^2)^2}\, f_3 + \frac{1+8\alpha_1+16\alpha_1^2+8\alpha_1^3+3\alpha_1^4}{1-\alpha_1^2}\, f_2^2 + \frac{4\,f_1^3}{(1-\alpha_1)^2}. \end{aligned}$$

Proof With the same arguments as in Section 2 of Weiß & Schweer (2016), Theorem 1.7 of Ibragimov (1962) is applicable. Furthermore, the expressions for $\sigma_{kl}^{(r)}$ with $k, l \geq 2$ are already known from Theorem 2.2 in Weiß & Schweer (2016), and $\sigma_{11}^{(r)}$ was derived before in the context of formula (11). Hence, to prove Lemma 2, it remains to compute the entries $\sigma_{12}^{(r)}$, $\sigma_{13}^{(r)}$ and $\sigma_{14}^{(r)}$ of the asymptotic covariance matrix $\boldsymbol{\Sigma}^{(r)}$.

We start with some auxiliary expressions. We have

$$Q_1^{(r)} \ := E\Big[\frac{(X_t)_{(r)}\, X_t}{M_t^r}\Big] \ = \ E\Big[\frac{E[(X_t)_{(r+1)} + r\,(X_t)_{(r)} \mid X_{t-1}, \ldots]}{M_t^r}\Big] \ = \ E[M_t + r] \ = \ f_1 + r. \quad \text{(A.3)}$$

Similarly, using that

$$E[M_t^2] \ = \ \alpha_0^2 + 2\alpha_0\alpha_1\, f_1 + \alpha_1^2\,(f_2 + f_1^2) \ = \ (\alpha_0 + \alpha_1\, f_1)^2 + \alpha_1^2\, f_2 \ = \ f_1^2 + \alpha_1^2\, f_2,$$

it follows that

$$\begin{aligned} Q_2^{(r)} &:= E\Big[\tfrac{(X_t)_{(r)}\, X_t^2}{M_t^r}\Big] \;=\; E\Big[\tfrac{E[(X_t)_{(r+2)}+(2r+1)\,(X_t)_{(r+1)}+r^2\,(X_t)_{(r)}\ \mid\ X_{t-1},\ldots]}{M_t^r}\Big] \\ &= E[M_t^2 + (2r+1)\, M_t + r^2] \\ &= r^2 + f_1^2 + \alpha_1^2\, f_2 + (2r+1) f_1 \;=\; r^2 + 2r\, f_1 + f_2 + f_1^2. \end{aligned} \tag{A.4}$$

Finally,

$$\begin{aligned} Q_{1,1}^{(r)} &:= E\Big[\tfrac{(X_t)_{(r)}\, X_t X_{t-1}}{M_t^r}\Big] \;=\; E\Big[\tfrac{X_{t-1}\, E[(X_t)_{(r+1)}+r\,(X_t)_{(r)}\ \mid\ X_{t-1},\ldots]}{M_t^r}\Big] \\ &= E\big[X_{t-1}\,(M_t+r)\big] \;=\; (r+\alpha_0)\, f_1 + \alpha_1\,(f_2+f_1^2) \\ &= r\, f_1 + \alpha_1\, f_2 + f_1\,(\alpha_0+\alpha_1\, f_1) \;=\; r\, f_1 + \alpha_1\, f_2 + f_1^2. \end{aligned} \tag{A.5}$$

Now we can start with computing $\sigma_{1j}^{(r)}$ for $j = 2,3,4$. For $k \geq 1$, we always have

$$E[Y_{k,1}^{(r)} Y_{0,j}^{(r)}] \;=\; E\big[E[Y_{k,1}^{(r)} Y_{0,j}^{(r)} \mid X_{k-1},\ldots]\big] \;=\; E\big[Y_{0,j}^{(r)}\ \underbrace{E[Y_{k,1}^{(r)} \mid X_{k-1},\ldots]}_{=0}\big] \;=\; 0. \tag{A.6}$$

Let us compute $\sigma_{12}^{(r)}$ first. For $k \geq 1$, by conditioning and using that $M_k = \alpha_0 + \alpha_1\, X_{k-1}$, we have

$$\begin{aligned} E[Y_{0,1}^{(r)} Y_{k,2}^{(r)}] &= E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_k\big] \;-\; f_1 \;=\; \alpha_1\, E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_{k-1}\big] + \alpha_0 \;-\; f_1 \\ &= \ldots \;=\; \alpha_1^k\, E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_0\big] + \alpha_0\,(1+\alpha_1+\ldots+\alpha_1^{k-1}) \;-\; f_1 \\ &= \alpha_1^k\, Q_1^{(r)} + \alpha_0\, \tfrac{1-\alpha_1^k}{1-\alpha_1} \;-\; f_1 \;=\; \alpha_1^k\,(Q_1^{(r)} - f_1) \;\overset{(A.3)}{=}\; \alpha_1^k\, r, \end{aligned}$$

which also holds for $k = 0$. Together with (A.6), it follows that

$$\sigma_{12}^{(r)} \;=\; \sum_{k=0}^{\infty} E[Y_{0,1}^{(r)} Y_{k,2}^{(r)}] \;=\; \sum_{k=0}^{\infty} r\, \alpha_1^k \;=\; \frac{r}{1-\alpha_1}.$$

Concerning $\sigma_{13}^{(r)}$, first note that the 2nd non-central moment of the Poisson distribution implies

$$E[X_t^2 \mid X_{t-1},\ldots] \;=\; M_t^2 + M_t \;=\; \alpha_1^2\, X_{t-1}^2 + \alpha_1(2\alpha_0+1)\, X_{t-1} + \alpha_0(\alpha_0+1).$$

Then we compute by successive conditioning that

$$\begin{aligned} E[Y_{0,1}^{(r)} Y_{k,3}^{(r)}] &= \alpha_1^2\, E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_{k-1}^2\big] + \alpha_1(2\alpha_0+1)\,(r\,\alpha_1^{k-1} + f_1) + \alpha_0(\alpha_0+1) \;-\; f_2 - f_1^2 \\ &= \alpha_1^2\, E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_{k-1}^2\big] + (2\alpha_0+1)\, r\,\alpha_1^k + f_1\,\big(1 + f_1(1-\alpha_1^2)\big) \;-\; f_2 - f_1^2 \\ &= \ldots \;=\; \alpha_1^{2k}\, E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_0^2\big] + (2\alpha_0+1)\, r\,\alpha_1^k(1+\alpha_1+\ldots+\alpha_1^{k-1}) \\ &\qquad + f_1\,\big(1 + f_1(1-\alpha_1^2)\big)(1+\alpha_1^2+\ldots+\alpha_1^{2(k-1)}) \;-\; f_2 - f_1^2 \\ &= \alpha_1^{2k}\, Q_2^{(r)} + (2\alpha_0+1)\, r\,\alpha_1^k\, \tfrac{1-\alpha_1^k}{1-\alpha_1} + (f_2+f_1^2)(1-\alpha_1^2)\, \tfrac{1-\alpha_1^{2k}}{1-\alpha_1^2} \;-\; f_2 - f_1^2 \\ &= \alpha_1^{2k}\,\big(Q_2^{(r)} - r\, \tfrac{2\alpha_0+1}{1-\alpha_1} - f_2 - f_1^2\big) + r\,\alpha_1^k\, \tfrac{2\alpha_0+1}{1-\alpha_1} \\ &\overset{(A.4)}{=} r\,\alpha_1^{2k}\,(r - \tfrac{1}{1-\alpha_1}) + r\,\alpha_1^k\,(2f_1 + \tfrac{1}{1-\alpha_1}). \end{aligned}$$

So it follows that

$$\sigma_{13}^{(r)} \;=\; r\,(2f_1+\tfrac{1}{1-\alpha_1}) \sum_{k=0}^{\infty} \alpha_1^k + r\,(r-\tfrac{1}{1-\alpha_1}) \sum_{k=0}^{\infty} \alpha_1^{2k} \;=\; \frac{2r\, f_1}{1-\alpha_1} + \frac{r^2}{1-\alpha_1^2} + \frac{r\,\alpha_1}{(1-\alpha_1)(1-\alpha_1^2)}.$$

Finally, combining the previous derivations, we compute $\sigma_{14}^{(r)}$ as

$$\begin{aligned} E[Y_{0,1}^{(r)} Y_{k,4}^{(r)}] &= \alpha_1\, E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_{k-1}^2\big] + \alpha_0\, E\big[\tfrac{(X_0)_{(r)}}{M_0^r}\, X_{k-1}\big] \;-\; \alpha_1 f_2 - f_1^2 \\ &= \alpha_1 \Big(r\,\alpha_1^{2(k-1)}\,(r - \tfrac{1}{1-\alpha_1}) + r\,\alpha_1^{k-1}\,(2f_1 + \tfrac{1}{1-\alpha_1}) \;+\; f_2 + f_1^2\Big) \\ &\quad + \alpha_0\,(r\,\alpha_1^{k-1} + f_1) \;-\; \alpha_1 f_2 - f_1^2 \\ &= \tfrac{r}{\alpha_1}\,(r - \tfrac{1}{1-\alpha_1})\,\alpha_1^{2k} + r\,\alpha_1^k\,\big(\tfrac{1}{1-\alpha_1} + f_1 + \tfrac{f_1}{\alpha_1}\big) \end{aligned}$$

for $k \geq 1$, while

$$E[Y_{0,1}^{(r)} Y_{0,4}^{(r)}] \;=\; Q_{1,1}^{(r)} - \alpha_1 f_2 - f_1^2 \;\overset{(A.5)}{=}\; r\,f_1.$$

Therefore,

$$\begin{aligned} \sigma_{14}^{(r)} &= r\,\big(\tfrac{1}{1-\alpha_1} + f_1 + \tfrac{f_1}{\alpha_1}\big)\,\textstyle\sum_{k=0}^{\infty} \alpha_1^k \;+\; \tfrac{r}{\alpha_1}\,\big(r - \tfrac{1}{1-\alpha_1}\big)\,\textstyle\sum_{k=0}^{\infty} \alpha_1^{2k} \\ &\quad - \tfrac{r}{\alpha_1}\,\big(r - \tfrac{1}{1-\alpha_1}\big) - r\,\big(\tfrac{1}{1-\alpha_1} + \tfrac{f_1}{\alpha_1}\big) \\ &= r\,\big(\tfrac{1}{(1-\alpha_1)^2} + \tfrac{f_1(1+\alpha_1)}{\alpha_1(1-\alpha_1)}\big) \;+\; \tfrac{r}{\alpha_1(1-\alpha_1^2)}\,\big(r - \tfrac{1}{1-\alpha_1}\big) \;-\; \tfrac{r^2}{\alpha_1} + \tfrac{r}{\alpha_1} - \tfrac{r\,f_1}{\alpha_1} \\ &= \tfrac{2\,r\,f_1}{1-\alpha_1} \;+\; \tfrac{r^2\,\alpha_1}{1-\alpha_1^2} \;+\; \tfrac{r\,\alpha_1^2}{(1-\alpha_1)(1-\alpha_1^2)}. \end{aligned}$$

This completes the proof.

In the next step, we apply the Delta method to derive the joint distribution of $(\widehat{C}_{1;\,r}, \hat{\alpha}_0, \hat{\alpha}_1)^\top$.

Corollary 1 *Let $(X_t)_{\mathbb{Z}}$ be a stationary INARCH(1) process. Then the distribution of $(\widehat{C}_{1;\,r}, \hat{\alpha}_0, \hat{\alpha}_1)^\top$ is asymptotically approximated by a normal distribution with mean vector $(1, \alpha_0, \alpha_1)^\top$ and covariance matrix $\frac{1}{T-1}\,\tilde{\boldsymbol{\Sigma}}^{(r)}$, where*

$$\tilde{\boldsymbol{\Sigma}}^{(r)} \;=\; \begin{pmatrix} \sum_{k=1}^r \binom{r}{k}^2 k!\, q_{0,k} & r & 0 \\ r & \frac{\alpha_0}{1-\alpha_1}\big(\alpha_0(1+\alpha_1) + \frac{1+2\alpha_1^4}{1+\alpha_1+\alpha_1^2}\big) & -\alpha_0(1+\alpha_1) - \frac{(1+2\alpha_1)\alpha_1^3}{1+\alpha_1+\alpha_1^2} \\ 0 & -\alpha_0(1+\alpha_1) - \frac{(1+2\alpha_1)\alpha_1^3}{1+\alpha_1+\alpha_1^2} & (1-\alpha_1^2)\big(1 + \frac{\alpha_1(1+2\alpha_1^2)}{\alpha_0(1+\alpha_1+\alpha_1^2)}\big) \end{pmatrix}.$$

Proof Define the function $\boldsymbol{g} : \mathbb{R}^4 \to \mathbb{R}^3$ by

$$g_1(\boldsymbol{y}) \;:=\; y_1, \quad g_2(\boldsymbol{y}) \;:=\; y_2\,\frac{y_3 - y_4}{y_3 - y_2^2}, \quad g_3(\boldsymbol{y}) \;:=\; \frac{y_4 - y_2^2}{y_3 - y_2^2}. \tag{A.7}$$

Note that $g_2\big(\cdot, f_1, f_2 + f_1^2, \alpha_1 f_2 + f_1^2\big) \;=\; \alpha_0$ and $g_3\big(\cdot, f_1, f_2 + f_1^2, \alpha_1 f_2 + f_1^2\big) \;=\; \alpha_1$.

From the proof of Theorem 4.2 in Weiß & Schweer (2016) (see p. 13 in Appendix B.4), we know that the Jacobian of $\boldsymbol{g}$ equals

$$\mathbf{J}_{\boldsymbol{g}}(\boldsymbol{y}) \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \dfrac{(y_3 - y_4)(y_3 + y_2^2)}{(y_3 - y_2^2)^2} & \dfrac{y_2(y_4 - y_2^2)}{(y_3 - y_2^2)^2} & \dfrac{-y_2}{y_3 - y_2^2} \\ 0 & \dfrac{2y_2(y_4 - y_3)}{(y_3 - y_2^2)^2} & \dfrac{y_2^2 - y_4}{(y_3 - y_2^2)^2} & \dfrac{1}{y_3 - y_2^2} \end{pmatrix},$$

such that $\mathbf{D} := \mathbf{J}_{\boldsymbol{g}}\big(1, f_1, f_2 + f_1^2, \alpha_1 f_2 + f_1^2\big)$ is given by

$$\begin{aligned} \mathbf{D} \;&=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \dfrac{(1-\alpha_1)(f_2 + 2f_1^2)}{f_2} & \dfrac{\alpha_1 f_1}{f_2} & -\dfrac{f_1}{f_2} \\ 0 & -\dfrac{2(1-\alpha_1)\,f_1}{f_2} & -\dfrac{\alpha_1}{f_2} & \dfrac{1}{f_2} \end{pmatrix} \\ &=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-\alpha_1)\big(1 + 2(1-\alpha_1^2)\,f_1\big) & \alpha_1(1-\alpha_1^2) & -(1-\alpha_1^2) \\ 0 & -2(1-\alpha_1)(1-\alpha_1^2) & -\dfrac{\alpha_1}{f_2} & \dfrac{1}{f_2} \end{pmatrix}. \end{aligned}$$

Now, let us look at

$$\tilde{\boldsymbol{\Sigma}}^{(r)} = \big(\tilde{\sigma}_{ij}^{(r)}\big) \;:=\; \mathbf{D}\boldsymbol{\Sigma}^{(r)}\mathbf{D}^\top,$$

where $\boldsymbol{\Sigma}^{(r)}$ is the covariance matrix from Lemma 2 above. The components $\tilde{\sigma}_{22}^{(r)}, \tilde{\sigma}_{23}^{(r)}, \tilde{\sigma}_{33}^{(r)}$ are already known from formula (11) in Weiß (2010) (or from Theorem 4.2 in Weiß & Schweer (2016)), and $\tilde{\sigma}_{11}^{(r)} = \sigma_{11}^{(r)}$ obviously holds.

So it remains to compute $\tilde{\sigma}_{12}^{(r)} = \sum_{j=2}^4 d_{11}d_{2j}\,\sigma_{1j}^{(r)}$ and $\tilde{\sigma}_{13}^{(r)} = \sum_{j=2}^4 d_{11}d_{3j}\,\sigma_{1j}^{(r)}$. We get

$$\begin{aligned}
\tilde{\sigma}_{12}^{(r)} &= (1-\alpha_1)\big(1+2(1-\alpha_1^2)\,f_1\big)\,\sigma_{12}^{(r)} + \alpha_1(1-\alpha_1^2)\,\sigma_{13}^{(r)} - (1-\alpha_1^2)\,\sigma_{14}^{(r)} \\
&= r + 2r\,(1-\alpha_1^2)\,f_1 + 2r\,f_1\,\alpha_1(1+\alpha_1) + r^2\,\alpha_1 + \tfrac{r\,\alpha_1^2}{1-\alpha_1} - 2\,r\,f_1\,(1+\alpha_1) - r^2\,\alpha_1 - \tfrac{r\,\alpha_1^2}{1-\alpha_1} \\
&= r,
\end{aligned}$$

as well as

$$\begin{aligned}
\tilde{\sigma}_{13}^{(r)} &= -2(1-\alpha_1)(1-\alpha_1^2)\,\sigma_{12}^{(r)} - \tfrac{\alpha_1}{f_2}\,\sigma_{13}^{(r)} + \tfrac{1}{f_2}\,\sigma_{14}^{(r)} \\
&= -2r\,(1-\alpha_1^2) - 2r\,\alpha_1(1+\alpha_1) - \tfrac{r^2\,\alpha_1}{f_1} - \tfrac{r\,\alpha_1^2}{f_1\,(1-\alpha_1)} + 2r\,(1+\alpha_1) + \tfrac{r^2\,\alpha_1}{f_1} + \tfrac{r\,\alpha_1^2}{f_1\,(1-\alpha_1)} \\
&= 0.
\end{aligned}$$

This completes the proof.

Using Corollary 1, we are able to approximate the variance of $\widehat{C}_{1;\,r}(\hat{\alpha}_0, \hat{\alpha}_1)$ by the asymptotic variance $\frac{1}{T-1}\,\sigma_{1;\,r}^2$ of $\widetilde{C}_{1;\,r}(\hat{\alpha}_0, \hat{\alpha}_1)$ according to (19):

$$\begin{aligned}
\sigma_{1;\,r}^2 &= \tilde{\sigma}_{11}^{(r)} + r^2\,q_{0,1}^2\,\tilde{\sigma}_{22}^{(r)} + r^2\,q_{1,1}^2\,\tilde{\sigma}_{33}^{(r)} - 2r\,q_{0,1}\,\tilde{\sigma}_{12}^{(r)} + 2r^2\,q_{0,1}q_{1,1}\tilde{\sigma}_{23}^{(r)} \\
&= \textstyle\sum_{k=1}^r \binom{r}{k}^2 k!\,q_{0,k} \;-\; 2r^2\,q_{0,1} \;+\; r^2\,q_{0,1}^2\,\frac{\alpha_0}{1-\alpha_1}\big(\alpha_0(1+\alpha_1) + \frac{1+2\alpha_1^4}{1+\alpha_1+\alpha_1^2}\big) \\
&\quad + r^2\,q_{1,1}^2\,(1-\alpha_1^2)\big(1 + \tfrac{\alpha_1(1+2\alpha_1^2)}{\alpha_0(1+\alpha_1+\alpha_1^2)}\big) \;-\; 2r^2\,q_{0,1}q_{1,1}\,\big(\alpha_0(1+\alpha_1) + \tfrac{(1+2\alpha_1)\alpha_1^3}{1+\alpha_1+\alpha_1^2}\big).
\end{aligned}$$

So the proof of formula (21) is complete.

A.2 Derivation of Equality (26)

First, we note that if the random variable Z follows a Poisson distribution with mean λ, and if $a > 0$, we have for $k = 1, 2, \ldots$

$$\begin{aligned}
E\left[\left(\frac{a}{a+Z}\right)^k\right] &= \int_0^1 \exp\big(-\lambda\,(1-s)\big)\,\frac{a^k}{(k-1)!}\,s^{a-1}\,\log^{k-1}\left(\frac{1}{s}\right)\,ds \\
&= \frac{a^k}{(k-1)!}\sum_{n=0}^{+\infty}\frac{(-1)^n\,\lambda^n}{n!}\int_0^1 (1-s)^n\,s^{a-1}\,\log^{k-1}\left(\frac{1}{s}\right)\,ds \\
&= a^k\sum_{n=0}^{+\infty}\frac{(-1)^n\,\lambda^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{(a+j)^k},
\end{aligned}$$

using the Dominated Convergence Theorem and the following result (formula 16 on page 552 of Gradshteyn & Ryzhic (2007))

$$\int_0^1 \left(\log\frac{1}{x}\right)^n (1-x^q)^m\,x^{p-1}\,dx \;=\; n!\sum_{k=0}^{m}\binom{m}{k}\frac{(-1)^k}{(p+kq)^{n+1}} \qquad \text{with } p, q > 0.$$

We note that for $k = 1$, the expression may be replaced by the equivalent one

$$E\left[\frac{a}{a+Z}\right] = \Gamma(a+1)\sum_{n=0}^{+\infty}\frac{(-1)^n}{\Gamma(a+n+1)}\lambda^n,$$

since

$$\frac{\Gamma(a+1)}{\Gamma(a+n+1)} = \frac{a}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{a+j}$$

as may be proved by recurrence.

Let us now consider that the moment generating function of M_1, $\mathrm{mgf}_{M_1}(u) = E[\exp(uM_1)]$, is defined for every $u \in (u_1; u_2)$, where $u_1 < 0 < u_2$ such that $\min\{-u_1, u_2\} = b > 2$. With these conditions, we will prove that

$$E\left[\frac{1}{M_t^l}\right] = \frac{1}{\alpha_1^l}\sum_{n=0}^{+\infty}\frac{(-1)^n}{n!}E[M_{t-1}^n]\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l},$$

that is, the change between the expectation and the infinite sum is allowed. For this purpose, let us consider s such that $0 < s < \frac{1}{2}\min\{-u_1, u_2\}$ and the function

$$H(t) = \int\sum_{n=0}^{+\infty}\frac{(-1)^n(tx)^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l}dP_{M_1}(x), \qquad t \in (-s; s).$$

Considering the functions

$$h_k(x) = \sum_{n=0}^{k}\frac{(-1)^n(tx)^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l} \qquad \text{with } k \in \mathbb{N}_0,$$

and $h(x) := h_\infty(x)$, we have for every x and for $k = 1, 2, \ldots$

$$|h_k(x)| \le \sum_{n=0}^{k}\frac{|tx|^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{1}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l} \le \left(\frac{\alpha_1}{\alpha_0}\right)^l\sum_{n=0}^{k}\frac{(2|tx|)^n}{n!} \le \left(\frac{\alpha_1}{\alpha_0}\right)^l\exp(2s|x|),$$

since $|t| < s$, and also $\lim_{k\to\infty}h_k(x) = h(x)$. Moreover,

$$\begin{aligned}\int\exp(2s|x|)\,dP_{M_1}(x) &\le \int_{-\infty}^{+\infty}\exp(2sx)\,dP_{M_1}(x) + \int_{-\infty}^{+\infty}\exp(-2sx)\,dP_{M_1}(x)\\ &= \mathrm{mgf}_{M_1}(2s) + \mathrm{mgf}_{M_1}(-2s) < +\infty.\end{aligned}$$

So, we may apply the Dominated Convergence Theorem and we obtain

$$H(t) = \int h(x)\,dP_{M_1}(x) = \lim_{k\to\infty}\sum_{n=0}^{k}\frac{(-1)^n t^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l}\int x^n dP_{M_1}(x),$$

that is,

$$E\left[\sum_{n=0}^{+\infty}\frac{(-1)^n(tM_{t-1})^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l}\right] = \sum_{n=0}^{+\infty}\frac{(-1)^n}{n!}E[t^nM_{t-1}^n]\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1}+j\right)^l},$$

for $t \in [-s; s]$. The result is valid for $t = 1$ if and only $s > 1$, which is possible as $\min\{-u_1, u_2\} > 2$, and so (26) follows.