

Research article

An alternative hyper-Poisson integer-valued GARCH model with application to polio, internet protocol and COVID-19 data

Kee Wah Fo¹, Seng Huat Ong^{2,3}, Choung Min Ng³, and You Beng Koh^{3,*}

¹ Faculty of Engineering and Quantity Surveying, INTI International University, Persiaran Bandar Baru Nilai, 71800 Nilai, Malaysia

² Institute of Actuarial Science and Data Analytics, UCSI University, 56000 Kuala Lumpur, Malaysia

³ Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

* **Correspondence:** Email: kohyoubeng@um.edu.my.

Abstract: Time series of counts are observed widely in actuarial science, finance, epidemiology and biology. These time series may exhibit over-, equi- and under-dispersion. The Poisson distribution is commonly used in count time series models, but it is restricted by the equality of mean and variance. Other distributions such as the generalized Poisson, double Poisson, hyper-Poisson, and COM-Poisson distributions have been proposed to replace the Poisson distribution to model the different levels of dispersion in time series of counts. These models have certain limitations such as complex expressions for the mean and variance which complicate the formulation as GARCH models. In this study, we propose an alternative hyper-Poisson (AHP) distribution, with simple forms of conditional mean and variance, for an integer-valued GARCH (INGARCH) model for time series of counts that also exhibit the different levels of dispersion. We demonstrate that the AHP-INGARCH model is comparable to some existing INGARCH models. Additionally, the model can cover a wider range of dispersion. The maximum likelihood estimation can be used to estimate the parameters of the proposed model. Applications to three real-life data sets related to polio, internet protocol and daily COVID-19 new deaths underscore the usefulness of the proposed model in studying both over-dispersed and under-dispersed time series of counts.

Keywords: alternative hyper-Poisson; integer-valued GARCH; time series of counts; polio; internet protocol; COVID-19

Mathematics Subject Classification: 60G10, 62M10, 62F10

1. Introduction

Various modelling approaches have been proposed for the time series of counts and recent reviews on this topic can be found in many sources ([5,15,16,21,22,24,32,34]). The time series of counts analysis in the literature can generally be classified into parameter-driven or observation-driven models. Zeger [38] initiated a class of parameter-driven models for time series of counts, which introduces autocorrelation as well as over-dispersion into the model through a latent process. Despite having wide applications in various fields, there is difficulty with efficient estimation for such parameter-driven models. Recently, Koh et al. [24] showed that the Monte Carlo Expectation Maximization (MCEM) algorithm is one of the alternative for the estimation of the parameters, and particle filter and smoother are useful approaches in inferencing the unobserved latent variables of the model. Jung et al. [21] and Jung and Tremayne [22] have reviewed both approaches of the parameter-driven and observation-driven models for time series of counts as well as the estimation and diagnostic tests performed for these time series models.

Observation-driven thinning-based models are commonly used to model over-dispersed or under-dispersed count data, see for example Weiß [33], Bourguignon and Weiß [3], Yang [37] and Kang [23]. Weiß [33] showed that the first-order nonnegative integer-value autoregressive (INAR(1)) model based on binomial thinning operator with Good- and power-law weighted Poisson-distributed innovations are particularly well-suited for modelling under-dispersed counts. Bourguignon and Weiß [3] proposed a new INAR(1) model for stationary count data processes with Bernoulli-geometric (BerG) marginal distributions, that can model time series of counts with over-, equi- and under-dispersion, from a new generalized thinning operator based on the convolution of binomial and negative binomial random variables. Recently, Yang [37] introduced a novel thinning operator based on the generalized Poisson distribution, called GP thinning operator, and Kang [23] defined a new thinning operator, called GSC thinning operator, based on a new discrete distribution proposed by Gómez-Déniz et al. [17], to construct a new INAR(1) model to capture the dispersion features of count time series count.

Another popular observation-driven approach for modelling time series of count are the INGARCH models. Heinen [19] proposed the autoregressive conditional Poisson (ACP) model, Ferland et al. [14] called it as integer-valued generalised autoregressive conditional heteroscedasticity (INGARCH) model, that has the advantage in the ease of application of the likelihood to evaluate the model. More specifically, the time series of counts $\{Y_t\}$ for $t \geq 1$ follows a conditional Poisson distribution with an autoregressive mean μ_t , which is defined as

$$Y_t | \mathcal{F}_{t-1} \sim \text{Poisson}(\mu_t)$$

$$E[Y_t | \mathcal{F}_{t-1}] = \mu_t = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \mu_{t-j}, \quad (1.1)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, q$, $p \geq 1$, $q \geq 0$ and \mathcal{F}_{t-1} denotes the information available on the series up to and including time $t - 1$. It is worth noting that the conditional mean (1.1), dependent on past observed counts and past conditional means, is similar to the conditional variance in the GARCH model by Bollerslev [2] as well as the conditional intensity of the autoregressive conditional duration model by Engle and Russell [12] for continuous-valued discrete-time data. However, the INGARCH model with Poisson distribution is suitable only for positive serial correlation and over-dispersion in the count data.

In this paper, we will focus on developing a class of observation-driven INGARCH-type models for modelling time series of counts that may exhibit over-, equi- and under-dispersion. The Poisson distribution, which has been commonly used to model time series of counts, is restrictive to the model variance equal to the mean; that is, for modelling equi-dispersed data. Therefore, Heinen [19] proposed the generalised double autoregressive conditional Poisson model which can model over-, equi- and under-dispersion of count time series. The double Poisson (DP) distribution proposed by Efron [11], which has two parameters, is used in this model. Bourguignon et al. [4] also used the double Poisson distribution to extend the INAR(1) for modelling count time series with over-, equi- and under-dispersion. However, Winkelmann [36] found that the results of the DP with its normalizing constant approximated by Efron's original method are not exact, making it a hurdle for applications. Moreover, Zhu [41] also argued that the DP distribution, having intractable normalizing constant and moments, is difficult to utilize for modelling and many of its properties remained unknown. Hence, some of the theoretical aspects of the resulting DP-INGARCH(1,1) models may be difficult to establish. On the other hand, Zou et al. [42] noted that while DP provides a good fit when the mean is high for all dispersion types, the fit is highly unreliable when the mean is small. With these shortcomings, the DP distribution may be unattractive for the analysis of time series despite being able to model over-, equi- and under-dispersion of the count data.

Zhu [39] presented the negative binomial INGARCH (NB-INGARCH) model that works well for an over-dispersed poliomyelitis monthly data, and claimed that the proposed model is better than the Poisson- and DP-INGARCH models. Although the mean-variance relationship of the NB enables over-dispersion to be captured, there is difficulty in handling data characterized by under-dispersion ([26]). Subsequently, Zhu [41] proposed a generalized Poisson INGARCH (GP-INGARCH) model as an alternative model to account for both over- and under-dispersion in time series of counts data. The generalized Poisson (GP) distribution has been studied extensively by many sources ([6,7,13]). Unfortunately, Zhu [40] states that the GP model is not a true probability model under certain conditions, leading to its inability to model some levels of under-dispersion.

Zhu [40] in turn addressed the weakness of this model by proposing a COM-Poisson INGARCH model and indicating that this model is a powerful competitor to the GP model. The COM-Poisson (COMP) distribution has been proposed by Conway and Maxwell [8] as a model for queuing systems with state-dependent service times. This distribution has been used widely after Shmueli et al. [29] further examined its statistical and probabilistic properties. The COMP distribution has two parameters with λ as the centering parameter and ν as the dispersion parameter. Zhu [40] noted that a COMP model based on the COMP formulation for λ would be difficult to interpret. To overcome this problem, a re-parameterization of the COMP distribution, $\mu = \lambda^{1/\nu}$, has been proposed by Guikema and Goffelt [18] to provide a clear centering parameter. Recently, Qian and Zhu [27] used the generalized COMP distribution, which has one more parameter than COMP distribution to handle heavy-tailed count time series.

Suitable to model both over- and under-dispersed data, the hyper-Poisson (HP) distribution is a popular distributions besides DP, GP and COMP distributions and belongs to the two-parameter family of discrete distributions. The HP distribution, introduced by Bardwell and Crow [1], is generated by the confluent function and admits over-dispersion as well as under-dispersion. Although the HP distribution was applied in generalized linear models and regression models for over-dispersed and under-dispersed count data (see Sáez-Castillo and Conde-Sánchez [28]), it is difficult to apply in INGARCH model. The mean and variance are in confluent hypergeometric form and hence, the conditional mean and conditional variance for INGARCH model are hard to derive. Kumar and Nair [25]

then introduced an alternative hyper-Poisson (AHP) distribution, which has simpler forms of the mean and variance.

In this paper, we introduce the AHP distribution as an INGARCH model for modelling over-dispersed and under-dispersed count time series data, and then compare it to existing INGARCH models. The AHP distribution has simpler forms for the mean and variance compared to the HP distribution, making it potentially more viable and useful choice as an INGARCH model. We also investigate the properties of the AHP distribution and derive a lower bound for one of the parameters to ensure a valid probability mass function (pmf) for the case of under-dispersion. Similar to Weiß [31], we provide a set of Yule-Walker type equations from which the autocorrelation function of general AHP-INGARCH(p, q) models can be obtained. In particular, we derive the equations of variance and autocorrelation function for AHP-INGARCH(1,1) model. It is worth to noting that the unconditional variance of AHP-INGARCH(1,1) model is more general than the unconditional variance of INGARCH model of Weiß [31]. The INGARCH model of Weiß [31] is a special case of AHP-INGARCH model when $\gamma = 1$. In the applications to three real-life data sets, we have included five existing INGARCH models, namely the Poisson-, NB-, COMP-, GP- and DP-INGARCH models in direct comparison against the proposed AHP-INGARCH model, highlighting the usefulness of AHP-INGARCH as a competitive model in practice.

The contents of this paper are organized as follows. Section 2 presents the AHP distribution along with its basic properties, while Section 3 outlines the proposed INGARCH model with the AHP distribution. The maximum likelihood estimation for the proposed model is briefly discussed in Section 4, followed by a simulation study to assess the performance of this estimation method in Section 5. Section 6 demonstrates the application of the proposed model to three real-life data sets, which exhibit different types of dispersions, in comparison to some existing models for time series of counts. Finally, Section 7 provides some concluding remarks.

2. Some properties of the alternative hyper-Poisson distribution

In this section, we give a brief review of the hyper-Poisson (HP) distribution and its alternative form. The two-parameter HP distribution is first introduced by Bardwell and Crow [1] through the following probability generating function (pgf):

$$G(u) = E[u^Z] = \frac{\phi(1; \gamma; \theta u)}{\phi(1; \gamma; \theta)}, \quad \theta > 0, \quad \gamma > 0,$$

where $\phi(a; b; c) = \sum_{k=0}^{\infty} \frac{(a)_k c^k}{(b)_k k!}$ is the confluent hypergeometric series, or known as the Kummer M function, in which $(a)_k = \frac{\Gamma(a+k)}{\Gamma(a)}$ is the Pochhammer symbol and $\Gamma(\cdot)$ is a gamma function. Sáez-Castillo and Conde-Sánchez [28] utilize the HP distribution to model over-dispersed and under-dispersed count data and showed that the range of values of the dispersion index of HP distribution is wide in both situations of over- and under-dispersion. However, the mean of HP distribution can only be expressed in term of the confluent hypergeometric series.

Kumar and Nair [25] then introduced an alternative form of HP distribution, called AHP. The pmf and pgf of AHP distribution ($AHP(\theta, \gamma)$) can be written, respectively as

$$P(Z = z) = \frac{\theta^z}{(\gamma)_z} \phi(1 + z; \gamma + z; -\theta), \quad \theta > 0, \quad \gamma > 0, \quad z = 0, 1, 2, \dots,$$

and

$$G(u) = \phi(1; \gamma; \theta(u-1)) = \sum_{k=0}^{\infty} \frac{(1)_k [\theta(u-1)]^k}{(\gamma)_k k!},$$

where Z is the AHP distributed random variable. Note that when $\gamma = 1$, we have $G(u) = \sum_{k=0}^{\infty} \frac{[\theta(u-1)]^k}{k!} = e^{\theta(u-1)}$, the pgf of a Poisson distribution with parameter θ .

The mean and variance of a AHP distributed random variable Z are, respectively, $\mu_Z = E[Z] = \frac{\theta}{\gamma}$ and $\sigma_Z^2 = \frac{\theta}{\gamma} \left[1 + \frac{\theta}{\gamma} \left(\frac{\gamma-1}{\gamma+1} \right) \right] = \mu_Z - \mu_Z^2 + \frac{2\gamma\mu_Z^2}{1+\gamma}$. However, we found that in order to guarantee the positivity of the variance, the parameter γ needs to be constrained by $\gamma > \frac{-1-\theta+\sqrt{\theta^2+6\theta+1}}{2}$ (or $\gamma > \frac{\mu_Z-1}{\mu_Z+1}$). Figure 1 shows the feasible range of γ for a given μ_Z ; we note that large γ is required for large μ_Z so that $\sigma_Z^2 > 0$.

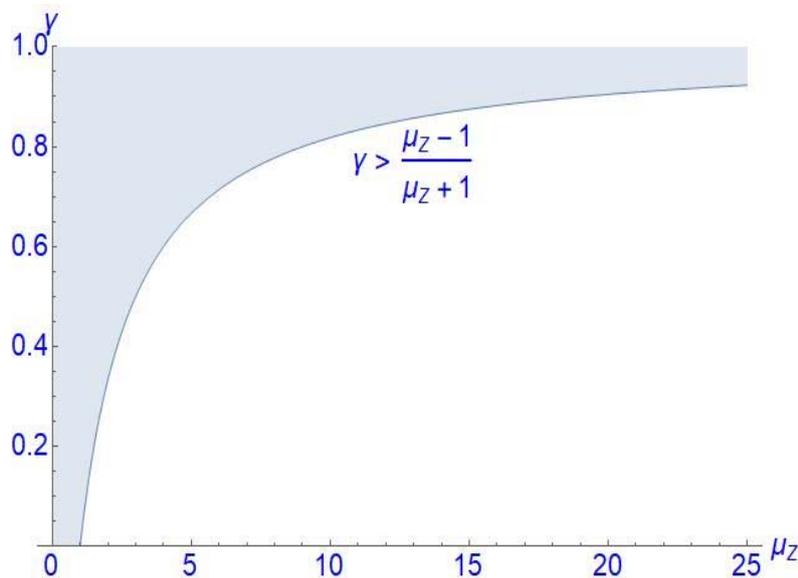


Figure 1. Feasible range of γ for a given mean μ_Z (shaded region).

The parameter constraint for the positivity of the variance can also be rewritten as $\theta < \theta_1$, for $0 < \gamma < 1$, where $\theta_1 = \frac{\gamma(\gamma+1)}{1-\gamma}$. However, in order to guarantee the positivity of the pmf of AHP distribution, the parameter θ needs to be constrained by $\theta < \theta_2$, where θ_2 is the solution of $\phi(\gamma-1; \gamma; \theta_2) = 0$ (see Appendix A.1 for the proof). Figure 2 shows the boundary lines of θ_1 and θ_2 for a given $\gamma \in (0,1)$. For an AHP distribution, the feasible region of θ for a given $\gamma \in (0,1)$ is the region below the solid line in Figure 2.

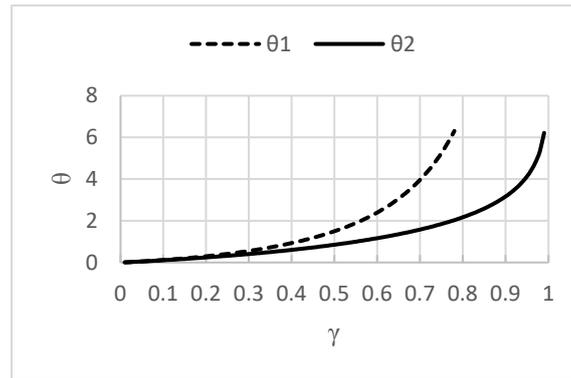


Figure 2. The boundary lines for the positivity of variance (dash line) and positivity of pmf (solid line) of an AHP distribution with $0 < \gamma < 1$.

The dispersion index of AHP distribution is given by

$$I_Z = \frac{\sigma_Z^2}{\mu_Z} = 1 - \mu_Z \left(\frac{1 - \gamma}{1 + \gamma} \right).$$

We note that the AHP distribution is under-dispersed ($0 < I_Z < 1$) when $\frac{-1 - \theta + \sqrt{\theta^2 + 6\theta + 1}}{2} < \gamma < 1$, equi-dispersed ($I_Z = 1$) when $\gamma = 1$, and over-dispersed ($I_Z > 1$) when $\gamma > 1$. Hence, the parameter γ is the dispersion parameter for AHP distribution. Bourguignon and Weiß [3] studied the dispersion behaviour of a BerG distribution, that is, the distribution of the convolution of a Bernoulli and a geometric random variable. They found that the BerG distribution can be used in count time series modelling with over-, equi- and under-dispersion. As noted by Bourguignon and Weiß [3], there is no feasible region for under-dispersion when the mean is larger than two for a BerG distribution. However, the AHP distribution can be seen to cover a wider range of dispersion regions compared to the BerG distribution.

In addition, the third and fourth central moments of an AHP random variable Z are obtained as follows:

$$\mu_3 = E[(Z - \mu_Z)^3] = \mu_Z - 3\mu_Z^2 + \frac{6\gamma\mu_Z^2}{1+\gamma} + 2\mu_Z^3 - \frac{6\gamma\mu_Z^3}{1+\gamma} + \frac{6\gamma^2\mu_Z^3}{(1+\gamma)(2+\gamma)},$$

$$\begin{aligned} \mu_4 = E[(Z - \mu_Z)^4] &= \mu_Z - 4\mu_Z^2 + \frac{14\gamma\mu_Z^2}{1+\gamma} + 6\mu_Z^3 - \frac{24\gamma\mu_Z^3}{1+\gamma} + \frac{36\gamma^2\mu_Z^3}{(1+\gamma)(2+\gamma)} - 3\mu_Z^4 \\ &+ \frac{12\gamma\mu_Z^4}{1+\gamma} - \frac{24\gamma^2\mu_Z^4}{(1+\gamma)(2+\gamma)} + \frac{24\gamma^3\mu_Z^4}{(1+\gamma)(2+\gamma)(3+\gamma)}. \end{aligned}$$

3. The AHP-INGARCH(p, q) model

Let $\{Y_t\}_{t \geq 1}$ denotes a univariate time series of counts and \mathcal{F}_{t-1} be the σ -field generated by $\{Y_{t-1}, Y_{t-2}, \dots\}$. We assume that process $\{Y_t\}$ is conditionally independent given \mathcal{F}_{t-1} and the conditional distribution of Y_t given \mathcal{F}_{t-1} is specified by an AHP distribution, that is

$$Y_t | \mathcal{F}_{t-1} \sim \text{AHP}(\theta_t, \gamma),$$

$$E[Y_t | \mathcal{F}_{t-1}] = \frac{\theta_t}{\gamma} = \mu_t = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \mu_{t-j}, \quad (3.1)$$

where $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, i = 1, \dots, p, j = 1, \dots, q, p \geq 1, q \geq 0$, and $\gamma > 0$. This model is denoted by AHP-INGARCH(p, q). Since the conditional distribution of Y_t given \mathcal{F}_{t-1} is specified by an AHP distribution, this model will be able to account for equi- ($\gamma = 1$), over- ($\gamma > 1$) and under-

dispersion ($\frac{-1-\theta_t + \sqrt{\theta_t^2 + 6\theta_t + 1}}{2} < \gamma < 1$). When $\gamma = 1$, the model (3.1) is equal to the Poisson-INGARCH model introduced by Heinen [19] and Ferland et al. [14]. Some of the important properties of the AHP-INGARCH(p, q) time series model are given next.

Theorem 3.1. *Let $\{Y_t\}$ be a weakly stationary process with range $\{0, 1, \dots\}$ following the AHP-INGARCH(p, q) model in (3.1). If $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$, then*

(i) *the unconditional expectation of Y_t is given by*

$$\mu_Y = E[Y_t] = \frac{\alpha_0}{(1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j)};$$

(ii) *the covariance $Cov[Y_t, \mu_{t-k}]$ fulfills*

$$Cov[Y_t, \mu_{t-k}] = \begin{cases} Cov[\mu_t, \mu_{t-k}], & k \geq 0, \\ Cov[Y_t, Y_{t-k}], & k < 0; \end{cases}$$

(iii) *the autocovariances $\gamma_Y(k) = Cov[Y_t, Y_{t-k}]$ and $\gamma_\mu(k) = Cov[\mu_t, \mu_{t-k}]$ will satisfy the following equations:*

$$\gamma_Y(k) = \sum_{i=1}^p \alpha_i \gamma_Y(|k-i|) + \sum_{j=1}^{\min(k-1, q)} \beta_j \gamma_Y(k-j) + \sum_{j=k}^q \beta_j \gamma_\mu(j-k), \quad k \geq 1, \quad (3.2)$$

$$\gamma_\mu(k) = \sum_{i=1}^{\min(k, p)} \alpha_i \gamma_\mu(k-i) + \sum_{i=k+1}^p \alpha_i \gamma_Y(i-k) + \sum_{j=1}^q \beta_j \gamma_\mu(|k-j|), \quad k \geq 0. \quad (3.3)$$

The proof of Theorem 3.1 is provided in Appendix A.2.

From (3.1), the conditional mean and conditional variance of Y_t are given by, respectively,

$$E(Y_t | \mathcal{F}_{t-1}) = \frac{\theta_t}{\gamma} = \mu_t,$$

and

$$V(Y_t | \mathcal{F}_{t-1}) = \frac{\theta_t}{\gamma} \left[1 + \frac{\theta_t (\gamma-1)}{\gamma (\gamma+1)} \right] = \mu_t \left[1 + \mu_t \frac{(\gamma-1)}{(\gamma+1)} \right].$$

Then, the unconditional mean and unconditional variance of Y_t are, respectively,

$$E(Y_t) = \mu_Y = \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j},$$

and

$$V(Y_t) = \sigma_Y^2 = E[V(Y_t | \mathcal{F}_{t-1})] + V[E(Y_t | \mathcal{F}_{t-1})] = \mu_Y + \mu_Y^2 \frac{\gamma-1}{\gamma+1} + \left(\frac{2\gamma}{\gamma+1} \right) V(\mu_t).$$

3.1. AHP-INGARCH(1,1) model

Consider the special case of an AHP-INGARCH(1,1) model. With arguments similar to those in Example 1 of Weiß [31], we obtain the unconditional variance of the AHP-INGARCH(1,1) model as

$$V(Y_t) = \left(\mu_Y + \mu_Y^2 \frac{\gamma-1}{\gamma+1} \right) \left(\frac{(\gamma+1)(1-\beta_1^2-2\alpha_1\beta_1)}{(\gamma+1)(1-\beta_1^2-2\alpha_1\beta_1)-2\gamma\alpha_1^2} \right),$$

with the variance of the conditional mean given by $V(\mu_t) = \frac{\alpha_1^2 \mu_Y (\gamma+1) + \alpha_1^2 \mu_Y^2 (\gamma-1)}{(\gamma+1)(1-\beta_1^2-2\alpha_1\beta_1)-2\gamma\alpha_1^2}$.

The autocovariance is

$$\begin{aligned} \gamma_Y(k) &= \alpha_1 \gamma_Y(k-1) + \beta_1 \gamma_Y(k-1) = (\alpha_1 + \beta_1)^{k-1} \gamma_Y(1) \\ &= (\alpha_1 + \beta_1)^{k-1} \left(\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma-1}{\gamma+1} \right) \left[\frac{(\gamma+1)(1-\beta_1^2-\alpha_1\beta_1)}{(\gamma+1)(1-\beta_1^2-2\alpha_1\beta_1)-2\gamma\alpha_1^2} \right], \quad k \geq 1, \end{aligned}$$

giving the autocorrelations as

$$\rho_Y(k) = (\alpha_1 + \beta_1)^{k-1} \frac{\alpha_1(1-\beta_1^2-\alpha_1\beta_1)}{(1-\beta_1^2-2\alpha_1\beta_1)} = (\alpha_1 + \beta_1)^{k-1} \frac{\alpha_1[1-\beta_1(\alpha_1+\beta_1)]}{1-(\alpha_1+\beta_1)^2+\alpha_1^2}, \quad k \geq 1.$$

The proof of this special case is provided in Appendix A.3.

Corollary 3.1. *Suppose that $\{Y_t\}$ following the AHP-INARCH(p) model with $q = 0$ in model (3.1) is second-order stationary, then the autocovariance function $\gamma_Y(k)$ satisfies the equation*

$$\gamma_Y(k) = \sum_{i=1}^p \alpha_i \gamma_Y(|k-i|), \quad k \geq 1.$$

3.2. AHP-INARCH(1) model

Consider the AHP-INARCH(1) model. The unconditional mean and unconditional variance are given by respectively,

$$E(Y_t) = \mu_Y = \frac{\alpha_0}{1-\alpha_1},$$

and

$$V(Y_t) = \sigma_Y^2 = \frac{\alpha_0}{(1-\alpha_1)^2} \left[\frac{(\gamma+1)(1-\alpha_1)+\alpha_0(\gamma-1)}{\gamma+1-2\gamma\alpha_1^2} \right].$$

From Corollary 3.1, one immediately obtains the autocovariance function of the AHP-INARCH(1) model as

$$\begin{aligned} \gamma_Y(k) &= (\alpha_1)^{k-1} \gamma_Y(1) = (\alpha_1)^{k-1} \left(\frac{\alpha_0}{(1-\alpha_1)^2} \right) \left[\frac{\alpha_1(\gamma+1)(1-\alpha_1)+\alpha_1\alpha_0(\gamma-1)}{\gamma+1-2\gamma\alpha_1^2} \right] \\ &= (\alpha_1)^k \left(\frac{\alpha_0}{(1-\alpha_1)^2} \right) \left(\frac{(\gamma+1)(1-\alpha_1)+\alpha_0(\gamma-1)}{\gamma+1-2\gamma\alpha_1^2} \right) = (\alpha_1)^k \gamma_Y(0). \end{aligned}$$

Hence, the autocorrelation function of the AHP-INARCH(1) model is

$$\rho_Y(k) = \frac{\gamma_Y(k)}{\gamma_Y(0)} = (\alpha_1)^k,$$

like in the standard AR(1) case.

4. Maximum likelihood estimation

In this section, we will discuss the maximum likelihood estimation (MLE) for the AHP-INGARCH (p, q) model (3.1), that is

$$Y_t | \mathcal{F}_{t-1} \sim AHP(\theta_t, \gamma),$$

$$E[Y_t | \mathcal{F}_{t-1}] = \frac{\theta_t}{\gamma} = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \frac{\theta_{t-j}}{\gamma}.$$

Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)^T$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^T$, $\boldsymbol{\lambda}^* = (\alpha_0, \boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)^T$, $\boldsymbol{\lambda} = (\gamma, \boldsymbol{\lambda}^{*T})^T$, and write the true value of $\boldsymbol{\lambda}$ as $\boldsymbol{\lambda}^0$. Suppose that the observation $\mathbf{Y} = (Y_1, \dots, Y_n)$ is fitted with the model (3.1). The conditional likelihood function at time t is

$$l_t(\boldsymbol{\lambda}) = f(y_t | \mathcal{F}_{t-1}; \boldsymbol{\lambda}) = \frac{\theta_t^{Y_t}}{(\gamma)_{Y_t}} \phi(1 + Y_t; \gamma + Y_t; -\theta_t),$$

where $\theta_t = \gamma \alpha_0 + \gamma \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \theta_{t-j}$. The conditional log-likelihood function is

$$l(\boldsymbol{\lambda}) = \ln \prod_{t=2}^n l_t(\boldsymbol{\lambda}) = \sum_{t=2}^n \{Y_t \ln \theta_t - \ln \Gamma(\gamma + Y_t) + \ln \Gamma(\gamma) + h(\theta_t)\},$$

where $h(\theta_t) = \ln[\phi(1 + Y_t; \gamma + Y_t; -\theta_t)]$. The score function is defined by

$$\frac{\partial l(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \sum_{t=2}^n \frac{\partial}{\partial \boldsymbol{\lambda}} \ln l_t(\boldsymbol{\lambda}),$$

with

$$\frac{\partial \ln l_t(\boldsymbol{\lambda})}{\partial \gamma} = \frac{Y_t}{\theta_t} \frac{\partial \theta_t}{\partial \gamma} - \psi(\gamma + Y_t) + \psi(\gamma) + \frac{\partial}{\partial \gamma} h(\theta_t),$$

$$\frac{\partial \ln l_t(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}^*} = \left[\frac{Y_t}{\theta_t} + \frac{\partial}{\partial \theta_t} h(\theta_t) \right] \frac{\partial \theta_t}{\partial \boldsymbol{\lambda}^*},$$

$$\frac{\partial \theta_t}{\partial \gamma} = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \frac{\partial \theta_{t-j}}{\partial \gamma}, \quad \frac{\partial \theta_t}{\partial \alpha_0} = \gamma + \sum_{j=1}^q \beta_j \frac{\partial \theta_{t-j}}{\partial \alpha_0},$$

$$\frac{\partial \theta_t}{\partial \alpha_i} = \gamma Y_{t-i} + \sum_{j=1}^q \beta_j \frac{\partial \theta_{t-j}}{\partial \alpha_i}, \quad i = 1, \dots, p,$$

$$\frac{\partial \theta_t}{\partial \beta_j} = \theta_{t-j} + \sum_{k=1}^q \beta_k \frac{\partial \theta_{t-k}}{\partial \beta_j}, j = 1, \dots, q.$$

where $\psi(x) = \frac{\partial}{\partial x} [\ln \Gamma(x)] = \frac{1}{\Gamma(x)} \frac{\partial}{\partial x} [\Gamma(x)]$ is the digamma function ([20]). The solution of the equation $\frac{\partial l(\lambda)}{\partial \lambda} = 0$, if it exists, gives the conditional MLE of λ , denoted by $\hat{\lambda}$.

The Hessian matrix is given by

$$H_n(\lambda) = - \sum_{t=2}^n \frac{\partial^2}{\partial \lambda \partial \lambda^T} \ln l_t(\lambda)$$

with

$$\frac{\partial^2 \ln l_t(\lambda)}{\partial \gamma^2} = \frac{y_t}{\theta_t} \frac{\partial^2 \theta_t}{\partial \gamma^2} - \frac{y_t}{\theta_t^2} \left(\frac{\partial \theta_t}{\partial \gamma} \right)^2 - \psi'(\gamma + y_t) + \psi'(\gamma) + \frac{\partial^2}{\partial \gamma^2} h(\theta_t),$$

$$\frac{\partial^2 \ln l_t(\lambda)}{\partial \gamma \partial \lambda^*} = \left[-\frac{y_t}{\theta_t^2} \frac{\partial \theta_t}{\partial \gamma} + \frac{\partial^2}{\partial \gamma \partial \theta_t} h(\theta_t) \right] \frac{\partial \theta_t}{\partial \lambda^*} + \left[\frac{y_t}{\theta_t} + \frac{\partial}{\partial \theta_t} h(\theta_t) \right] \frac{\partial^2 \theta_t}{\partial \gamma \partial \lambda^*},$$

$$\frac{\partial^2 \ln l_t(\lambda)}{\partial \lambda^* \partial \lambda^{*T}} = \left[-\frac{y_t}{\theta_t^2} + \frac{\partial^2}{\partial \theta_t^2} h(\theta_t) \right] \frac{\partial \theta_t}{\partial \lambda^*} \frac{\partial \theta_t}{\partial \lambda^{*T}} + \left[\frac{y_t}{\theta_t} + \frac{\partial}{\partial \theta_t} h(\theta_t) \right] \frac{\partial^2 \theta_t}{\partial \lambda^* \partial \lambda^{*T}},$$

$$\frac{\partial^2 \theta_t}{\partial \alpha_0^2} = 0, \quad \frac{\partial^2 \theta_t}{\partial \alpha_i^2} = 0, \quad \frac{\partial^2 \theta_t}{\partial \alpha_0 \partial \alpha_i} = 0, \quad i = 1, \dots, p,$$

$$\frac{\partial^2 \theta_t}{\partial \alpha_0 \partial \beta_j} = \frac{\partial \theta_{t-j}}{\partial \alpha_0} + \sum_{k=1}^q \beta_k \frac{\partial^2 \theta_{t-k}}{\partial \alpha_0 \partial \beta_j}, \quad \frac{\partial^2 \theta_t}{\partial \beta_j^2} = 2 \frac{\partial \theta_{t-j}}{\partial \beta_j} + \sum_{k=1}^q \beta_k \frac{\partial^2 \theta_{t-k}}{\partial \beta_j^2}, \quad j = 1, \dots, q,$$

$$\frac{\partial^2 \theta_t}{\partial \alpha_i \partial \beta_j} = \frac{\partial \theta_{t-j}}{\partial \alpha_i} + \sum_{k=1}^q \beta_k \frac{\partial^2 \theta_{t-k}}{\partial \alpha_i \partial \beta_j}, \quad i = 1, \dots, p, \quad j = 1, \dots, q,$$

$$\frac{\partial^2 \theta_t}{\partial \beta_j \partial \beta_k} = \frac{\partial \theta_{t-j}}{\partial \beta_k} + \frac{\partial \theta_{t-k}}{\partial \beta_j} + \sum_{l=1}^q \beta_l \frac{\partial^2 \theta_{t-l}}{\partial \beta_k \partial \beta_j}, \quad j, k = 1, \dots, q$$

where $\psi'(x) = \frac{\partial}{\partial x} [\psi(x)]$ is the trigamma function. According to White [35], the standard errors of $\hat{\lambda}$ can be computed from the robust sandwich matrix $H_n^{-1}(\hat{\lambda}) S_n(\hat{\lambda}) H_n^{-1}(\hat{\lambda})$, where $S_n(\lambda) = \sum_{t=2}^n \frac{\partial \ln l_t(\lambda)}{\partial \lambda} \frac{\partial \ln l_t(\lambda)}{\partial \lambda^T}$, and $H_n(\lambda) = - \sum_{t=2}^n \frac{\partial^2 \ln l_t(\lambda)}{\partial \lambda \partial \lambda^T}$.

5. Monte Carlo simulation study

Since the autocorrelation function of AHP-INGARCH(1,1) model is derived, the parameter of AHP-INGARCH(1,1) model can also be estimated according to the Yule-Walker approach. It is of interest to compare the efficiency of the Yule-Walker (YW) estimator with the maximum likelihood (ML) estimator. In this section, we conduct a Monte Carlo simulation study to investigate the performance of the YW and ML estimators for the proposed AHP-INGARCH(1,1) model. The simulations are computed by using the R programming language. The data set Y_1, Y_2, \dots, Y_n is generated in accordance to model (3.1) for different sample sizes n and different parameter values. Due to space constraints, we showed only the simulation results for the data with fixed mean equal to one and parameters $(\alpha_0 = 0.6, \alpha_1 = 0.3, \beta_1 = 0.1)$ for several sample sizes ($n = 100, 200$ and 500) with

over-dispersion ($\gamma = 2$), equi-dispersion ($\gamma = 1$) and under-dispersion ($\gamma = 0.8$) in Table 1. Similar results are obtained for other sets of parameter values considered. The parameter estimates and mean square errors for the parameters $(\alpha_0, \alpha_1, \beta_1, \gamma)$ are computed over 1000 replications. In the MLE estimation, we use sample mean as the initial mean value, namely $\mu_1 = \frac{1}{n} \sum_{i=1}^n Y_i$.

Based on the results from the simulation study, the parameter estimates approach the true values along with small mean square errors at all three types of dispersions as the sample size increases.

Table 1. Yule-Walker (YW) and maximum likelihood estimates (MLE) for simulated AHP-INGARCH(1, 1) model with $(\alpha_0 = 0.6, \alpha_1 = 0.3, \beta_1 = 0.1, \gamma)$ and different sample sizes n over 1000 replications (Mean squared errors in parentheses).

Dispersion	n	Estimates				
			$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\gamma}$
$\gamma = 2$ (over-dispersion)	100	YW	0.6688 (0.1866)	0.2825 (0.0188)	0.0486 (0.2298)	2.0157 (1.3083)
		MLE	0.5699 (0.0387)	0.2854 (0.0113)	0.1471 (0.0411)	2.0355 (0.9536)
	200	YW	0.6686 (0.0913)	0.2957 (0.0177)	0.0339 (0.1070)	2.0541 (0.6345)
		MLE	0.5905 (0.0249)	0.2907 (0.0070)	0.1193 (0.0262)	2.0767 (0.5949)
	500	YW	0.6189 (0.0257)	0.2910 (0.0036)	0.0859 (0.0290)	2.0191 (0.1926)
		MLE	0.5953 (0.0121)	0.2931 (0.0026)	0.1083 (0.0128)	2.0237 (0.1751)
$\gamma = 1$ (equi-dispersion)	100	YW	0.6755 (0.2987)	0.2842 (0.0386)	0.0327 (0.3618)	1.0375 (0.1024)
		MLE	0.5794 (0.0388)	0.2688 (0.0114)	0.1472 (0.0413)	1.0436 (0.0871)
	200	YW	0.6670 (0.0868)	0.2798 (0.0095)	0.0499 (0.0924)	1.0186 (0.0490)
		MLE	0.5932 (0.0263)	0.2763 (0.0059)	0.1278 (0.0281)	1.0194 (0.0354)
	500	YW	0.6159 (0.0246)	0.2926 (0.0028)	0.0925 (0.0287)	1.0097 (0.0149)
		MLE	0.5947 (0.0130)	0.2923 (0.0024)	0.1145 (0.0145)	1.0130 (0.0128)
$\gamma = 0.8$ (under-dispersion)	100	YW	0.6566 (0.1908)	0.2841 (0.0255)	0.0358 (0.2377)	0.8560 (0.0614)
		MLE	0.5611 (0.0353)	0.2759 (0.0107)	0.1418 (0.0379)	0.8513 (0.0387)
	200	YW	0.6418 (0.0710)	0.2828 (0.0125)	0.0510 (0.0877)	0.8205 (0.0192)
		MLE	0.5801 (0.0230)	0.2776 (0.0055)	0.1212 (0.0254)	0.8284 (0.0166)
	500	YW	0.6121 (0.0228)	0.2878 (0.0023)	0.0837 (0.0281)	0.8165 (0.0076)
		MLE	0.5892 (0.0110)	0.2880 (0.0021)	0.1073 (0.0128)	0.8185 (0.0064)

6. Comparison of INGARCH models

This section describes the analysis of three real-life data sets fitted by the Poisson, NB, DP, GP, COMP, and AHP distributed INGARCH-type modes. Based on the sample mean and variance, the first data set exhibits over-dispersion and the last two data sets exhibit under-dispersion. The MLE

results are obtained using the R software through the Box constraints optimization (L-BFGS-B) function with the initial conditional mean equals to the sample mean. Statistical consistency between the predictive and observed distributions is investigated by plotting the histogram of probability integral transform (PIT) (see, Dawid [10]). According to Czado et al. [9], from the PIT histogram, under-dispersed predictive distribution is signified by a U-shaped histogram whereas over-dispersed predictive distribution is indicated by a hump or inverse-U shaped histogram. When central tendencies are biased, a skewed histogram is observed.

6.1. Polio series

The polio data is a time series of length 168, which has been fitted by Zhu [39] using the NB-INGARCH model. The data series is the monthly number of cases of poliomyelitis reported by the U.S. Centers for Disease Control from 1970 to 1983 with a sample mean of 1.33 and a sample variance of 3.50. These statistics indicate that the series is over-dispersed and the sample first-order autocorrelation coefficient (FOAC), $\hat{\rho}(1) = 0.2948$.

Figure 3 describes the original data, autocorrelation function (ACF) and partial autocorrelation function (PACF) of the polio series while Table 2 summarizes the result of six fitted INGARCH(1,1) models.

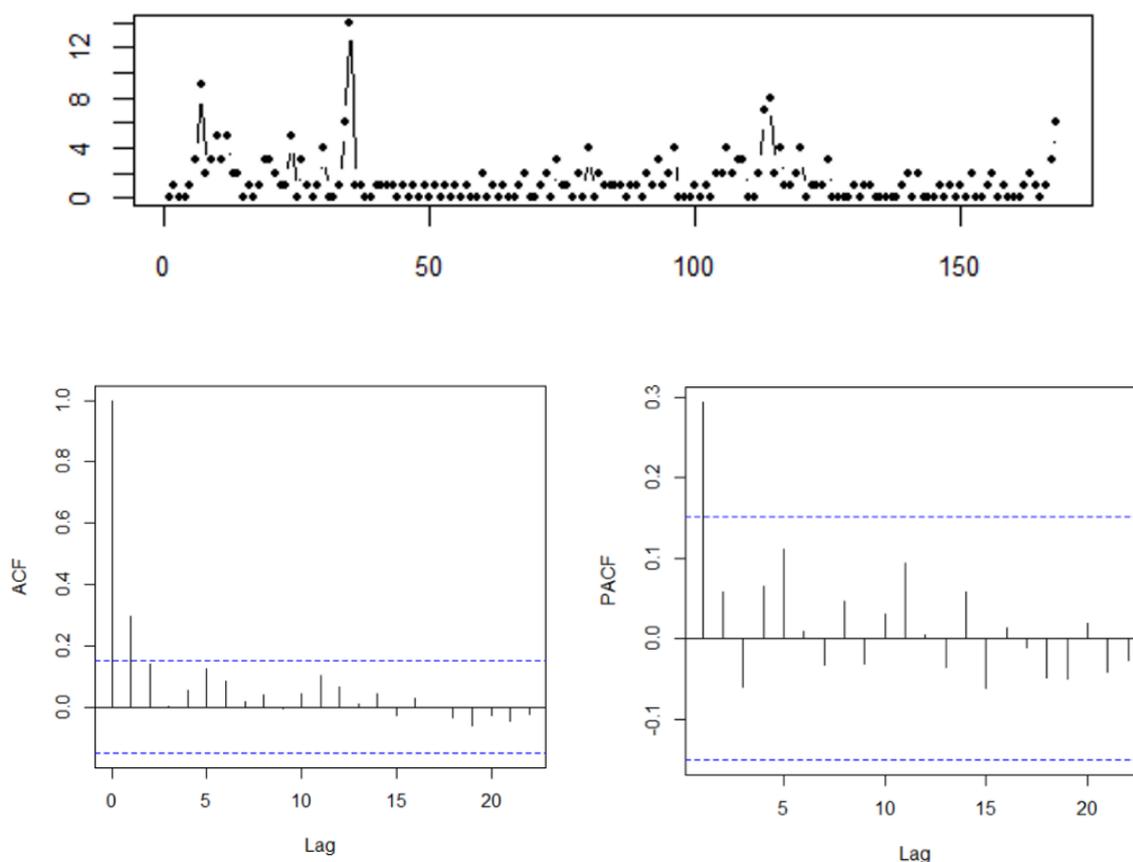


Figure 3. Plots of time series, ACF and PACF of polio cases from 1970 to 1983.

Table 2. Polio series: Parameter estimates with Poisson-, NB-, DP-, GP-, COMP-, and AHP-INGARCH(1,1) models, standard errors are shown in parentheses.

Model	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{r}/\hat{\gamma}/\hat{\phi}/\hat{\nu}/\hat{\gamma}$	AIC	BIC
Poisson	0.6357 (0.1702)	0.3515 (0.0678)	0.1846 (0.1342)		562.08	571.40
NB	0.6075 (0.2275)	0.3643 (0.1029)	0.1982 (0.1858)	1.6346 (0.4326)	520.47	532.87
DP	0.6357 (0.2278)	0.3515 (0.0907)	0.1846 (0.1796)	0.5585 (0.0611)	529.33	541.73
GP	0.3645 (0.4105)	0.1647 (0.0859)	0.5689 (0.3497)	1.4089 (0.1083)	528.08	540.48
COMP	0.0529 (0.0399)	0.1845 (0.0713)	0.1670 (0.1896)	0.2546 (0.0524)	524.37	536.77
AHP	0.6418 (0.2063)	0.4214 (0.1082)	0.1344 (0.1536)	4.1310 (2.0243)	521.15	533.55

Based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC) for all the models, the AHP-INGARCH(1,1) model is found to be comparable to the NB-INGARCH(1,1) but better than the Poisson, DP-, GP- and COMP-INGARCH(1,1) models for this over-dispersed polio dataset. It is worth noting that the values of the parameter estimates for NB-INGARCH(1,1) model are different from Zhu [39] because they fixed the parameter $r = 2$.

Figure 4 shows the PIT histogram of the six INGARCH(1,1) models. Overall, Poisson-INGARCH(1,1) model have a U-shaped histogram while DP-INGARCH(1,1) model has a skewed histogram. Although not uniform, the NB-, GP-, COMP- and AHP-INGARCH(1,1) have only a slight hump in the histogram indicating better performance than the Poisson- and DP-INGARCH(1,1) models.

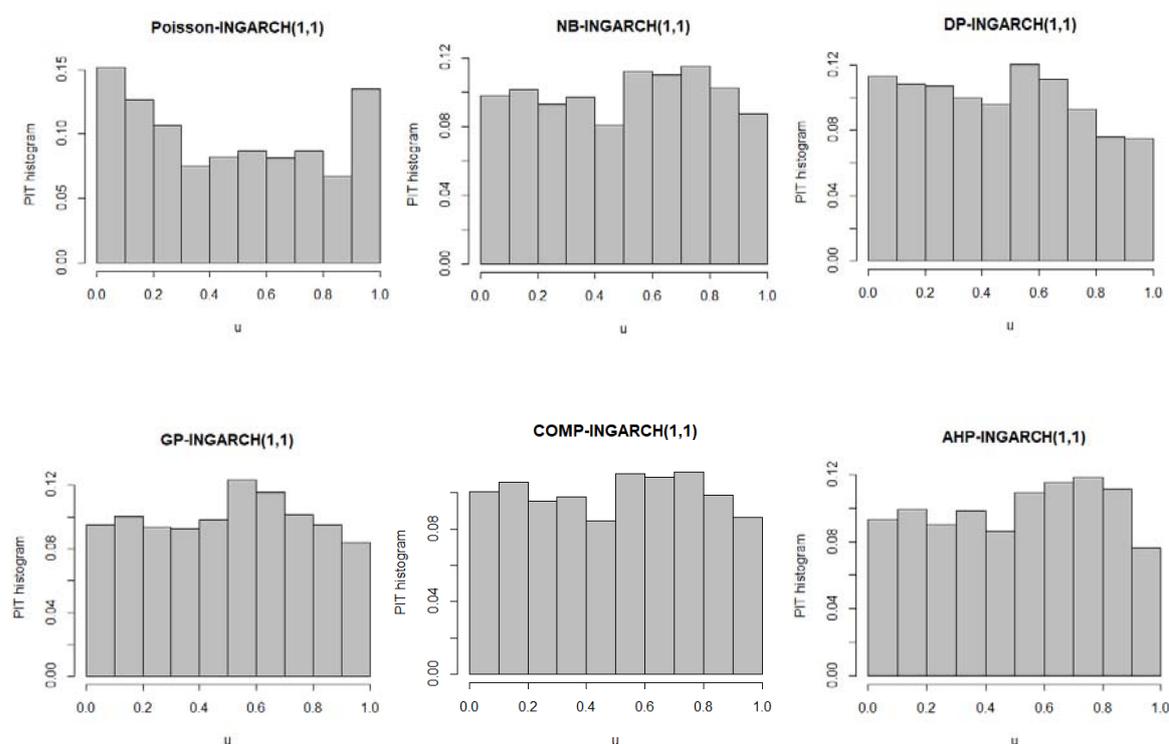


Figure 4. PIT histogram for Poisson-, NB-, DP-, GP-, COMP- and AHP-INGARCH (1,1) models of polio series.

Comparison of the estimated means, variances, dispersion index and FOACs within the fitted Poisson, NB, DP, GP, COMP and AHP models for polio data and another two under-dispersion data sets that will be discussed later in the next sections are summarized in Table 3. The dispersion index for NB- and AHP-INGARCH(1,1) models are closest to the dispersion index for the polio data compared to the Poisson-, DP-, GP- and COMP-INGARCH(1,1) models.

Table 3. Sample and estimated mean, variance, dispersion index (I_Z) and FOAC under the Poisson, NB, DP, GP, COMP and AHP models.

Data	Model	Sample	Poisson	NB	DP	GP	COMP	AHP
Polio	Mean	1.3333	1.3701	1.3888	1.3701	1.3683	1.9624	1.4449
	Variance	3.5050	1.6076	3.4813	2.8786	2.8756	2.0349	4.0534
	I_Z	2.6288	1.1733	2.5067	2.1010	2.1016	1.0370	2.8053
	FOAC	0.2948	0.3787	0.3966	0.3787	0.1963	0.1908	0.4489
IP	Mean	1.2863	1.2917	1.2916	1.2917	1.2917	1.3154	1.2913
	Variance	1.2052	1.4039	1.4039	1.4589	1.1777	1.2184	1.2806
	I_Z	0.9370	1.0869	1.0869	1.1294	0.9117	0.9263	0.9917
	FOAC	0.2925	0.2827	0.6927	0.2827	0.2834	0.2825	0.2830
COVID-19	Mean	0.8416	0.5032	0.5045	0.5045	0.4001	0.7396	0.6022
	Variance	0.7347	0.5814	0.5829	0.4981	0.3315	0.5985	0.5547
	I_Z	0.8729	1.1555	1.1554	0.9873	0.8285	0.8091	0.9212
	FOAC	0.2374	0.2082	0.2081	0.2081	0.1891	0.2383	0.2993

6.2. Internet Protocol (IP) count series

The IP counts data is a time series of length 241 that gives the number of different IP addresses registered within periods of 2-min length at the server of the Department of Statistics of the University of Würzburg in 29 November 2005 between 10 a.m and 6 p.m. This data set has been investigated by Weiß [30] and Zhu [40,41]. The data examined is under-dispersed since the variance (1.2052) is smaller than the mean (1.2863) with the sample FOAC of 0.2925. Figure 5 presents the IP counts time series, ACF and PACF, respectively.

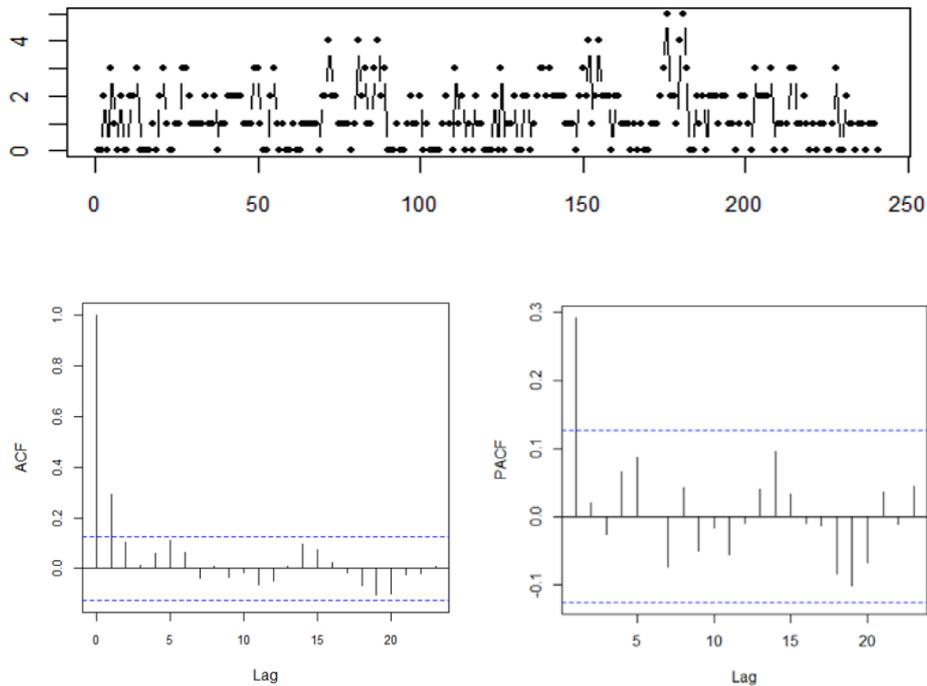


Figure 5. Plots of time series, ACF and PACF of IP counts series.

Zhu [41] fitted the Poisson INAR(1), the Poisson-, DP- and GP- INARCH(1) and INGARCH(1,1) models to the data and claimed that the GP-INARCH(1) model is more appropriate for this time series. Comparison of existing INARCH(1) and AHP-INARCH(1) models are summarized in Table 4. It is worth noting that the over-dispersed NB-INARCH(1) model of Zhu [39] is added in our analysis for comparison. As seen from Table 4, the estimated value of the NB-INARCH(1) parameter \hat{r} is large and significant, but unreliable. This is because the differentiation of the likelihood function with respect to the parameter r is problematic as reported in Zhu [39]. We obtain the same conclusion as Zhu [40], that is, (i) the estimated dispersion index of GP-, COMP- and AHP-INARCH(1) models are less than 1, which indicate that the data set is under-dispersed count time series, and Poisson-, DP- and NB-INARCH(1) models wrongly indicate that the data set is over-dispersed; (ii) GP model performs better in AIC and Poisson INARCH(1) performs better in BIC, but the difference of AIC and BIC values between other models are rather small. Typically, a difference in AIC value less than 2 is considered not significant (Zhu [40]). Based on Figure 6, the PIT histogram of the AHP, GP and COMP models shows approximate uniformity, while Poisson-, NB- and DP-INARCH(1) are slightly hump-shaped.

Table 4. IP counts series: Parameter estimates with Poisson-, NB-, DP-, GP-, COMP- and AHP-INARCH(1) models, standard errors are shown in parentheses.

Model	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{r}/\hat{\nu}/\hat{\phi}/\hat{\nu}/\hat{\nu}$	AIC	BIC
Poisson	0.9265 (0.1007)	0.2827 (0.0684)		675.20	682.15
NB	0.9264 (0.1007)	0.2828 (0.0684)	39722.23 (253.85)	677.20	687.62
DP	0.9265 (0.1026)	0.2827 (0.0697)	0.9623 (0.0878)	677.02	687.44
GP	0.9256 (0.0917)	0.2834 (0.0622)	0.9157 (0.0409)	673.78	684.19
COMP	1.0492 (0.1157)	0.2825 (0.0636)	1.2674 (0.1790)	674.81	685.22
AHP	0.9258 (0.0974)	0.2830 (0.0653)	0.8796 (0.0937)	675.90	686.32

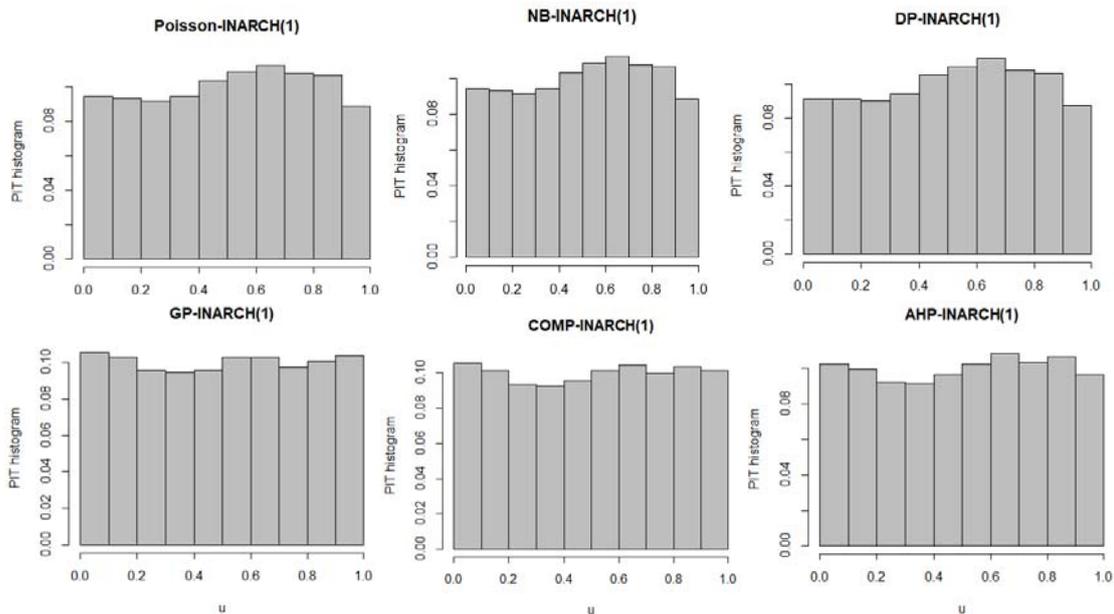


Figure 6. PIT histogram for Poisson-, NB-, DP-, GP-, COMP- and AHP-INARCH (1) models of IP counts series.

6.3. COVID-19 series

The COVID-19 data is a time series of length 101 that gives the number of daily COVID-19 new deaths recorded in Saudi Arabia between 26th January 2023 to 6th May 2023. This data is publicly available at the website <https://ourworldindata.org/>. The data examined is under-dispersed since the variance (0.7347) is smaller than the mean (0.8416) with the sample FOAC of 0.2374. Figure 7 presents the COVID-19 new deaths time series, ACF and PACF, respectively.

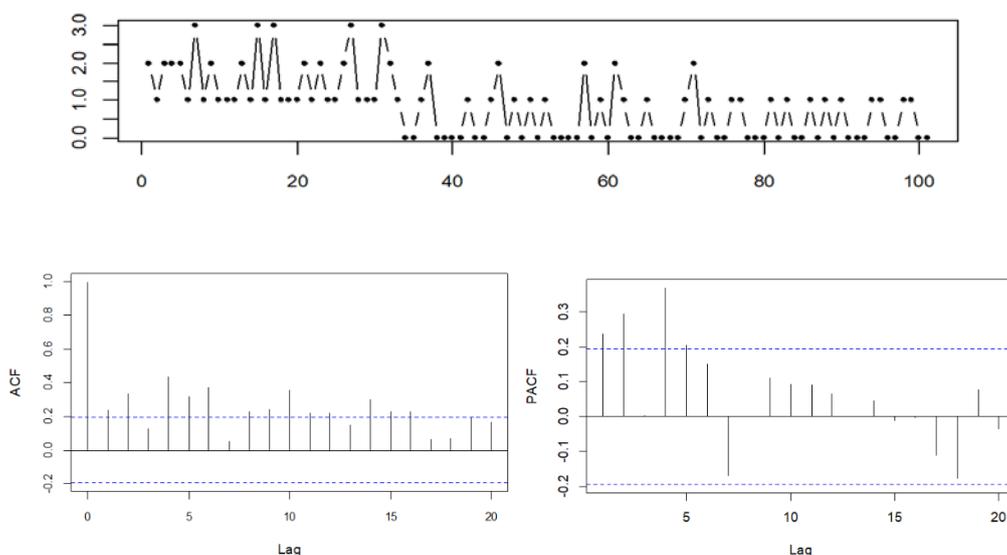


Figure 7. Plots of time series, ACF and PACF of daily COVID-19 new deaths.

Table 5 summarizes the results of six INGARCH(1,1) models for the COVID-19 new deaths data. Based on AIC and BIC, the AHP-INGARCH(1,1) model is the best. The estimated value of the NB-INGARCH(1,1) parameter \hat{r} is again large and significant, but unreliable. Besides the over-dispersed Poisson- and NB-INGARCH(1,1) models, the dispersion index of DP-, GP-, COMP- and AHP-INGARCH(1,1) are all less than 1 which indicate that these models are capable of capturing under-dispersed feature from the empirical data (see Table 3). All the PIT histograms in Figure 8 visibly deviate from uniformity (with a slight preference for the AHP-INGARCH(1,1) model). In particular, the PIT histograms of Poisson-, NB- and DP-INGARCH(1,1) models are slightly hump-shaped while GP-, COMP- and AHP-INGARCH(1,1) models are slightly U-shaped.

Table 5. Daily COVID-19 new deaths: Parameter estimates with Poisson-, NB-, DP-, GP-, COMP-, and AHP-INGARCH(1,1) models, standard errors are shown in parentheses.

Model	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{r}/\hat{\gamma}/\hat{\phi}/\hat{v}/\hat{\gamma}$	AIC	BIC
Poisson	0.0130 (0.0268)	0.0890 (0.0472)	0.8851 (0.0645)		223.01	230.76
NB	0.0131 (0.0268)	0.0891 (0.0472)	0.8850 (0.0646)	16860.79 (267.79)	225.01	235.31
DP	0.0131 (0.0248)	0.0891 (0.0436)	0.8850 (0.0597)	1.1702 (0.1655)	223.83	234.13
GP	0.0075 (0.0225)	0.0736 (0.0360)	0.9077 (0.0537)	0.8503 (0.0502)	219.61	229.91
COMP	0.0520 (0.0413)	0.1012 (0.0423)	0.8710 (0.0580)	1.9497 (0.4518)	219.35	229.64
AHP	0.0172 (0.0265)	0.1220 (0.0501)	0.8493 (0.0639)	0.4912 (0.1070)	217.55	227.85

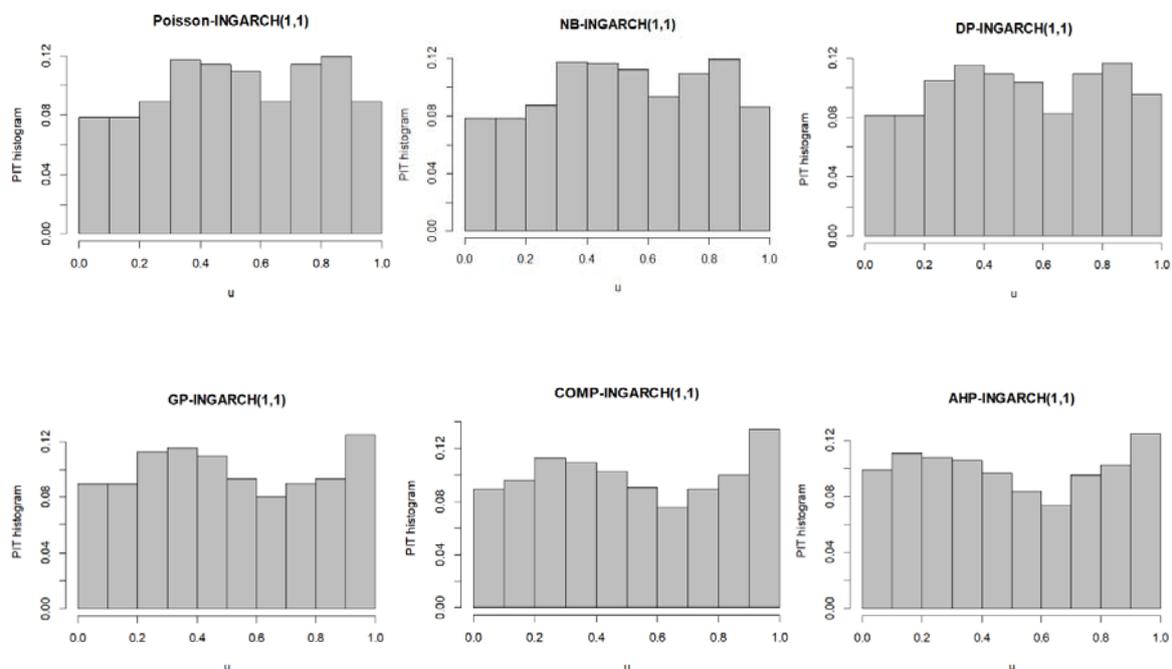


Figure 8. PIT histogram for Poisson-, NB-, DP-, GP-, COMP- and AHP-INGARCH(1,1) models of daily COVID-19 new deaths series.

7. Conclusions

This work introduces a new family of AHP-INGARCH models for analysing over-dispersed and under-dispersed count time series data. The advantages of the AHP-INGARCH-type model are (i) the ease in obtaining the model mean and variance compared to distributions like the COMP and DP distributions, making the AHP distribution a useful choice for an INGARCH model, and (ii) the ability to accommodate a wider range of dispersion, placing the AHP distribution as a flexible model for applications. The application of the AHP-INGARCH models to three real-life data sets clearly demonstrates the model competitiveness in studying both over-dispersed and under-dispersed data. Since γ is the dispersion parameter of AHP distribution, one of the potential future research is to relax the assumption of constant dispersion by allowing time-varying dispersion. Another direction of future research is to extend the results to multivariate cases.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgements

The authors would like to thank the editor and the referees for their valuable comments.

Seng Huat Ong is supported by the Ministry of Higher Education grant FRGS/1/2020/STG06/SYUC/02/1 and UCSI University grant REIG-FBM-2022/050. The authors would like to thank INTI International University for funding the publication of this work.

Conflict of interest

All authors declare no conflicts of interest in this paper.

Appendix A

A.1 Parameter constraint of AHP distribution

Let $Z \sim AHP(\theta, \gamma)$. The probability mass function of Z can be written as

$$P(Z = z) = \frac{\theta^z}{(\gamma)_z} \phi(1 + z; \gamma + z; -\theta) = \frac{\theta^z e^{-\theta}}{(\gamma)_z} \phi(\gamma - 1; \gamma + z; \theta)$$

where the second equality is obtained from Johnson et al. [20]. Since $\frac{\theta^z e^{-\theta}}{(\gamma)_z} > 0$ for $\theta > 0$ and $\gamma > 0$. The positivity of the pmf of Z depends on the confluent hypergeometric function

$$\phi(\gamma - 1; \gamma + z; \theta) = \sum_{k=0}^{\infty} \frac{(\gamma - 1)_k \theta^k}{(\gamma + z)_k k!}.$$

It is easy to see that $\phi(\gamma - 1; \gamma + z; \theta) > 0$ when $\gamma \geq 1$. Hence, the positivity of the pmf of Z is guaranteed for $\gamma \geq 1$.

For $0 < \gamma < 1$, we note that $(\gamma + z)_k \geq (\gamma)_k$ for all $z \geq 0$ and $k \geq 0$. Since

$$\phi(\gamma - 1; \gamma + z; \theta) = \sum_{k=0}^{\infty} \frac{(\gamma - 1)_k \theta^k}{(\gamma + z)_k k!} \geq \sum_{k=0}^{\infty} \frac{(\gamma - 1)_k \theta^k}{(\gamma)_k k!} = \phi(\gamma - 1; \gamma; \theta),$$

we have

$$P(Z = z) = \frac{\theta^z e^{-\theta}}{(\gamma)_z} \phi(\gamma - 1; \gamma + z; \theta) \geq P(Z = 0) = e^{-\theta} \phi(\gamma - 1; \gamma; \theta)$$

for all $z \geq 1$. Thus, for $0 < \gamma < 1$, the positivity of the pmf of Z is guaranteed as long as $P(Z = 0) > 0$. This condition can be further simplified as

$$\begin{aligned} \phi(\gamma - 1; \gamma; \theta) &= \sum_{k=0}^{\infty} \frac{(\gamma - 1)_k \theta^k}{(\gamma)_k k!} = \sum_{k=0}^{\infty} \frac{\Gamma(\gamma - 1 + k) \Gamma(\gamma) \theta^k}{\Gamma(\gamma - 1) \Gamma(\gamma + k) k!} \\ &= (\gamma - 1) \sum_{k=0}^{\infty} \frac{\theta^k}{(\gamma - 1 + k) k!} > 0 \end{aligned}$$

It is worth to note that the confluent hypergeometric function $\phi(\gamma - 1; \gamma; \theta)$ is a decreasing function on θ . Hence, for a given $\gamma \in (0, 1)$, $\phi(\gamma - 1; \gamma; \theta) > 0$ iff $\theta < \theta^*$, where θ^* is the solution of $\phi(\gamma - 1; \gamma; \theta^*) = 0$.

A.2 Proof of Theorem 3.1

(i) Let $\mu_Y = E[Y_t]$ denote the unconditional expectation of Y_t if it exists. Then by the tower property of conditional expectation, we have

$$\begin{aligned} \mu_Y &= E[E[Y_t | \mathcal{F}_{t-1}]] = \alpha_0 + \sum_{i=1}^p \alpha_i E[Y_{t-i}] + \sum_{j=1}^q \beta_j E[\mu_{t-j}] = \alpha_0 + \sum_{i=1}^p \alpha_i \mu_Y + \sum_{j=1}^q \beta_j \mu_Y \\ &= \frac{\alpha_0}{\left(1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j\right)}, \end{aligned}$$

where $E[\mu_{t-j}] = E[E[Y_{t-j} | \mathcal{F}_{t-j-1}]] = \mu_Y$ by the tower property of conditional expectation.

(ii) By the definition of covariance and the tower property, we have

$$\begin{aligned} \text{Cov}[Y_t - \mu_t, \mu_{t-k}] &= E[E[(Y_t - \mu_t)(\mu_{t-k} - \mu_Y) | \mathcal{F}_{t-1}]] \\ &= E[(\mu_{t-k} - \mu_Y) \cdot E[Y_t - \mu_t | \mathcal{F}_{t-1}]] \\ &= E[(\mu_{t-k} - \mu_Y) \cdot (\mu_Y - \mu_Y)] \\ &= 0 \end{aligned}$$

for $k \geq 0$, where μ_t is \mathcal{F}_{t-1} measurable, i.e., $E[\mu_t|\mathcal{F}_{t-1}] = \mu_t$. The proof is completed by noticing that

$$\text{Cov}[Y_t - \mu_t, \mu_{t-k}] = \text{Cov}[Y_t, \mu_{t-k}] - \text{Cov}[\mu_t, \mu_{t-k}] = 0.$$

For $k < 0$, by the same argument, we have

$$\begin{aligned} \text{Cov}[Y_{t-k} - \mu_{t-k}, Y_t] &= E[E[(Y_{t-k} - \mu_{t-k})(Y_t - \mu_Y)|\mathcal{F}_{t-1}]] \\ &= E[(Y_t - \mu_Y) \cdot E[Y_{t-k} - \mu_{t-k}|\mathcal{F}_{t-k-1}]] \\ &= E[(Y_t - \mu) \cdot (\mu_{t-k} - \mu_{t-k})] \\ &= 0 \end{aligned}$$

Again, we use the property that μ_{t-k} is \mathcal{F}_{t-k-1} measurable, i.e., $E[\mu_{t-k}|\mathcal{F}_{t-k-1}] = \mu_{t-k}$ and $E[Y_{t-k}|\mathcal{F}_{t-k-1}] = \mu_{t-k}$. The proof is completed by noticing that

$$\text{Cov}[Y_{t-k} - \mu_{t-k}, Y_t] = \text{Cov}[Y_{t-k}, Y_t] - \text{Cov}[\mu_{t-k}, Y_t] = 0.$$

(iii) Finally, applying part (ii), we have for $k \geq 0$,

$$\begin{aligned} \gamma_\mu(k) &= \text{Cov}[\mu_t, \mu_{t-k}] = \text{Cov}\left[\left(\alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \mu_{t-j}\right), \mu_{t-k}\right] \\ &= \sum_{i=1}^p \alpha_i \cdot \text{Cov}[Y_{t-i}, \mu_{t-k}] + \sum_{j=1}^q \beta_j \cdot \text{Cov}[\mu_{t-j}, \mu_{t-k}] \\ &= \sum_{i=1}^{\min(k,q)} \alpha_i \cdot \text{Cov}[\mu_{t-i}, \mu_{t-k}] + \sum_{i=k+1}^p \alpha_i \cdot \text{Cov}[Y_{t-i}, Y_{t-k}] + \sum_{j=1}^q \beta_j \gamma_\mu(|k-j|) \\ &= \sum_{i=1}^{\min(k,q)} \alpha_i \cdot \gamma_\mu(|k-i|) + \sum_{i=k+1}^p \alpha_i \cdot \gamma_Y(i-k) + \sum_{j=1}^q \beta_j \gamma_\mu(|k-j|). \end{aligned}$$

Again, applying part (ii), we obtain for $k \geq 1$,

$$\begin{aligned} \gamma_Y(k) &= \text{Cov}[Y_t, Y_{t-k}] = \text{Cov}[\mu_t, Y_{t-k}] = \text{Cov}\left[\left(\alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \mu_{t-j}\right), Y_{t-k}\right] \\ &= \sum_{i=1}^p \alpha_i \cdot \text{Cov}[Y_{t-i}, Y_{t-k}] + \sum_{j=1}^q \beta_j \cdot \text{Cov}[\mu_{t-j}, Y_{t-k}] \\ &= \sum_{i=1}^p \alpha_i \cdot \gamma_Y(|k-i|) + \sum_{j=1}^{\min(k-1,q)} \beta_j \cdot \text{Cov}[Y_{t-j}, Y_{t-k}] + \sum_{j=k}^q \beta_j \cdot \text{Cov}[\mu_{t-j}, \mu_{t-k}] \\ &= \sum_{i=1}^p \alpha_i \gamma_Y(|k-i|) + \sum_{j=1}^{\min(k-1,q)} \beta_j \gamma_Y(k-j) + \sum_{j=k}^q \beta_j \gamma_\mu(j-k). \end{aligned}$$

A.3 Proof of Special Case AHP-INGARCH(1,1)

(i) *Unconditional variance*

From Eq (3.2), we obtain for $k \geq 2$ that

$$\gamma_Y(k) = \alpha_1 \gamma_Y(k-1) + \beta_1 \gamma_Y(k-1) = (\alpha_1 + \beta_1)^{k-1} \gamma_Y(1). \quad (*)$$

Again, from Eq (3.2), for $k = 1$, we have

$$\begin{aligned} \gamma_Y(1) &= \alpha_1 \gamma_Y(0) + \beta_1 \gamma_Y(0) = \alpha_1 V(Y_t) + \beta_1 V(\mu_t) \\ &= \alpha_1 (E[V(Y_t | \mathcal{F}_{t-1})] + V(E[Y_t | \mathcal{F}_{t-1}])) + \beta_1 V(\mu_t) \\ &= \alpha_1 \left(E \left[\mu_t \left(1 + \mu_t \left(\frac{\gamma-1}{\gamma+1} \right) \right) \right] + V(\mu_t) \right) + \beta_1 V(\mu_t) \\ &= \alpha_1 \left(E(\mu_t) + E[\mu_t^2] \frac{(\gamma-1)}{(\gamma+1)} \right) + (\alpha_1 + \beta_1) V(\mu_t) \\ &= \alpha_1 \mu_Y + \alpha_1 [V(\mu_t) + \mu^2] \frac{(\gamma-1)}{(\gamma+1)} + (\alpha_1 + \beta_1) V(\mu_t) \\ &= \alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{(\gamma-1)}{(\gamma+1)} + \left(\alpha_1 \frac{(\gamma-1)}{(\gamma+1)} + \alpha_1 + \beta_1 \right) V(\mu_t) \\ &= \alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{(\gamma-1)}{(\gamma+1)} + \left(\frac{2\alpha_1 \gamma + \beta_1 (\gamma+1)}{(\gamma+1)} \right) V(\mu_t) \end{aligned} \quad (**)$$

To determine an expression for $V(\mu_t)$, note first that for $k \geq 1$, Eq (3.3) can be written as

$$\gamma_\mu(k) = \alpha_1 \gamma_\mu(k-1) + \beta_1 \gamma_\mu(k-1) = (\alpha_1 + \beta_1)^k V(\mu_t).$$

For $k = 0$, Eq (3.3) can be written as

$$\begin{aligned} \gamma_\mu(0) &= V(\mu_t) = \alpha_1 \gamma_Y(1) + \beta_1 \gamma_\mu(1) \\ &= \alpha_1 \left[\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{(\gamma-1)}{(\gamma+1)} + \left(\alpha_1 \frac{(\gamma-1)}{(\gamma+1)} + \alpha_1 + \beta_1 \right) V(\mu_t) \right] + \beta_1 [(\alpha_1 + \beta_1) V(\mu_t)] \\ &= \alpha_1^2 \mu_Y + \alpha_1^2 \mu_Y^2 \frac{(\gamma-1)}{(\gamma+1)} + \left[\alpha_1^2 \frac{(\gamma-1)}{(\gamma+1)} + (\alpha_1 + \beta_1)^2 \right] V(\mu_t) \\ &= \left(\alpha_1^2 \mu_Y + \alpha_1^2 \mu_Y^2 \frac{(\gamma-1)}{(\gamma+1)} \right) / \left(1 - \alpha_1^2 \frac{(\gamma-1)}{(\gamma+1)} - (\alpha_1 + \beta_1)^2 \right) \\ &= \frac{\alpha_1^2 \mu_Y (\gamma + 1) + \alpha_1^2 \mu_Y^2 (\gamma - 1)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \end{aligned}$$

Therefore, the unconditional variance of Y_t can be obtained from

$$V(Y_t) = E[V(Y_t | \mathcal{F}_{t-1})] + V(E[Y_t | \mathcal{F}_{t-1}]) = E \left[\mu_t \left(1 + \mu_t \left(\frac{\gamma-1}{\gamma+1} \right) \right) \right] + V(\mu_t)$$

$$\begin{aligned}
&= \mu_Y + \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} + \left(\frac{2\gamma}{\gamma + 1} \right) V(\mu_t) \\
&= \left(\mu_Y + \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right) + \left(\frac{2\gamma}{\gamma + 1} \right) \frac{\alpha_1^2 \mu_Y (\gamma + 1) + \alpha_1^2 \mu_Y^2 (\gamma - 1)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \\
&= \left(\mu_Y + \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right) + \frac{2\gamma \alpha_1^2 \left(\mu_Y + \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \\
&= \left(\mu_Y + \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right) \left(1 + \frac{2\gamma \alpha_1^2}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \right) \\
&= \left(\mu_Y + \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right) \left(\frac{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \right).
\end{aligned}$$

(ii) *Autocovariance function*

Given that $V(\mu_t) = \frac{\alpha_1^2 \mu_Y (\gamma + 1) + \alpha_1^2 \mu_Y^2 (\gamma - 1)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} = \frac{\alpha_1 (\gamma + 1) (\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma - 1}{\gamma + 1})}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2}$ and from Eqs (*) and (**), we obtain for $k \geq 2$ that

$$\begin{aligned}
\gamma_Y(k) &= (\alpha_1 + \beta_1)^{k-1} \left[\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} + \left(\frac{2\alpha_1 \gamma + \beta_1 (\gamma + 1)}{\gamma + 1} \right) V(\mu_t) \right] \\
&= (\alpha_1 + \beta_1)^{k-1} \left[\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} + \left(\frac{2\alpha_1 \gamma + \beta_1 (\gamma + 1)}{\gamma + 1} \right) \left(\frac{\alpha_1 (\gamma + 1) (\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma - 1}{\gamma + 1})}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \right) \right] \\
&= (\alpha_1 + \beta_1)^{k-1} \left(\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right) \left[1 + \frac{2\gamma \alpha_1^2 + \alpha_1 \beta_1 (\gamma + 1)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \right] \\
&= (\alpha_1 + \beta_1)^{k-1} \left(\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right) \left[\frac{(\gamma + 1)(1 - \beta_1^2 - \alpha_1 \beta_1)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \right].
\end{aligned}$$

Note that for $k = 1$, we have

$$\gamma_Y(1) = \left(\alpha_1 \mu_Y + \alpha_1 \mu_Y^2 \frac{\gamma - 1}{\gamma + 1} \right) \left[\frac{(\gamma + 1)(1 - \beta_1^2 - \alpha_1 \beta_1)}{(\gamma + 1)(1 - \beta_1^2 - 2\alpha_1 \beta_1) - 2\gamma \alpha_1^2} \right].$$

References

1. G. E. Bardwell, E. L. Crow, A two-parameter family of hyper-Poisson distributions, *J. Amer. Stat. Assoc.*, **59** (1964), 133–141. <http://dx.doi.org/10.1080/01621459.1964.10480706>
2. T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, *J. Economet.*, **31** (1986), 307–327. [http://dx.doi.org/10.1016/0304-4076\(86\)90063-1](http://dx.doi.org/10.1016/0304-4076(86)90063-1)
3. M. Bourguignon, C. H. Weiß, An INAR (1) process for modeling count time series with equidispersion, underdispersion and overdispersion, *Test*, **26** (2017), 847–868. <http://dx.doi.org/10.1007/s11749-017-0536-4>
4. M. Bourguignon, J. Rodrigues, M. Santos-Neto, Extended Poisson INAR (1) processes with equidispersion, underdispersion and overdispersion, *J. Appl. Stat.*, **46** (2018), 101–118. <http://dx.doi.org/10.1080/02664763.2018.1458216>
5. A. C. Cameron, P. K. Trivedi, *Regression analysis of count Data*, 2 Eds., Cambridge University Press, New York, 2013. <https://doi.org/10.1017/CBO9781139013567>

6. P. C. Consul, *Generalized Poisson distributions: Properties and applications*, 1 Ed., Marcel Dekker, New York, 1989.
7. P. C. Consul, F. Famoye, Generalized Poisson regression model, *Commun. Stat. -Theory Methods.*, **21** (1992), 89–109. <http://dx.doi.org/10.1080/03610929208830766>
8. R. W. Conway, W. L. Maxwell, A queuing model with state dependent service rates, *J. Ind. Eng.*, **12** (1962), 132–136.
9. C. Czado, T. Gneiting, L. Held, Predictive model assessment for count data, *Biometrics*, **65** (2009), 1254–1261. <https://dx.doi.org/10.1111/j.1541-0420.2009.01191.x>
10. A. P. Dawid, Present position and potential developments: Some personal views: Statistical theory: The prequential approach, *J. Roy. Stat. Soc. A*, **147** (1984), 278–292. <https://dx.doi.org/10.2307/2981683>
11. B. Efron, Double exponential families and their use in generalized linear regression, *J. Amer. Stat. Assoc.*, **81** (1986), 709–721. <https://dx.doi.org/10.1080/01621459.1986.10478327>
12. R. F. Engle, J. R. Russell, Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica*, **66** (1998), 1127–1162. <https://dx.doi.org/10.2307/2999632>
13. F. Famoye, Restricted generalized Poisson regression model, *Commun. Stat. -Theory Methods.*, **22** (1993), 1335–1354. <https://dx.doi.org/10.1080/03610929308831089>
14. R. Ferland, A. Latour, D. Oraichi, Integer-valued GARCH processes. *J. Time Ser. Anal.*, **27** (2006), 923–942. <https://dx.doi.org/10.1111/j.1467-9892.2006.00496.x>
15. K. Fokianos, A. Rahbek, D. Tjøstheim, Poisson autoregression, *J. Amer. Stat. Assoc.*, **104** (2009), 1430–1439. <https://dx.doi.org/10.2139/ssrn.1323291>
16. K. Fokianos, Some recent progress in count time series, *Statistics*, **45** (2011), 49–58. <https://dx.doi.org/10.1080/02331888.2010.541250>
17. E. Gómez-Déniz, J. M. Sarabia, E. Calderín-Ojeda, A new discrete distribution with actuarial applications, *Insur. Math. Econ.*, **48** (2011), 406–412. <https://doi.org/10.1016/j.insmatheco.2011.01.007>
18. S. D. Guikema, J. P. Goffelt, A flexible count data regression model for risk analysis, *Risk Anal*, **28** (2008), 213–223. <https://dx.doi.org/10.1111/j.1539-6924.2008.01014.x>
19. A. Heinen, Modelling time series count data: An autoregressive conditional Poisson model, *SSRN E-journal*, 2003. <https://dx.doi.org/10.2139/ssrn.1117187>
20. N. L. Johnson, A. W. Kemp, S. Kotz, *Univariate discrete distributions*, 3 Eds., John Wiley & Sons, New Jersey, 2005.
21. R. C. Jung, M. Kukuk, R. Liesenfeld, Time series of count data: Modeling, estimation and diagnostics, *Comput. Stat. Data. Anal.*, **51** (2006), 2350–2364. <https://dx.doi.org/10.1016/j.csda.2006.08.001>
22. R. C. Jung, A. Tremayne, Useful models for time series of counts or simply wrong ones? *AStA. Adv. Stat. Anal.*, **95** (2011), 59–91. <https://dx.doi.org/10.1007/s10182-010-0139-9>
23. Y. Kang, D. Wang, K. Yang, Y. Zhang, A new thinning-based INAR (1) process for underdispersed or overdispersed counts. *J. Korean Stat. Soc.*, **49** (2020), 324–349. <https://doi.org/10.1007/s42952-019-00010-2>
24. Y. B. Koh, N. A. Bukhari, I. Mohamed, Parameter-driven state-space model for integer-valued time series with application, *J. Stat. Comput. Simul.*, **89** (2019), 1394–1409. <https://doi.org/10.1080/00949655.2019.1582653>
25. C. S. Kumar, B. U. Nair, An alternative hyper-Poisson distribution, *Statistica*, **72** (2012), 357–369. <https://dx.doi.org/10.6092/issn.1973-2201/3652>

26. D. Lord, S. D. Guikema, S. R. Geedipally, Application of the Conway-Maxwell-Poisson generalized linear model for analyzing motor vehicle crashes, *Accid. Anal. Prev.*, **40** (2008), 1123–1134. <https://dx.doi.org/10.1016/j.aap.2007.12.003>
27. L. Qian, F. Zhu, A flexible model for time series of counts with overdispersion or underdispersion, zero-inflation and heavy-tailedness, *Commun. Math. Stat.*, (2023). <https://doi.org/10.1007/s40304-022-00327-1>
28. A. J. Sáez-Castillo, A. Conde-Sánchez, A hyper-Poisson regression model for overdispersed and underdispersed count data, *Comput. Stat. Data Anal.*, **61** (2013), 148–157. <https://dx.doi.org/10.1016/j.csda.2012.12.009>
29. G. Shmueli, T. P. Minka, J. B. Kadane, S. Borle, P. Boatwright, A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution, *J. Roy. Stat. Soc. C*, **54** (2005), 127–142. <https://dx.doi.org/10.1111/j.1467-9876.2005.00474.x>
30. C. H. Weiß, Controlling correlated processes of Poisson counts, *Qual. Reliab. Eng. Int.*, **23** (2007), 741–754. <https://dx.doi.org/10.1002/qre.875>
31. C. H. Weiß, Modelling time series of counts with overdispersion, *Stat. Method. Appl.*, **18** (2009), 507–519. <https://dx.doi.org/10.1007/s10260-008-0108-6>
32. C. H. Weiß, Thinning operations for modeling time series of counts—a survey, *AStA Adv. Stat. Anal.*, **92** (2008), 319–341. <https://dx.doi.org/10.1007/s10182-008-0072-3>
33. C. H. Weiß, Integer-valued autoregressive models for counts showing underdispersion, *J. Appl. Stat.*, **40** (2013), 1931–1948. <https://doi.org/10.1080/02664763.2013.800034>
34. C. H. Weiß, *An Introduction to Discrete-valued Times Series*, 1 Ed., John Wiley & Sons Ltd, Hoboken, 2018. <https://doi.org/10.1002/9781119097013#>
35. H. White, Maximum likelihood estimation of misspecified models, *Econometrica*, **50** (1982), 1–25. <https://dx.doi.org/10.2307/1912526>
36. R. Winkelmann, *Econometric analysis of count data*, 5 Eds., Springer Science & Business Media, New York, 2008.
37. K. Yang, K. Yao, D. Wang, H. Li, Y. Diao, Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued autoregressive processes, *Metrika*, **82** (2019), 863–889. <https://doi.org/10.1007/s00184-019-00714-9>
38. S. L. Zeger, A regression model for time series of counts, *Biometrika*, **75** (1988), 621–629. <https://dx.doi.org/10.2307/2336303>
39. F. Zhu, A negative binomial integer-valued GARCH model, *J. Time Ser. Anal.*, **32** (2011), 54–67. <https://dx.doi.org/10.1111/j.1467-9892.2010.00684.x>
40. F. Zhu, Modeling time series of counts with COM-Poisson INGARCH models, *Math. Comput. Model.*, **56** (2012), 191–203. <https://dx.doi.org/10.1016/j.mcm.2011.11.069>
41. F. Zhu, Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models, *J. Math. Anal. Appl.*, **389** (2012), 58–71. <https://dx.doi.org/10.1016/j.jmaa.2011.11.042>
42. Y. Zou, S. R. Geedipally, D. Lord, Evaluating the double Poisson generalized linear model, *Accid. Anal. Prev.*, **59** (2013), 497–505. <https://dx.doi.org/10.1016/j.aap.2013.07.017>