Title	Asymptotic Properties of Estimators in Some Nonnegative Integer-Valued Time Series Models
Author(s)	曽, 小強
Citation	北海道大学. 博士(経済学) 甲第15882号
Issue Date	2024-03-25
DOI	10.14943/doctoral.k15882
Doc URL	http://hdl.handle.net/2115/92311
Type	theses (doctoral)
File Information	Zeng_Xiaoqiang.pdf

Asymptotic Properties of Estimators
in Some Nonnegative Integer-Valued Time Series Models

Xiaoqiang Zeng

Graduate School of Economics and Business, Hokkaido University

February 2024

Preface

The analysis of count time series has made rapid progress during the last few decades. There has been a huge literature about the formulations of models. Among them, we focus on nonnegative integer-valued autoregressive process of the first-order (INAR(1)) and alternative dependent counting nonnegative INAR process of the first-order (ADCINAR(1)). Both statistical models are semi-parametric in the sense that we do not impose any distributional assumption about the innovation. The primary contribution of this thesis is to give the closed-form expressions for higher autocumulant functions and propose the bias-corrections of the commonly used estimators.

The first part of this thesis is concerned with stationary INAR(1) process under a general innovation.

(i) The third, fourth, fifth, and sixth autocumulant functions of the stationary INAR(1) process are derived explicitly.

(ii) An analytical bias-correction of a class of estimators is studied.

(iii) Asymptotic theory about the Whittle likelihood estimation is presented. Also, the Wald-type test about the equidispersion is constructed, on the basis of the estimators for the innovation mean and variance.

On the other hand, the second part of this thesis is concerned with stationary ADCINAR(1) process under a general innovation.

(i) The third and fourth autocumulant functions of the stationary ADCINAR(1) process are derived explicitly, together with the structure about arbitrary higher autocumulant functions.

(ii) The two-step conditional least squares (CLS) estimator for the new parameter in the stationary ADCINAR(1) process is revisited. Also, a nonparametric (lag window-type) bias-correction and an analytical bias-correction of the Yule–Walker and CLS estimators are developed.

Contents

Chapter 1

Introduction

Analysis of count time series has been widely applied in various domains of economics, medical statistics, etc. For instance, to enhance efficiency and improve customer satisfaction, many service companies invest a lot of money and manpower for research of queuing systems. Park and Oh (1997) showed that, for certain situations, the queue length process in $M/M/\infty$ system can be represented as nonnegative integer-valued autoregressive process of the first-order (INAR(1)). Also, the survival rate of experimental animals over time is an important metric in drug development. Some examples of time series count data are the number of patients in a hospital at specific points in time, the failure count of machinery or equipment, and so on.

Over the last four decades, there has been a growing interest in modelling the INAR-types through the so-called thinning operation, unlike the traditional autoregressive model (see, e.g., Brockwell and Davis (1987)). Perhaps, the most fundamental operation is the binomial thinning due to Steutel and van Harn (1979). We refer the readers to the review paper and book by Weiß (2008,2018) for a huge literature about the formulation of models, their probabilistic aspects, and estimation or hypothesis testing.

Al-Osh and Alzaid (1987) was one of the pioneers who considered the INAR(1) process with the Poisson marginals, i.e., $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$, where $\alpha\circ$ is the binomial thinning operator with $\alpha \in [0, 1)$, and the innovation $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables according to the Poisson distribution $\mathrm{Po}((1-\alpha)\mu)$ for $\mu > 0$. They studied the Yule–Walker (YW), conditional least squares (CLS), and (conditional or full) maximum likelihood (ML) estimators for the parameter α. Park and Oh (1997) additionally established the asymptotic normality of the sample mean and sample autocovariance as well as the YW and CLS estimators. Bourguignon and Vasconcellos (2015a) considered the stationary INAR(1) process under the power series innovation, and conducted the simulations for the YW, CLS, and conditional ML estimators for the parameter α. Yang et al. (2018) applied the empirical likelihood method for the stationary mixed INAR(1) process under a general innovation.

Although the above-mentioned results are, of course, fundamental, there are several reasons that motivate us to study additional asymptotic properties of some estimators in the INAR-type processes.

First, theoretical results in the literature were often discussed for the INAR process with the Poisson marginals. It has the mathematical elegance due to the equidispersion property, i.e., the mean is equal to the variance. However, the mean and variance of other distributions may be not the same, so that, in practice, the overdispersed (or underdispersed) case is more important.

Second, Schweer and Weiß (2016) derived, for any positve integer r, the $(r+1)$th autocumulant function of the INAR(1) process with the Poisson marginals. The formulas are applicable only for the Poisson case. Considering the importance of higher cumulant functions in time series analysis, we will derive the closed-form expressions of the third, fourth, fifth, and sixth autocumulant functions for a non-Poisson case.

Third, the frequency domain analysis is a standard tool in the stationary processes (see, e.g., Brockwell and Davis (1987)). Although some simulation results about the Whittle estimation in the INAR(1) process with the Poisson marginals were given by da Silva and Oliveira (2004) (see also Zhang and Wang (2015) for random coefficient INAR(1) process), to the best of our knowledge, there is no theoretical results about the frequency domain analysis for the stationary INAR(1) process.

Fourth, we are interested in an alternative dependent counting nonnegative INAR process of the first-order (ADCINAR(1)), which was recently proposed by Nastić et al. (2017) (see also Ristić et al. (2013)), i.e., $Y_t = \alpha \diamond_{\vartheta} Y_{t-1} + \varepsilon_t$, where $\alpha \diamond_{\vartheta}$ is an alternative generalized binomial thinning operator with $0 \le \alpha \le \vartheta < 1 (\vartheta \ne 0)$, and $\{\varepsilon_t\}$ is an innovation (we do not assume its distributional form). Considering again that higher autocumulant functions are the basic in the asymptotic theory of time series, we will explicitly derive the third and fourth autocumulant functions of the stationary ADCINAR(1) process, together with the structure about arbitrary higher autocumulant functions.

Fifth, Nastić et al. (2017) considered the estimation of the new parameter ϑ in the ADCINAR(1) process under the specific innovation $\{\varepsilon_t\}$ (see Theorem 3 in Nastić et al. (2017); the marginal distribution of $\{Y_t\}$ is then geometric distribution Geo$(\mu/(1+\mu))$ for $\mu > 0$) and gave asymptotic normality when other parameter (α, μ) is unrealistically known. However, in practice, the distributional assumptions on the innovation $\{\varepsilon_t\}$ can not be specified in advance. Also, the conditional expectation is given by $E(Y_t|Y_{t-1}) = \alpha Y_{t-1} + E(\varepsilon_t)$, hence, α is an important parameter to be inferenced. Thus, we will revisit asymptotic properties of an estimator for ϑ without assuming that α is known.

Finally, it is well know that the estimators are biased in a finite-sample. Some authors constructed an analytical bias-correction in the stationary INAR(1) process (see Bourguignon and Vasconcellos (2015b) and Weiß and Schweer (2016)). One of the major results in this thesis is to derive the asymptotic expansions of the biases of the YW and CLS estimators for the parameter α in the stationary INAR(1) and ADCINAR(1) processes. Not surprisingly, the resulting analytical bias-corrected YW and CLS estimators are complicated for the ADCINAR(1) case. Therefore, the primary contribution is to develop, in a nonparametric way, lag window-type bias-corrected YW and CLS estimators, without computing the closed-form expression for the asymptotic expansions of the biases.

The thesis consists of two parts. The first part is concerned with the INAR(1) process (Chapters 2–4).

In the second part, the ADCINAR(1) process is considered (Chapters 5 and 6). The rest of this thesis is organized as follows.

Chapter 2 contains the definition of the INAR(1) process and describe its probabilistic aspects. We provide the closed-form expressions for the third, fourth, fifth, and sixth autocumulant functions of the stationary INAR(1) process under a general innovation. The asymptotic normality of sample mean and sample autocovariance are also established. Chapter 3 discusses a class of estimators for the parameter α which includes the YW, Burg, and method of moment estimators as special cases. We establish the strong consistency and asymptotic normality of such a general estimator in the stationary INAR(1) process. After deriving, explicitly, the asymptotic expansion for the bias of the general estimator, we construct an analytical bias-corrected estimator for the parameter α and study higher-order comparison among different estimators, in terms of the mean squared error. Chapter 4 is concerned with asymptotic theory about the frequency domain analysis in the stationary INAR(1) process. The strong consistency and asymptotic normality of the Whittle estimators for α, μ_ε, and σ^2_ε are established, where μ_ε and σ^2_ε are the innovation mean and variance. We propose the Wald-type tests about the equidispersion, on the basis of the estimators for μ_ε and σ^2_ε.

On the other hand, Chapter 5 is concerned with the stationary ADCINAR(1) process under a general innovation. We provide the closed-form expressions for the third and fourth autocumulant functions of the stationary ADCINAR(1) process, together with the structure about arbitrary higher autocumulant functions. Chapter 6 revisits two-step CLS (2CLS) estimator for the new parameter ϑ. Also, the lag window-type bias-correction and the analytical bias-correction of the commonly used YW and CLS estimators for the parameter α are developed. The merit of the lag window-type bias-correction is that there is no need for computing the closed-form expression of the biases and then constructing its sample analogue.

Chapter 7 contains the real data analyses using IP count data and Download count data, available in Weiß (2018). We demonstrate the usefulness of two stationary INAR(1) and ADCINAR(1) processes. Especially, we focus on (i) the equidispersion tests developed in Section 4.3, and (ii) the CLS estimation (Section 6.2) for the stationary ADCINAR(1) process without the specific distributional form of the innovation.

Finally, Chapter 8 concludes this thesis, together with some future issues.

Chapter 2

Probabilistic and statistical properties of INAR(1) process

2.1 Introduction

Our main interest is the statistical analysis of count time series. Due to the limitations of the traditional autoregressive-type models, Steutel and van Harn (1979) proposed a binomial thinning operator. Since then, there has been a huge literature about modelling time series count data, during the last four decades. Al-Osh and Alzaid (1987) mainly considered the nonnegative integer-valued autoregressive process of the first-order (INAR(1)) with the Poisson marginals, i.e., $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$, where $\alpha\circ$ is the binomial thinning operator with $\alpha \in [0, 1)$, and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (IID) random variables according to the Poisson distribution $\mathrm{Po}((1 - \alpha)\mu)$ for $\mu > 0$. It is easy to see that the autocorrelation function at lag $u(\geq 0)$ is α^u. Further, da Silva and Oliveira (2004) discussed Yule–Walker type equation of the third raw automoment (and autocumulant) function. Schweer and Weiß (2016) derived, for any positve integer r, the $(r + 1)$th autocumulant function of the INAR(1) process with the Poisson marginals. However, to the best of our knowledge, there is no paper dealing with higher autocumulant functions of the stationary INAR(1) process under a non-Poisson innovation. Considering that higher autocumulant functions are fundamental, we derive them explicity.

The rest of this chapter is organized as follows. Section 2.2 gives the definition and some useful properties of the binomial thinning operator. After the introduction of the INAR(1) process and its basic formulas of the moment, variance, and autocorrelation function at lag $u(\geq 0)$, Section 2.3 derives, explicitly, the third, fourth, fifth, and sixth autocumulant functions of the stationary INAR(1) process under a general innovation. Section 2.4 establishes asymptotic normality of the sample mean and sample autocovariance. The technical proof of Proposition 2.1 is postponed to Section 2.5.

2.2 Binomial thinning operator

Before describing the INAR(1) process, we introduce the binomial thinning operator due to Steutel and van Harn (1979). For any $\alpha \in [0, 1]$ and nonnegative integer-valued random variable Y, we define the binomial thinning operation by

$$\alpha \circ Y = \begin{cases} 0, & Y = 0, \\ \sum_{j=1}^{Y} B_j(\alpha), & Y = 1, 2, \ldots, \end{cases}$$

where $\{B_j(\alpha)\}$, referred to as counting series, is a sequence of IID random variables, independent of Y, such that

$$P[B_j(\alpha) = 1] = 1 - P[B_j(\alpha) = 0] = \alpha.$$

Given $Y = 1, 2, \ldots$, $\alpha \circ Y$ has a binomial distribution $\mathrm{Bin}(Y, \alpha)$, by definition. Thus, using

$$E(u^{\alpha \circ Y}|Y) = \begin{cases} 1, & Y = 0, \\ (1 - \alpha + \alpha u)^Y, & Y = 1, 2, \ldots, \end{cases}$$

the probability generating function (pgf) of $\alpha \circ Y$ is given by

$$E(u^{\alpha \circ Y}) = E[(1 - \alpha + \alpha u)^Y].$$

We list other properties of the thinning operation used repeatedly in this thesis (some of them are found in da Silva and Oliveira (2004), with slight corrections or extensions):

Lemma 2.1. *(i)* $0 \circ Y = 0$ *and* $1 \circ Y = Y$.

(ii) For $\beta, \gamma \in [0, 1]$, $\beta \circ (\gamma \circ Y) \overset{d}{=} (\beta\gamma) \circ Y \overset{d}{=} \gamma \circ (\beta \circ Y)$, *where* $\overset{d}{=}$ *stands for equal in distribution.*

Proof By definition, (i) is trivial, and (ii) is shown by means of the pgf

$$E[u^{\beta \circ (\gamma \circ Y)}] = E[(1 - \beta + \beta u)^{\gamma \circ Y}] = E[(1 - \gamma + \gamma(1 - \beta + \beta u))^Y] = E[(1 - \beta\gamma + \beta\gamma u)^Y] = E[u^{(\beta\gamma) \circ Y}].$$

□

Lemma 2.2. *(i) For a function* G *(e.g., we set* $G(Y) \equiv 1$ *or* $G(Y) = Y - E(Y)$*),*

$$\begin{aligned} E[G(Y)(\alpha \circ Y)] &= \alpha E[G(Y)Y], \\ E[G(Y)(\alpha \circ Y)^2] &= \alpha^2 E[G(Y)Y^2] + \alpha(1 - \alpha)E[G(Y)Y], \\ E[G(Y)(\alpha \circ Y)^3] &= \alpha^3 E[G(Y)Y^3] + 3\alpha^2(1 - \alpha)E[G(Y)Y^2] + \alpha(1 - \alpha)(1 - 2\alpha)E[G(Y)Y], \end{aligned}$$

$$
\begin{aligned}
E[G(Y)(\alpha \circ Y)^4] &= \alpha^4 E[G(Y)Y^4] + 6\alpha^3(1-\alpha)E[G(Y)Y^3] + \alpha^2(1-\alpha)(7-11\alpha)E[G(Y)Y^2] \\
&\quad + \alpha(1-\alpha)(1-6\alpha+6\alpha^2)E[G(Y)Y], \\
E[G(Y)(\alpha \circ Y)^5] &= \alpha^5 E[G(Y)Y^5] + 10\alpha^4(1-\alpha)E[G(Y)Y^4] + 5\alpha^3(1-\alpha)(5-7\alpha)E[G(Y)Y^3] \\
&\quad + 5\alpha^2(1-\alpha)(3-12\alpha+10\alpha^2)E[G(Y)Y^2] \\
&\quad + \alpha(1-\alpha)(1-14\alpha+36\alpha^2-24\alpha^3)E[G(Y)Y], \\
E[G(Y)(\alpha \circ Y)^6] &= \alpha^6 E[G(Y)Y^6] + 15\alpha^5(1-\alpha)E[G(Y)Y^5] + 5\alpha^4(1-\alpha)(13-17\alpha)E[G(Y)Y^4] \\
&\quad + 15\alpha^3(1-\alpha)(6-20\alpha+15\alpha^2)E[G(Y)Y^3] \\
&\quad + \alpha^2(1-\alpha)(31-239\alpha+476\alpha^2-274\alpha^3)E[G(Y)Y^2] \\
&\quad + \alpha(1-\alpha)(1-30\alpha+150\alpha^2-240\alpha^3+120\alpha^4)E[G(Y)Y]
\end{aligned}
$$

(it is implicitly assumed that the expectations in the right-hand side exist).

(ii) Also,

$$E[G(Y)(\alpha \circ Y - \alpha E(Y))] = \alpha E[G(Y)(Y - E(Y))],$$

hence,

$$Cov[G(Y), \alpha \circ Y] = \alpha E[G(Y)(Y - E(Y))]$$

(it is implicitly assumed that $E[G(Y)(Y - E(Y))]$ *exists).*

(iii) Furthermore,

$$
\begin{aligned}
E[G(Y)(\alpha \circ Y - \alpha E(Y))^2] &= \alpha^2 E[G(Y)(Y-E(Y))^2] + \alpha(1-\alpha)\Big\{E[G(Y)(Y-E(Y))] + E[G(Y)]E(Y)\Big\}, \\
E[G(Y)(\alpha \circ Y - \alpha E(Y))^3] &= \alpha^3 E[G(Y)(Y-E(Y))^3] \\
&\quad + 3\alpha^2(1-\alpha)\Big\{E[G(Y)(Y-E(Y))^2] + E[G(Y)(Y-E(Y))]E(Y)\Big\} \\
&\quad + \alpha(1-\alpha)(1-2\alpha)\Big\{E[G(Y)(Y-E(Y))] + E[G(Y)]E(Y)\Big\}, \\
E[G(Y)(\alpha \circ Y - \alpha E(Y))^4] &= \alpha^4 E[G(Y)(Y-E(Y))^4] \\
&\quad + 6\alpha^3(1-\alpha)\Big\{E[G(Y)(Y-E(Y))^3] + E[G(Y)(Y-E(Y))^2]E(Y)\Big\} \\
&\quad + \alpha^2(1-\alpha)(7-11\alpha)E[G(Y)(Y-E(Y))^2] \\
&\quad + 2\alpha^2(1-\alpha)(5-7\alpha)E[G(Y)(Y-E(Y))]E(Y) \\
&\quad + 3\alpha^2(1-\alpha)^2E[G(Y)][E(Y)]^2 \\
&\quad + \alpha(1-\alpha)(1-6\alpha+6\alpha^2)\Big\{E[G(Y)(Y-E(Y))] + E[G(Y)]E(Y)\Big\}, \\
E[G(Y)(\alpha \circ Y - \alpha E(Y))^5] &= \alpha^5 E[G(Y)(Y-E(Y))^5]
\end{aligned}
$$

$$
\begin{aligned}
&+ 10\alpha^4(1-\alpha)\Big\{E[G(Y)(Y-E(Y))^4] + E[G(Y)(Y-E(Y))^3]E(Y)\Big\}\\
&+ 5\alpha^3(1-\alpha)(5-7\alpha)E[G(Y)(Y-E(Y))^3]\\
&+ 10\alpha^3(1-\alpha)(4-5\alpha)E[G(Y)(Y-E(Y))^2]E(Y)\\
&+ 15\alpha^3(1-\alpha)^2E[G(Y)(Y-E(Y))][E(Y)]^2\\
&+ 5\alpha^2(1-\alpha)(3-12\alpha+10\alpha^2)E[G(Y)(Y-E(Y))^2]\\
&+ 5\alpha^2(1-\alpha)(5-18\alpha+14\alpha^2)E[G(Y)(Y-E(Y))]E(Y)\\
&+ 10\alpha^2(1-\alpha)^2(1-2\alpha)E[G(Y)][E(Y)]^2\\
&+ \alpha(1-\alpha)(1-14\alpha+36\alpha^2-24\alpha^3)\Big\{E[G(Y)(Y-E(Y))] + E[G(Y)]E(Y)\Big\}
\end{aligned}
$$

(it is implicitly assumed that the expectations in the right-hand side exist).

Proof (i) Let

$$M_\ell(x,p) = \sum_{i=1}^{\ell} p^{i-1} S\binom{\ell}{i} \sum_{j=1}^{i} s\binom{i}{j} x^j = \sum_{j=1}^{\ell} x^j \sum_{i=j}^{\ell} p^{i-1} S\binom{\ell}{i} s\binom{i}{j}$$

be a polynomial of degree ℓ in x (without constant term) for any $p \in [0,1]$ and positive integer ℓ, where $s\binom{i}{j}$ and $S\binom{\ell}{i}$ are the Stirling numbers of the first and second kind, respectively; for these definitions, see Olver et al. (2010; Section 26.8). Given Y, $\alpha \circ Y$ is distributed as the binomial distribution $\mathrm{Bin}(Y,\alpha)$, so that, for any positive integer ℓ,

$$E[(\alpha \circ Y)^\ell | Y] = \alpha M_\ell(Y,\alpha) \quad \text{(e.g., Johnson et al. (2005)).}$$

The result follows from $E[G(Y)(\alpha \circ Y)^\ell] = E[G(Y)E[(\alpha \circ Y)^\ell | Y]]$.

(ii)&(iii) Use (i) and the binomial theorem; $(A+B)^m = \sum_{i=0}^{m} {}_mC_i A^{m-i}B^i$, $m = 1,2,\ldots$. □

In addition to the moment of $\alpha \circ Y$, the second, third, fourth, fifth, and sixth central moments (or cumulants) of $\alpha \circ Y$ are derived as direct consequences of Lemma 2.2(i,iii) with $G(Y) \equiv 1$.

Corollary 2.1. *(i) When $E(Y)$ exists, then,*

$$E(\alpha \circ Y) = \alpha E(Y).$$

(ii) The jth central moments of $\alpha \circ Y$, $j = 2,3,4,5,6$ (when $E(Y^j)$ exists), are given by

$$
\begin{aligned}
E[(\alpha \circ Y - \alpha E(Y))^2] &= \alpha^2 E[(Y-E(Y))^2] + \alpha(1-\alpha)E(Y),\\
E[(\alpha \circ Y - \alpha E(Y))^3] &= \alpha^3 E[(Y-E(Y))^3] + 3\alpha^2(1-\alpha)E[(Y-E(Y))^2] + \alpha(1-\alpha)(1-2\alpha)E(Y),\\
E[(\alpha \circ Y - \alpha E(Y))^4] &= \alpha^4 E[(Y-E(Y))^4] + 6\alpha^3(1-\alpha)\{E[(Y-E(Y))^3] + E[(Y-E(Y))^2]E(Y)\}\\
&\quad + \alpha^2(1-\alpha)(7-11\alpha)E[(Y-E(Y))^2] + 3\alpha^2(1-\alpha)^2[E(Y)]^2
\end{aligned}
$$

$$
\begin{aligned}
&+ \alpha(1-\alpha)(1-6\alpha+6\alpha^2)E(Y),\\
E[(\alpha\circ Y-\alpha E(Y))^5] &= \alpha^5 E[(Y-E(Y))^5] + 10\alpha^4(1-\alpha)\Big\{E[(Y-E(Y))^4] + E[(Y-E(Y))^3]E(Y)\Big\}\\
&+ 5\alpha^3(1-\alpha)(5-7\alpha)E[(Y-E(Y))^3]\\
&+ 10\alpha^3(1-\alpha)(4-5\alpha)E[(Y-E(Y))^2]E(Y)\\
&+ 5\alpha^2(1-\alpha)(3-12\alpha+10\alpha^2)E[(Y-E(Y))^2] + 10\alpha^2(1-\alpha)^2(1-2\alpha)[E(Y)]^2\\
&+ \alpha(1-\alpha)(1-14\alpha+36\alpha^2-24\alpha^3)E(Y),\\
E[(\alpha\circ Y-\alpha E(Y))^6] &= \alpha^6 E[(Y-E(Y))^6] + 15\alpha^5(1-\alpha)\Big\{E[(Y-E(Y))^5] + E[(Y-E(Y))^4]E(Y)\Big\}\\
&+ 5\alpha^4(1-\alpha)(13-17\alpha)E[(Y-E(Y))^4]\\
&+ 10\alpha^4(1-\alpha)(11-13\alpha)E[(Y-E(Y))^3]E(Y)\\
&+ 15\alpha^3(1-\alpha)(6-20\alpha+15\alpha^2)E[(Y-E(Y))^3]\\
&+ 15\alpha^3(1-\alpha)(12-36\alpha+25\alpha^2)E[(Y-E(Y))^2]E(Y)\\
&+ \alpha^2(1-\alpha)(31-249\alpha+476\alpha^2-274\alpha^3)E[(Y-E(Y))^2]\\
&+ 5\alpha^2(1-\alpha)(5-31\alpha+52\alpha^2-26\alpha^3)[E(Y)]^2\\
&+ \alpha(1-\alpha)(1-30\alpha+150\alpha^2-240\alpha^3+120\alpha^4)E(Y)\\
&+ 45\alpha^4(1-\alpha)^2E[(Y-E(Y))^2][E(Y)]^2 + 15\alpha^3(1-\alpha)^3[E(Y)]^3.
\end{aligned}
$$

Recall that, for a random variable X with finite sixth moment,

$$
\begin{aligned}
V(X) &= E[(X-E(X))^2] \quad (\text{note } Cum_2(X) = V(X)),\\
Cum_3(X) &= E[(X-E(X))^3],\\
Cum_4(X) &= E[(X-E(X))^4] - 3[V(X)]^2,\\
Cum_5(X) &= E[(X-E(X))^5] - 10Cum_3(X)V(X),\\
Cum_6(X) &= E[(X-E(X))^6] - 15Cum_4(X)V(X) - 10[Cum_3(X)]^2 - 15[V(X)]^3.
\end{aligned}
$$

Here, instead of writing $Cum(\underbrace{X,\dots,X}_{j \text{ times}})$, we use the notation $Cum_j(X)$ for the jth cumulant, where $j = 2, 3, \dots$.

Corollary 2.2. *The jth cumulants of $\alpha\circ Y$, $j = 2,3,4,5,6$ (when $E(Y^j)$ exists), are given by*

$$
\begin{aligned}
V(\alpha\circ Y) &= \alpha^2V(Y) + \alpha(1-\alpha)E(Y),\\
Cum_3(\alpha\circ Y) &= \alpha^3Cum_3(Y) + 3\alpha^2(1-\alpha)V(Y) + \alpha(1-\alpha)(1-2\alpha)E(Y),\\
Cum_4(\alpha\circ Y) &= \alpha^4Cum_4(Y) + 6\alpha^3(1-\alpha)Cum_3(Y) + \alpha^2(1-\alpha)(7-11\alpha)V(Y)\\
&+ \alpha(1-\alpha)(1-6\alpha+6\alpha^2)E(Y),
\end{aligned}
$$

$$
\begin{aligned}
Cum_5(\alpha \circ Y) &= \alpha^5 Cum_5(Y) + 10\alpha^4(1-\alpha)Cum_4(Y) + 5\alpha^3(1-\alpha)(5-7\alpha)Cum_3(Y) \\
&\quad + 5\alpha^2(1-\alpha)(3-12\alpha+10\alpha^2)V(Y) + \alpha(1-14\alpha+36\alpha^2-24\alpha^3)E(Y), \\
Cum_6(\alpha \circ Y) &= \alpha^6 Cum_6(Y) + 15\alpha^5(1-\alpha)Cum_5(Y) + 5\alpha^4(1-\alpha)(13-17\alpha)Cum_4(Y) \\
&\quad + 15\alpha^3(1-\alpha)(6-20\alpha+15\alpha^2)Cum_3(Y) \\
&\quad + \alpha^2(1-\alpha)(31-239\alpha+476\alpha^2-274\alpha^3)V(Y) \\
&\quad + \alpha(1-\alpha)(1-30\alpha+150\alpha^2-240\alpha^3+120\alpha^4)E(Y).
\end{aligned}
$$

2.3 INAR(1) process and higher autocumulant functions

Using the binomial thinning operator $\alpha\circ$, the INAR(1) process is defined by

$$
Y_t = \alpha \circ Y_{t-1} + \varepsilon_t, \quad t = 0, \pm 1, \ldots, \tag{2.1}
$$

where $\{\varepsilon_t\}$, referred to as an innovation, is a sequence of IID nonnegative integer-valued random variables, such that ε_t and Y_{t-i} are independent for all $i \geq 1$. Al-Osh and Alzaid (1987) gave this definition, but mainly focused on the Poisson innovation. Bourguignon and Vasconcellos (2015a) considered the stationary INAR(1) process under the power series (PS) innovation. In what follows, the mean, variance, jth cumulant, and jth raw moment of ε_t are denoted by $\mu_\varepsilon = E(\varepsilon_t)$, $\sigma_\varepsilon^2 = V(\varepsilon_t)$, $\kappa_{j,\varepsilon} = Cum_j(\varepsilon_t)$, and $\mu'_{j,\varepsilon} = E(\varepsilon_t^j)$, respectively.

We emphasize that, throughout this thesis, except for some simulation experiments, we do not assume the distributional form about the innovation $\{\varepsilon_t\}$. In this sense, we treat (2.1) to be semi-parametric.

When $\alpha \in [0, 1)$, the INAR(1) process $\{Y_t\}$ is then strictly stationary and ergodic (Du and Li (1991)), whose mean μ_Y and variance σ_Y^2 are, respectively, given by

$$
\mu_Y = \frac{\mu_\varepsilon}{1-\alpha}, \quad \sigma_Y^2 = \frac{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}{1-\alpha^2},
$$

since

$$
E(Y_t) = E(\alpha \circ Y_{t-1}) + \mu_\varepsilon = \alpha E(Y_{t-1}) + \mu_\varepsilon
$$

(use Corollary 2.1(i)) and, by independence between Y_{t-1} and ε_t,

$$
V(Y_t) = V(\alpha \circ Y_{t-1} + \varepsilon_t) = V(\alpha \circ Y_{t-1}) + \sigma_\varepsilon^2 = \alpha^2 V(Y_{t-1}) + \alpha(1-\alpha)E(Y_{t-1}) + \sigma_\varepsilon^2
$$

(use Corollary 2.2). Here, when μ_ε exists, μ_Y exists; when $\mu'_{2,\varepsilon}$ exists (in this case, μ_ε and σ_ε^2 exist), σ_Y^2

exists. The dispersion index is thus given by

$$\frac{\sigma_Y^2}{\mu_Y} = \frac{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}{(1+\alpha)\mu_\varepsilon} = 1 + \frac{1}{1+\alpha}\Big(\frac{\sigma_\varepsilon^2}{\mu_\varepsilon} - 1\Big),$$

where the left-hand-side is equal to 1 when $\sigma_\varepsilon^2 = \mu_\varepsilon$, as in the Poisson case (with the equidispersion property). It follows that the stationary INAR(1) process (2.1) is overdispersed (underdispersed) when $\sigma_\varepsilon^2 > \mu_\varepsilon$ ($\sigma_\varepsilon^2 < \mu_\varepsilon$). This characterization is unique feature of the INAR(1) process, which enables us to construct a novel equidispersion test of the INAR(1) process; see Chapter 4.

Further, for any positive integer u, Y_t is independent of $\sum_{i=0}^{u-1} \alpha^i \circ \varepsilon_{t+u-i}$, so that, using Lemma 2.2(ii),

$$Cov(Y_t, Y_{t+u}) = Cov\Big(Y_t, \alpha^u \circ Y_t + \sum_{i=0}^{u-1} \alpha^i \circ \varepsilon_{t+u-i}\Big) = Cov(Y_t, \alpha^u \circ Y_t) = \alpha^u \sigma_Y^2, \qquad (2.2)$$

hence, the autocorrelation function at lag u of the stationary INAR(1) process is given by α^u, as in the usual stationary autoregressive process of the first-order (e.g., Brockwell and Davis (1987)).

Similarly, it is possible to compute the third, fourth, fifth, and sixth cumulants of Y_t (denoted by $\kappa_{j,Y}$), as follows. Noting that

$$Cum_j(Y_t) = Cum_j(\alpha \circ Y_{t-1} + \varepsilon_t) = Cum_j(\alpha \circ Y_{t-1}) + \kappa_{j,\varepsilon}$$

by independence between Y_{t-1} and ε_t, and using Corollary 2.2, we have

$$\begin{aligned}
Cum_3(Y_t) &= \alpha^3 Cum_3(Y_{t-1}) + 3\alpha^2(1-\alpha)V(Y_{t-1}) + \alpha(1-\alpha)(1-2\alpha)E(Y_{t-1}) + \kappa_{3,\varepsilon},\\
Cum_4(Y_t) &= \alpha^4 Cum_4(Y_{t-1}) + 6\alpha^3(1-\alpha)Cum_3(Y_{t-1}) + \alpha^2(1-\alpha)(7-11\alpha)V(Y_{t-1})\\
&\quad + \alpha(1-\alpha)(1-6\alpha+6\alpha^2)E(Y_{t-1}) + \kappa_{4,\varepsilon},\\
Cum_5(Y_t) &= \alpha^5 Cum_5(Y_{t-1}) + 10\alpha^4(1-\alpha)Cum_4(Y_{t-1}) + 5\alpha^3(1-\alpha)(5-7\alpha)Cum_3(Y_{t-1})\\
&\quad + 5\alpha^2(1-\alpha)(3-12\alpha+10\alpha^2)V(Y_{t-1})\\
&\quad + \alpha(1-\alpha)(1-14\alpha+36\alpha^2-24\alpha^3)E(Y_{t-1}) + \kappa_{5,\varepsilon},\\
Cum_6(Y_t) &= \alpha^6 Cum_6(Y_{t-1}) + 15\alpha^5(1-\alpha)Cum_5(Y_{t-1}) + 5\alpha^4(1-\alpha)(13-17\alpha)Cum_4(Y_{t-1})\\
&\quad + 15\alpha^3(1-\alpha)(6-20\alpha+15\alpha^2)Cum_3(Y_{t-1})\\
&\quad + \alpha^2(1-\alpha)(31-239\alpha+476\alpha^2-274\alpha^3)V(Y_{t-1})\\
&\quad + \alpha(1-\alpha)(1-30\alpha+150\alpha^2-240\alpha^3+120\alpha^4)E(Y_{t-1}) + \kappa_{6,\varepsilon},
\end{aligned}$$

hence, by strictly stationarity of $\{Y_t\}$,

$$\kappa_{3,Y} = \frac{1}{1-\alpha^3}\{3\alpha^2(1-\alpha)\sigma_Y^2 + \alpha(1-\alpha)(1-2\alpha)\mu_Y + \kappa_{3,\varepsilon}\},$$

$$\kappa_{4,Y} = \frac{1}{1-\alpha^4}\{6\alpha^3(1-\alpha)\kappa_{3,Y} + \alpha^2(1-\alpha)(7-11\alpha)\sigma_Y^2 + \alpha(1-\alpha)(1-6\alpha+6\alpha^2)\mu_Y + \kappa_{4,\varepsilon}\},$$

$$\kappa_{5,Y} = \frac{1}{1-\alpha^5}\{10\alpha^4(1-\alpha)\kappa_{4,Y} + 5\alpha^3(1-\alpha)(5-7\alpha)\kappa_{3,Y} + 5\alpha^2(1-\alpha)(3-12\alpha+10\alpha^2)\sigma_Y^2$$
$$+ \alpha(1-\alpha)(1-14\alpha+36\alpha^2-24\alpha^3)\mu_Y + \kappa_{5,\varepsilon}\},$$

$$\kappa_{6,Y} = \frac{1}{1-\alpha^6}\{15\alpha^5(1-\alpha)\kappa_{5,Y} + 5\alpha^4(1-\alpha)(13-17\alpha)\kappa_{4,Y} + 15\alpha^3(1-\alpha)(6-20\alpha+15\alpha^2)\kappa_{3,Y}$$
$$+ \alpha^2(1-\alpha)(31-239\alpha+476\alpha^2-274\alpha^3)\sigma_Y^2$$
$$+ \alpha(1-\alpha)(1-30\alpha+150\alpha^2-240\alpha^3+120\alpha^4)\mu_Y + \kappa_{6,\varepsilon}\}.$$

Remark 2.1. If $\mu'_{J,\varepsilon}$ exists for $J \geq 3$ (in this case, μ_ε, σ_ε^2, and $\kappa_{j,\varepsilon}$, $j = 3, \ldots, J$, exist), then, $\kappa_{j,Y}$ exists for $j = 3, \ldots, J$.

Before presenting the higher autocumulant functions of the stationary INAR(1) process $\{Y_t\}$, we introduce the following notations:

$$Q_{2:3,Y} = \kappa_{3,Y} - \sigma_Y^2,$$
$$Q_{2:4,Y} = \kappa_{4,Y} - 3\kappa_{3,Y} + 2\sigma_Y^2,$$
$$Q_{2:5,Y} = \kappa_{5,Y} - 6\kappa_{4,Y} + 11\kappa_{3,Y} - 6\sigma_Y^2,$$
$$Q_{2:6,Y} = \kappa_{6,Y} - 10\kappa_{5,Y} + 35\kappa_{4,Y} - 50\kappa_{3,Y} + 24\sigma_Y^2.$$

Remark 2.2. The cumulant generating function of a random variable X that is distributed according to a Poisson distribution with a parameter $\lambda \geq 0$, denoted by Po(λ), is given by $\log E[e^{tX}] = \lambda(e^t - 1)$, hence, all cumulants of $X \sim$ Po(λ) are equal to λ. Therefore, $Q_{2:3,Y} = Q_{2:4,Y} = Q_{2:5,Y} = Q_{2:6,Y} = 0$ for the INAR(1) process $\{Y_t\}$ with the Poisson marginals, i.e., these quantities show the departure from the Poisson distribution, parallel to the fact that all higher cumulants of the normal distribution vanish.

For the asymptotic theory of the stationary time series, the autocumulant functions, equivalently, the central automoment functions, are fundamental. Recall that, given a stationary process $\{X_t\}$ with mean μ_X, the $(r+1)$th central automoment function, where r is a positive integer, is defined by

$$\mu_X(u_1, \ldots, u_r) = E\Big[(X_t - \mu_X)\prod_{i=1}^{r}(X_{t+u_i} - \mu_X)\Big]$$

for nonnegative integers $u_r \geq \cdots \geq u_1 \geq 0$. The second central automoment function is nothing but the autocovariance function $\gamma_X(\cdot)$ and the third central automoment function is equal to the third autocumulant function $\gamma_X(\cdot,\cdot)$, i.e.,

$$\gamma_X(u_1) = \mu_X(u_1) \text{ and } \gamma_X(u_1, u_2) = \mu_X(u_1, u_2).$$

Further, the fourth autocumulant function $\gamma_X(\cdot,\cdot,\cdot)$ is given by

$$\gamma_X(u_1,u_2,u_3) = \mu_X(u_1,u_2,u_3) - \gamma_X(u_1)\gamma_X(u_3-u_2) - \gamma_X(u_2)\gamma_X(u_3-u_1) - \gamma_X(u_3)\gamma_X(u_2-u_1).$$

Proposition 2.1. *In addition to the autocovariance function $\gamma_Y(u) = \alpha^u\sigma_Y^2$ (see (2.2)), the third, fourth, fifth, and sixth autocumulant functions of the stationary INAR(1) process $\{Y_t\}$ are, respectively, given by*

$$\gamma_Y(u,v) = \alpha^v(\alpha^u Q_{2:3,Y} + \sigma_Y^2), \tag{2.3}$$

$$\gamma_Y(u,v,w) = \alpha^w\{\alpha^{v+u}Q_{2:4,Y} + (\alpha^u + 2\alpha^v)Q_{2:3,Y} + \sigma_Y^2\}, \tag{2.4}$$

$$\begin{aligned}\gamma_Y(u,v,w,x) = \alpha^x\{&\alpha^{w+v+u}Q_{2:5,Y} + (\alpha^{v+u} + 2\alpha^{w+u} + 3\alpha^{w+v})Q_{2:4,Y}\\ &+ (\alpha^u + 2\alpha^v + 4\alpha^w)Q_{2:3,Y} + \sigma_Y^2\},\end{aligned} \tag{2.5}$$

$$\begin{aligned}\gamma_Y(u,v,w,x,y) = \alpha^y\{&\alpha^{x+w+v+u}Q_{2:6,Y} + (\alpha^{w+v+u} + 2\alpha^{x+v+u} + 3\alpha^{x+w+u} + 4\alpha^{x+w+v})Q_{2:5,Y}\\ &+ (\alpha^{v+u} + 2\alpha^{w+u} + 3\alpha^{w+v} + 4\alpha^{x+u} + 6\alpha^{x+v} + 9\alpha^{x+w})Q_{2:4,Y}\\ &+ (\alpha^u + 2\alpha^v + 4\alpha^w + 8\alpha^x)Q_{2:3,Y} + \sigma_Y^2\}\end{aligned} \tag{2.6}$$

for $y \ge x \ge w \ge v \ge u \ge 0$ (the proof is postponed to Section 2.5).

Remark 2.3. Schweer and Weiß (2016) showed that the $(r+1)$th autocumulant function of the INAR(1) process with the Poisson marginals is given by $Cum(Y_t, Y_{t+u_1}, \dots, Y_{t+u_r}) = \alpha^{u_r}\mu_Y$ for any positive integer r and nonnegative integers $u_r \ge \cdots \ge u_1 \ge 0$.

Remark 2.4. When $\alpha \in [0,1)$, the INAR(1) process (2.1) is strictly stationary, having the representation $Y_t = \sum_{i=0}^{\infty}\alpha^i \circ \varepsilon_{t-i}$ (see Al-Osh and Alzaid (1987)). Using this, we have another derivation for the mean μ_Y and variance σ_Y^2, as well as the second, third, and fourth automoment functions, i.e., for $w \ge v \ge u \ge 0$,

$$\begin{aligned}&\mu_Y(u) = \alpha^u\sigma_Y^2, \quad \mu_Y(u,v) = \alpha^v(\alpha^u Q_{2:3,Y} + \sigma_Y^2),\\ &\mu_Y(u,v,w) = \alpha^w\{\alpha^{v+u}Q_{2:4,Y} + (\alpha^u + 2\alpha^v)Q_{2:3,Y} + \sigma_Y^2\} + (\alpha^{w-v+u} + 2\alpha^{w+v-u})\sigma_Y^4.\end{aligned}$$

2.4 Asymptotic normality of sample mean and sample autocovariance

Suppose that the observation $\{Y_1,\dots,Y_n\}$ of length n is generated by $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$, where we assume that $\alpha \in [0,1)$ (in that case, the INAR(1) process is strictly stationary and ergodic; see Du and Li (1991)) and that $E(\varepsilon_t^J)$ exists for some $J \ge 2$ (see Remark 2.1). Let $\overline{Y} = (1/n)\sum_{t=1}^n Y_t$ and $\widehat{\gamma}(h) = (1/n)\sum_{t=1}^{n-h}(Y_t - \overline{Y})(Y_{t+h} - \overline{Y})$, $h = 0,1,\dots,n-1$.

The following asymptotic normality of the sample mean and sample autocovariance is a non-Poisson extension of Park and Oh (1997). Note that some formulas in their lemmas for the INAR(1) process with the Poisson marginals were corrected here (the result below is, however, not intended just to make the correction).

Proposition 2.2. *Suppose that $\{Y_t\}$ is INAR(1) process with $\alpha \in [0,1)$ and that $E(\varepsilon_t^J)$ exists for some integer $J \geq 4$. The following hold for a fixed positive integer m.*

(i) Let $\widetilde{\gamma}(h) = (1/n)\sum_{t=1}^{n}(Y_t - \mu_Y)(Y_{t+h} - \mu_Y)$, $h = 0, 1, \ldots, m$. We have

$$\sqrt{n}[\overline{Y} - \mu_Y, \widetilde{\gamma}(0) - \gamma_Y(0), \ldots, \widetilde{\gamma}(m) - \gamma_Y(m)]^T \xrightarrow{d} \mathrm{N}(0, \Xi),$$

where

$$\Xi = \begin{pmatrix} \Xi_{(\mu\mu)} & \Xi_{(\gamma\mu)}^T \\ \Xi_{(\gamma\mu)} & \Xi_{(\gamma\gamma)} \end{pmatrix}$$

is the $(m+2) \times (m+2)$ matrix such that $\Xi_{(\mu\mu)} = (1+\alpha)\sigma_Y^2/(1-\alpha)$, $\Xi_{(\gamma\mu)} = (\Xi_0, \Xi_1, ..., \Xi_m)^T$ is $(m+1) \times 1$ vector, whose Ith element is given by

$$\Xi_I = \frac{1+\alpha+\alpha^2}{1-\alpha^2}\alpha^I Q_{2:3,Y} + \Big(\frac{1+\alpha}{1-\alpha} + I\Big)\alpha^I \sigma_Y^2,$$

and $\Xi_{(\gamma\gamma)}$ is the $(m+1) \times (m+1)$ symmetric matrix $(\Xi_{IJ}; I, J = 0, 1, \ldots, m)$, whose (I,J)th element, for $I \geq J$, is given by

$$\begin{aligned}
\Xi_{IJ} = &\frac{Q_{2:4,Y}}{1-\alpha^2}(1+\alpha^2)\alpha^{I+J} \\
&+ \frac{Q_{2:3,Y}}{1-\alpha^2}\{(1+\alpha+\alpha^2)\alpha^I - \alpha(\alpha^{2I-J} + \alpha^{2I+J} + \alpha^{I+2J}) + 2(1+\alpha)^2\alpha^{I+J}\} \\
&+ \frac{\sigma_Y^2}{1-\alpha}\{1 + I - J + \alpha(1 - I + J)\}\alpha^I \\
&+ \frac{\sigma_Y^4}{1-\alpha^2}[\{1 + I - J + \alpha^2(1 - I + J)\}\alpha^{I-J} + \{1 + I + J + \alpha^2(1 - I - J)\}\alpha^{I+J}].
\end{aligned}$$

(ii) $\sqrt{n}[\overline{Y} - \mu_Y, \widehat{\gamma}(0) - \gamma_Y(0), \ldots, \widehat{\gamma}(m) - \gamma_Y(m)]^T \xrightarrow{d} \mathrm{N}(0, \Xi)$.

Proof (i) For each integer $M(> 2m)$, let

$$\widetilde{\gamma}^{(M)}(h) = \frac{1}{n}\sum_{t=1}^{n}(Y_t^{(M)} - \mu_{Y^{(M)}})(Y_{t+h}^{(M)} - \mu_{Y^{(M)}}), \quad h = 0, 1, \ldots, m, \quad \text{and } \overline{Y}^{(M)} = \frac{1}{n}\sum_{t=1}^{n} Y_t^{(M)},$$

where $Y_t^{(M)} = \sum_{i=0}^{M}\alpha^i \circ \varepsilon_{t-i}$ is a truncated version of $Y_t = \sum_{i=0}^{\infty}\alpha^i \circ \varepsilon_{t-i}$. Here, $\{Y_t^{(M)}\}$ is strictly stationary, whose mean and autocovariance function at lag $u(\geq 0)$ are, respectively, given by

$$\begin{aligned}
\mu_{Y^{(M)}} &= \frac{1-\alpha^{M+1}}{1-\alpha}\mu_\varepsilon, \\
\gamma_{Y^{(M)}}(u) &= \begin{cases} 0, & u > M, \\ \alpha^u\Big[\frac{1-\alpha^{2(M+1-u)}}{1-\alpha^2}\sigma_\varepsilon^2 + \Big\{\frac{1-\alpha^{M+1-u}}{1-\alpha} - \frac{1-\alpha^{2(M+1-u)}}{1-\alpha^2}\Big\}\mu_\varepsilon\Big], & 0 \leq u \leq M. \end{cases}
\end{aligned}$$

Since $\{Y_t^{(M)} - \mu_{Y^{(M)}}, (Y_t^{(M)} - \mu_{Y^{(M)}})^2 - \gamma_{Y^{(M)}}(0), ..., (Y_t^{(M)} - \mu_{Y^{(M)}})(Y_{t+m}^{(M)} - \mu_{Y^{(M)}}) - \gamma_{Y^{(M)}}(m)\}$ is strictly stationary and $(M+m)$-dependent, we can use a finite-dependent central limit theorem and the Cramér–Wold device to show that

$$\sqrt{n}[\overline{Y}^{(M)} - \mu_{Y^{(M)}}, \widetilde{\gamma}^{(M)}(0) - \gamma_{Y^{(M)}}(0), \ldots, \widetilde{\gamma}^{(M)}(m) - \gamma_{Y^{(M)}}(m)]^T \xrightarrow{d} \mathrm{N}(0, \Xi^{(M)}),$$

where

$$\Xi^{(M)} = \begin{pmatrix} \Xi^{(M)}_{(\mu\mu)} & \left(\Xi^{(M)}_{(\gamma\mu)}\right)^T \\ \Xi^{(M)}_{(\gamma\mu)} & \Xi^{(M)}_{(\gamma\gamma)} \end{pmatrix}$$

is the $(m+1) \times (m+1)$ matrix such that

$$\begin{aligned} \Xi^{(M)}_{(\mu\mu)} &= \lim_{n\to\infty} nV(\overline{Y}^{(M)}), \\ \Xi^{(M)}_{(\gamma\mu)} &= [\lim_{n\to\infty} nCov(\overline{Y}^{(M)}, \widetilde{\gamma}^{(M)}(0)), \ldots, \lim_{n\to\infty} nCov(\overline{Y}^{(M)}, \widetilde{\gamma}^{(M)}(m))]^T, \end{aligned}$$

and $\Xi^{(M)}_{(\gamma\gamma)}$ is the $(m+1) \times (m+1)$ matrix $(\Xi^{(M)}_{IJ}; I, J = 0, 1, ..., m)$, whose (I, J)th element is given by

$$\Xi^{(M)}_{IJ} = \lim_{n\to\infty} nCov(\widetilde{\gamma}^{(M)}(I), \widetilde{\gamma}^{(M)}(J)).$$

On the other hand, we introduce $\Delta_t^{(M)} = (Y_t - \mu_Y) - (Y_t^{(M)} - \mu_{Y^{(M)}})$, $t = 1, \ldots, n$, and then rewrite

$$(\overline{Y} - \mu_Y) - (\overline{Y}^{(M)} - \mu_{Y^{(M)}}) = \frac{1}{n}\sum_{t=1}^{n} \Delta_t^{(M)}$$

and, for $h = 0, 1, \ldots, m$,

$$\begin{aligned} &\widetilde{\gamma}(h) - \gamma_Y(h) - \{\widetilde{\gamma}^{(M)}(h) - \gamma_{Y^{(M)}}(h)\} \\ &= \frac{1}{n}\sum_{t=1}^{n}[(Y_t - \mu_Y)(Y_{t+h} - \mu_Y) - (Y_t^{(M)} - \mu_{Y^{(M)}})(Y_{t+h}^{(M)} - \mu_{Y^{(M)}})] - \{\gamma_Y(h) - \gamma_{Y^{(M)}}(h)\} \\ &= X^{(M)}_{n,h} + Y^{(M)}_{n,h} + Z^{(M)}_{n,h} - \{\gamma_Y(h) - \gamma_{Y^{(M)}}(h)\}, \qquad (2.7) \end{aligned}$$

where

$$X^{(M)}_{n,h} = \frac{1}{n}\sum_{t=1}^{n}(Y_t^{(M)} - \mu_{Y^{(M)}})\Delta_{t+h}^{(M)}, \;\; Y^{(M)}_{n,h} = \frac{1}{n}\sum_{t=1}^{n}\Delta_t^{(M)}(Y_{t+h}^{(M)} - \mu_{Y^{(M)}}), \;\; \text{and} \;\; Z^{(M)}_{n,h} = \frac{1}{n}\sum_{t=1}^{n}\Delta_t^{(M)}\Delta_{t+h}^{(M)}.$$

It is easy to see that

$$V[\sqrt{n}(\overline{Y} - \overline{Y}^{(M)})] = \frac{1}{n}V\Big(\sum_{t=1}^{n}\Delta_t^{(M)}\Big) \to 0 \quad \text{(uniformly in } n\text{) as } M \to \infty.$$

To verify the uniform asymptotical negligibility of (2.7) after multiplication by $\sqrt{n}$, we have only to show that, for $h = 0, 1, \ldots, m$,

$$
\begin{aligned}
&V[\sqrt{n}\{\widetilde{\gamma}(h) - \widetilde{\gamma}^{(M)}(h)\}] \\
&= nV[X_{n,h}^{(M)} + Y_{n,h}^{(M)} + Z_{n,h}^{(M)}] \\
&= n|V[X_{n,h}^{(M)}] + V[Y_{n,h}^{(M)}] + V[Z_{n,h}^{(M)}] + 2Cov[X_{n,h}^{(M)}, Y_{n,h}^{(M)}] + 2Cov[Y_{n,h}^{(M)}, Z_{n,h}^{(M)}] + 2Cov[X_{n,h}^{(M)}, Z_{n,h}^{(M)}]| \\
&\leq n\Big\{\sqrt{V[X_{n,h}^{(M)}]} + \sqrt{V[Y_{n,h}^{(M)}]} + \sqrt{V[Z_{n,h}^{(M)}]}\Big\}^2 \\
&\leq 3n\{V[X_{n,h}^{(M)}] + V[Y_{n,h}^{(M)}] + V[Z_{n,h}^{(M)}]\} \to 0 \quad \text{(uniformly in } n\text{) as } M \to \infty.
\end{aligned}
$$

Thus, we can see (e.g., Brockwell and Davis (1987)) that

$$\sqrt{n}[\overline{Y} - \mu_Y, \widetilde{\gamma}(0) - \gamma_Y(0), \ldots, \widetilde{\gamma}(m) - \gamma_Y(m)]^T \xrightarrow{d} \mathrm{N}(0, \Xi),$$

where $\Xi = \lim_{M\to\infty} \Xi^{(M)}$.

(ii) We have, for $h = 0, 1, \ldots, m$,

$$
\begin{aligned}
&\sqrt{n}[\widetilde{\gamma}(h) - \widehat{\gamma}(h)] \\
&= \frac{1}{\sqrt{n}}\Big[\sum_{t=1}^{n}(Y_t - \mu_Y)(Y_{t+h} - \mu_Y) - \sum_{t=1}^{n-h}(Y_t - \overline{Y})(Y_{t+h} - \overline{Y})\Big] \\
&= \frac{1}{\sqrt{n}}\Big[\sum_{t=1}^{n-h}(Y_t - \mu_Y)(Y_{t+h} - \mu_Y) - \sum_{t=1}^{n-h}(Y_t - \overline{Y})(Y_{t+h} - \overline{Y})\Big] + \frac{1}{\sqrt{n}}\sum_{t=n-h+1}^{n}(Y_t - \mu_Y)(Y_{t+h} - \mu_Y) \\
&= \frac{1}{\sqrt{n}}\Big[\sum_{t=1}^{n-h}\{(Y_t - \mu_Y) + (Y_{t+h} - \mu_Y) - (\overline{Y} - \mu_Y)\}\Big](\overline{Y} - \mu_Y) + \frac{1}{\sqrt{n}}\sum_{t=n-h+1}^{n}(Y_t - \mu_Y)(Y_{t+h} - \mu_Y) \\
&= O_p(1/\sqrt{n}) \\
&= o_p(1).
\end{aligned}
$$

Then, the result follows from Slutsky's theorem and (i). □

Poisson INAR(1) process

The stationary marginal distribution of $\{Y_t\}$, if it exists, is determined by the equation of the pgf;

$$E[u^{Y_t}] = E[u^{\alpha\circ Y_t}]E[u^{\varepsilon_t}].$$

(I) Suppose that the INAR(1) process (2.1), with $\alpha \in [0, 1)$, has the marginal Poisson distribution

Po(λ), where $\lambda > 0$. The pgfs of Y_t and $\alpha \circ Y_t$ are, respectively, given by

$$E[u^{Y_t}] = \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} e^{-\lambda} u^y = e^{\lambda(u-1)},$$
$$E[u^{\alpha \circ Y_t}] = E[(1-\alpha+\alpha u)^{Y_t}] = \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} e^{-\lambda}(1-\alpha+\alpha u)^y = e^{\alpha\lambda(u-1)},$$

so that

$$E[u^{\varepsilon_t}] = \frac{E[u^{Y_t}]}{E[u^{\alpha \circ Y_t}]} = e^{(1-\alpha)\lambda(u-1)},$$

i.e., ε_t must be distributed according to the Poisson distribution Po($(1-\alpha)\lambda$).

(II) Conversely, if ε_t is distributed according to the Poisson distribution Po($(1-\alpha)\lambda$), where $\alpha \in [0, 1)$ and $\lambda > 0$, the pgf of $Y_t = \sum_{j=0}^{\infty} \alpha^j \circ \varepsilon_{t-j}$, Y_t is then given by

$$E[u^{Y_t}] = \prod_{j=0}^{\infty} E[u^{\alpha^j \circ \varepsilon_{t-j}}] = \prod_{j=0}^{\infty} e^{\alpha^j (1-\alpha)\lambda(u-1)} = e^{\lambda(u-1)},$$

i.e., the marginal distribution of the stationary INAR(1) process is the Poisson distribution Po(λ).

PS distribution

Let $S \subset \{0, 1, \ldots\}$. We say (Noack (1950)) that a random variable X is distributed according to the PS distribution with one parameter $\theta(> 0)$, if its probability mass function is given by

$$P(X = x) = \frac{a(x)\theta^x}{C(\theta)}, \quad x \in S,$$

where $a(x) \geq 0$ depends only on x, and $C(\theta) = \sum_{x \in S} a(x)\theta^x$. Then, we have

$$E[X] = \frac{\theta}{C(\theta)} \frac{dC(\theta)}{d\theta} = \theta \frac{d \log C(\theta)}{d\theta},$$
$$V[X] = \frac{\theta^2}{C(\theta)} \frac{d^2 C(\theta)}{d\theta^2} + \frac{\theta}{C(\theta)} \frac{dC(\theta)}{d\theta} - \left[\frac{\theta}{C(\theta)} \frac{dC(\theta)}{d\theta}\right]^2 = \theta^2 \frac{d^2 \log C(\theta)}{d\theta^2} + \theta \frac{d \log C(\theta)}{d\theta}.$$

Also, its pgf is obviously given by

$$E[u^X] = \frac{C(\theta u)}{C(\theta)}.$$

The PS distribution, with $S = \{0, 1, \ldots\}$, contains the Poisson or negative binomial (NB) distributon, as follows.

- Let $\theta > 0$. The Poisson distribution, denoted by Po(θ), corresponds to the case $a(x) = (x!)^{-1}$, where $C(\theta) = e^{\theta}$. For $X \sim \text{Po}(\theta)$, it is well known that $E[X] = V[X] = \theta$ (equidispersion); besides, all cumulants of X are equal to θ, since the cumulant generating function is given by $\log E[e^{tX}] = \theta(e^t - 1)$.

- Let $0 < \theta < 1$. Given a positive integer r, the NB distribution, denoted by $\text{NB}(r, \theta)$, corresponds to the case $a(x) = \Gamma(r + x)/[x!\Gamma(r)]$, where $C(\theta) = (1 - \theta)^{-r}$. It is not difficult to see that, for $X \sim \text{NB}(r, \theta)$, $E[X] = \theta r(1 - \theta)^{-1}$ and $V[X] = \theta r(1 - \theta)^{-2}$ (overdispersion), since

$$\frac{\theta r(1 - \theta)^{-2}}{\theta r(1 - \theta)^{-1}} = \frac{1}{1 - \theta} = 1 + \frac{\theta}{1 - \theta} > 1.$$

Moreover,

$$Cum_3(X) = \frac{r\theta(1 + \theta)}{(1 - \theta)^3}, \;\; Cum_4(X) = \frac{r\theta(1 + 4\theta + \theta^2)}{(1 - \theta)^4}.$$

2.5 Proof of Proposition 2.1

Note that, for any positive integer $h > 0$, we have $Y_{t+h} = \alpha^h \circ Y_t + \sum_{j=0}^{h-1} \alpha^j \circ \varepsilon_{t+h-j}$, which implies that, for $u > 0$ and $v \geq 0$, $\alpha^v \circ Y_t$ is independent of $\sum_{j=0}^{u-1} \alpha^j \circ \varepsilon_{t+u-j}$. For the derivation of the $(r + 1)$th autocumulant function of the stationary INAR(1) process $\{Y_t\}$, where $r = 2, 3, 4, 5$, it suffices to compute

$$Cum(Y_t, Y_{t+u_1}, \ldots, Y_{t+u_r}) = Cum(Y_t, \alpha^{u_1} \circ Y_t, \ldots, \alpha^{u_r} \circ Y_t), \quad 0 \leq u_1 \leq \cdots \leq u_r,$$

by utilizing the formulas of the cumulants in terms of the central moments (e.g., McCullagh (2018)). Namely, this step is done through the computation of the central moment

$$E\Big[(Y_t - \mu_Y) \prod_{i=1}^{r} (\alpha^{u_i} \circ Y_t - \alpha^{u_i} \mu_Y)\Big].$$

The following four lemmas (it is implicitly assumed that the expectations in the right-hand side exist), which are extensions of Lemma 2.2(ii,iii), are prepared (the proof will be given at the end of this section).

Lemma 2.3. *For $\beta, \gamma \in [0, 1]$,*

$$\begin{aligned} &E[G(Y)(\beta \circ Y - \beta E(Y))(\gamma \circ (\beta \circ Y) - \gamma\beta E(Y))] \\ &= \gamma\beta^2 E[G(Y)(Y - E(Y))^2] + \gamma\beta(1 - \beta)\Big\{E[G(Y)(Y - E(Y))] + E[G(Y)]E(Y)\Big\}. \end{aligned}$$

Lemma 2.4. *For $\beta, \gamma, \delta \in [0, 1]$,*

$$\begin{aligned} &E[G(Y)(\beta \circ Y - \beta E(Y))(\gamma \circ (\beta \circ Y) - \gamma\beta E(Y))(\delta \circ (\gamma \circ (\beta \circ Y)) - \delta\gamma\beta E(Y))] \\ &= \delta\gamma^2\beta^3 E[G(Y)(Y - E(Y))^3] \end{aligned}$$

$$+\delta\gamma\beta^2(1+2\gamma-3\gamma\beta)\Big\{E[G(Y)(Y-E(Y))^2]+E[G(Y)(Y-E(Y))]E(Y)\Big\}$$
$$+\delta\gamma\beta(1-\beta)(1-2\gamma\beta)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}.$$

Lemma 2.5. *For* $\beta,\gamma,\delta,\eta\in[0,1]$,

$$\begin{aligned}
&E[G(Y)(\beta\circ Y-\beta E(Y))(\gamma\circ(\beta\circ Y)-\gamma\beta E(Y))(\delta\circ(\gamma\circ(\beta\circ Y))-\delta\gamma\beta E(Y))\\
&\quad\times(\eta\circ(\delta\circ(\gamma\circ(\beta\circ Y)))-\eta\delta\gamma\beta E(Y))]\\
&=\eta\delta^2\gamma^3\beta^4E[G(Y)(Y-E(Y))^4]\\
&\quad+\eta\delta\gamma^2\beta^3(1+2\delta+3\delta\gamma-6\delta\gamma\beta)\Big\{E[G(Y)(Y-E(Y))^3]+E[G(Y)(Y-E(Y))^2]E(Y)\Big\}\\
&\quad+\eta\delta\gamma\beta(\beta+2\gamma\beta+4\delta\gamma\beta-3\gamma\beta^2-6\delta\gamma\beta^2-9\delta\gamma^2\beta^2+11\delta\gamma^2\beta^3)E[G(Y)(Y-E(Y))^2]\\
&\quad+\eta\delta\gamma\beta(\beta+3\gamma\beta+6\delta\gamma\beta-4\gamma\beta^2-8\delta\gamma\beta^2-12\delta\gamma^2\beta^2+14\delta\gamma^2\beta^3)E[G(Y)(Y-E(Y))]E(Y)\\
&\quad+\eta\delta\gamma\beta(\gamma\beta+2\delta\gamma\beta-\gamma\beta^2-2\delta\gamma\beta^2-3\delta\gamma^2\beta^2+3\delta\gamma^2\beta^3)E[G(Y)][E(Y)]^2\\
&\quad+\eta\delta\gamma\beta(1-\beta-2\gamma\beta-4\delta\gamma\beta+2\gamma\beta^2+4\delta\gamma\beta^2+6\delta\gamma^2\beta^2-6\delta\gamma^2\beta^3)\\
&\quad\quad\times\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}.
\end{aligned}$$

Lemma 2.6. *For* $\beta,\gamma,\delta,\eta,\iota\in[0,1]$,

$$\begin{aligned}
&E[(Y-E(Y))(\beta\circ Y-\beta E(Y))(\gamma\circ(\beta\circ Y)-\gamma\beta E(Y))(\delta\circ(\gamma\circ(\beta\circ Y))-\delta\gamma\beta E(Y))\\
&\quad\times(\eta\circ(\delta\circ(\gamma\circ(\beta\circ Y)))-\eta\delta\gamma\beta E(Y))(\iota\circ(\eta\circ(\delta\circ(\gamma\circ(\beta\circ Y))))-\iota\eta\delta\gamma\beta E(Y))]\\
&=\iota\eta^2\delta^3\gamma^4\beta^5E[(Y-E(Y))^6]\\
&\quad+\iota\eta\delta^2\gamma^3\beta^4(1+2\eta+3\eta\delta+4\eta\delta\gamma-10\eta\delta\gamma\beta)\Big\{E[(Y-E(Y))^5]+E[(Y-E(Y))^4]E(Y)\Big\}\\
&\quad+\iota\eta\delta\gamma\beta(\gamma\beta^2+2\delta\gamma\beta^2+3\delta\gamma^2\beta^2+4\eta\delta\gamma\beta^2+6\eta\delta\gamma^2\beta^2+9\eta\delta^2\gamma^2\beta^2-6\delta\gamma^2\beta^3-12\eta\delta\gamma^2\beta^3\\
&\quad\quad-18\eta\delta^2\gamma^2\beta^3-24\eta\delta^2\gamma^3\beta^3+35\eta\delta^2\gamma^3\beta^4)E[(Y-E(Y))^4]\\
&\quad+\iota\eta\delta\gamma\beta(\gamma\beta^2+3\delta\gamma\beta^2+5\delta\gamma^2\beta^2+6\eta\delta\gamma\beta^2+10\eta\delta\gamma^2\beta^2+15\eta\delta^2\gamma^2\beta^2-9\delta\gamma^2\beta^3-18\eta\delta\gamma^2\beta^3\\
&\quad\quad-27\eta\delta^2\gamma^2\beta^3-36\eta\delta^2\gamma^3\beta^3+50\eta\delta^2\gamma^3\beta^4)E[(Y-E(Y))^3]E(Y)\\
&\quad+\iota\eta\delta\gamma\beta(\beta+2\gamma\beta+4\delta\gamma\beta+8\eta\delta\gamma\beta-3\gamma\beta^2-6\delta\gamma\beta^2-9\delta\gamma^2\beta^2-12\eta\delta\gamma\beta^2-18\eta\delta\gamma^2\beta^2-27\eta\delta^2\gamma^2\beta^2\\
&\quad\quad+11\delta\gamma^2\beta^3+22\eta\delta\gamma^2\beta^3+33\eta\delta^2\gamma^2\beta^3+44\eta\delta^2\gamma^3\beta^3-50\eta\delta^2\gamma^3\beta^4)E[(Y-E(Y))^3]\\
&\quad+\iota\eta\delta\gamma\beta(\beta+3\gamma\beta+7\delta\gamma\beta+14\eta\delta\gamma\beta-4\gamma\beta^2-9\delta\gamma\beta^2-14\delta\gamma^2\beta^2-18\eta\delta\gamma\beta^2-28\eta\delta\gamma^2\beta^2-42\eta\delta^2\gamma^2\beta^2\\
&\quad\quad+16\delta\gamma^2\beta^3+32\eta\delta\gamma^2\beta^3+48\eta\delta^2\gamma^2\beta^3+64\eta\delta^2\gamma^3\beta^3-70\eta\delta^2\gamma^3\beta^4)V(Y)E(Y)\\
&\quad+\iota\eta\delta\gamma\beta(\delta\gamma\beta^2+2\delta\gamma^2\beta^2+2\eta\delta\gamma\beta^2+4\eta\delta\gamma^2\beta^2+6\eta\delta^2\gamma^2\beta^2-3\delta\gamma^2\beta^3-6\eta\delta\gamma^2\beta^3-9\eta\delta^2\gamma^2\beta^3\\
&\quad\quad-12\eta\delta^2\gamma^3\beta^3+15\eta\delta^2\gamma^3\beta^4)V(Y)[E(Y)]^2\\
&\quad+\iota\eta\delta\gamma\beta(1-\beta-2\gamma\beta-4\delta\gamma\beta-8\eta\delta\gamma\beta+2\gamma\beta^2+4\delta\gamma\beta^2+6\delta\gamma^2\beta^2+8\eta\delta\gamma\beta^2+12\eta\delta\gamma^2\beta^2+18\eta\delta^2\gamma^2\beta^2\\
&\quad\quad-6\delta\gamma^2\beta^3-12\eta\delta\gamma^2\beta^3-18\eta\delta^2\gamma^2\beta^3-24\eta\delta^2\gamma^3\beta^3+24\eta\delta^2\gamma^3\beta^4)V(Y).
\end{aligned}$$

We are ready to prove Proposition 2.1. Hereafter, $\langle N\rangle$ means that there are similar N terms obtained under index permutations

Third autocumulant function

We have, for $v \geq u \geq 0$,

$$\begin{aligned} Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t) &= E[(Y_t - \mu_Y)(\alpha^u \circ Y_t - \alpha^u \mu_Y)(\alpha^v \circ Y_t - \alpha^v \mu_Y)] \\ &= \alpha^{v+u}\kappa_{3,Y} + \alpha^v(1-\alpha^u)\sigma_Y^2 \quad \text{(by Lemma 2.3)} \\ &= \alpha^{v+u}(\kappa_{3,Y} - \sigma_Y^2) + \alpha^v \sigma_Y^2. \end{aligned}$$

Fourth autocumulant function

Using

$$Cov(\beta \circ Y, \gamma \circ \beta \circ Y) = \gamma V(\beta \circ Y) = \gamma\beta^2 V(Y) + \gamma\beta(1-\beta)E(Y) \quad \text{(by Corollaries 2.1 and 2.2)},$$

we have, for $w \geq v \geq u \geq 0$,

$$\begin{aligned} & Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t, \alpha^w \circ Y_t) \\ &= E[(Y_t - \mu_Y)(\alpha^u \circ Y_t - \alpha^u \mu_Y)(\alpha^v \circ Y_t - \alpha^v \mu_Y)(\alpha^w \circ Y_t - \alpha^w \mu_Y)] \\ &\quad - \Big[Cov(Y_t, \alpha^u \circ Y_t) Cov(\alpha^v \circ Y_t, \alpha^w \circ Y_t)\langle 3\rangle\Big] \\ &= \alpha^{w+v+u} E[(Y_t - \mu_Y)^4] + (\alpha^{w+u} + 2\alpha^{w+v} - 3\alpha^{w+v+u})(\kappa_{3,Y} + \sigma_Y^2 \mu_Y) \\ &\quad + (\alpha^w - \alpha^{w+u} - 2\alpha^{w+v} + 2\alpha^{w+v+u})\sigma_Y^2 \\ &\quad - \Big[\alpha^u \sigma_Y^2 \{\alpha^{w+v}\sigma_Y^2 + \alpha^w(1-\alpha^v)\mu_Y\}\langle 3\rangle\Big] \quad \text{(by Lemma 2.4)} \\ &= \alpha^{w+v+u}(\kappa_{4,Y} - 3\kappa_{3,Y} + 2\sigma_Y^2) + (\alpha^{w+u} + 2\alpha^{w+v})(\kappa_{3,Y} - \sigma_Y^2) + \alpha^w \sigma_Y^2. \end{aligned}$$

Fifth autocumulant function

Using

$$\begin{aligned} & Cum(\beta \circ Y, \gamma \circ \beta \circ Y, \delta \circ \gamma \circ \beta \circ Y) \\ &= \delta\gamma^2 Cum_3(\beta \circ Y) + \delta\gamma(1-\gamma)V(\beta \circ Y) \quad \text{(by Lemmas 2.1(ii) and 2.3)} \\ &= \delta\gamma^2\beta^3 Cum_3(Y) + 3\delta\gamma^2\beta^2(1-\beta)V(Y) + \delta\gamma^2\beta(1-\beta)(1-2\beta)E(Y) \\ &\quad + \delta\gamma(1-\gamma)\beta^2 V(Y) + \delta\gamma(1-\gamma)\beta(1-\beta)E(Y) \quad \text{(by Corollary 2.2)} \\ &= \delta\gamma^2\beta^3 Cum_3(Y) + (\delta\gamma\beta^2 + 2\delta\gamma^2\beta^2 - 3\delta\gamma^2\beta^3)V(Y) + (\delta\gamma\beta - \delta\gamma\beta^2 - 2\delta\gamma^2\beta^2 + 2\delta\gamma^2\beta^3)E(Y), \end{aligned}$$

we have, for $x \geq w \geq v \geq u \geq 0$,

$$
\begin{aligned}
&\Big[Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t) Cov(\alpha^w \circ Y_t, \alpha^x \circ Y_t)\langle 6\rangle\Big] \\
&+ \Big[Cum(\alpha^u \circ Y_t, \alpha^v \circ Y_t, \alpha^w \circ Y_t) Cov(Y_t, \alpha^x \circ Y_t)\langle 4\rangle\Big] \\
&= \Big[\{\alpha^{v+u}\kappa_{3,Y} + \alpha^v(1-\alpha^u)\sigma_Y^2\}\{\alpha^{x+w}\sigma_Y^2 + \alpha^x(1-\alpha^w)\mu_Y\}\langle 6\rangle\Big] \\
&\quad + \Big[\{\alpha^{x+w+v}\kappa_{3,Y} + (\alpha^{x+v} + 2\alpha^{x+w} - 3\alpha^{x+w+v})\sigma_Y^2 + (\alpha^x - \alpha^{x+v} - 2\alpha^{x+w} + 2\alpha^{x+w+v})\mu_Y\}\alpha^u\sigma_Y^2\langle 4\rangle\Big] \\
&= 10\alpha^{x+w+v+u}\kappa_{3,Y}\sigma_Y^2 \\
&\quad + (\alpha^{x+v+u} + 2\alpha^{x+w+u} + 3\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\kappa_{3,Y}\mu_Y \\
&\quad + 3(\alpha^{x+v+u} + 2\alpha^{x+w+u} + 3\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\sigma_Y^4 \\
&\quad + (\alpha^{x+u} + 3\alpha^{x+v} + 6\alpha^{x+w} - 4\alpha^{x+v+u} - 8\alpha^{x+w+u} - 12\alpha^{x+w+v} + 14\alpha^{x+w+v+u})\sigma_Y^2\mu_Y,
\end{aligned}
$$

hence,

$$
\begin{aligned}
&Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t, \alpha^w \circ Y_t, \alpha^x \circ Y_t) \\
&= E[(Y_t - \mu_Y)(\alpha^u \circ Y_t - \alpha^u\mu_Y)(\alpha^v \circ Y_t - \alpha^v\mu_Y)(\alpha^w \circ Y_t - \alpha^w\mu_Y)(\alpha^x \circ Y_t - \alpha^x\mu_Y)] \\
&\quad - \Big[Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t) Cov(\alpha^w \circ Y_t, \alpha^x \circ Y_t)\langle 6\rangle\Big] \\
&\quad - \Big[Cum(\alpha^v \circ Y_t, \alpha^w \circ Y_t, \alpha^x \circ Y_t) Cov(Y_t, \alpha^u \circ Y_t)\langle 4\rangle\Big] \\
&= \alpha^{x+w+v+u}E[(Y_t - \mu_Y)^5] \\
&\quad + (\alpha^{x+v+u} + 2\alpha^{x+w+u} + 3\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\big\{E[(Y_t - \mu_Y)^4] + \kappa_{3,Y}\mu_Y\big\} \\
&\quad + (\alpha^{x+u} + 2\alpha^{x+v} + 4\alpha^{x+w} - 3\alpha^{x+v+u} - 6\alpha^{x+w+u} - 9\alpha^{x+w+v} + 11\alpha^{x+w+v+u})\kappa_{3,Y} \\
&\quad + (\alpha^{x+u} + 3\alpha^{x+v} + 6\alpha^{x+w} - 4\alpha^{x+v+u} - 8\alpha^{x+w+u} - 12\alpha^{x+w+v} + 14\alpha^{x+w+v+u})\sigma_Y^2\mu_Y \\
&\quad + (\alpha^x - \alpha^{x+u} - 2\alpha^{x+v} - 4\alpha^{x+w} + 2\alpha^{x+v+u} + 4\alpha^{x+w+u} + 6\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\sigma_Y^2 \\
&\quad - 10\alpha^{x+w+v+u}\kappa_{3,Y}\sigma_Y^2 \\
&\quad - (\alpha^{x+v+u} + 2\alpha^{x+w+u} + 3\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\kappa_{3,Y}\mu_Y \\
&\quad - (\alpha^{x+u} + 3\alpha^{x+v} + 6\alpha^{x+w} - 4\alpha^{x+v+u} - 8\alpha^{x+w+u} - 12\alpha^{x+w+v} + 14\alpha^{x+w+v+u})\sigma_Y^2\mu_Y \\
&\quad - 3(\alpha^{x+v+u} + 2\alpha^{x+w+u} + 3\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\sigma_Y^4 \quad \text{(by Lemma 2.5)} \\
&= \alpha^{x+w+v+u}\kappa_{5,Y} \\
&\quad + (\alpha^{x+v+u} + 2\alpha^{x+w+u} + 3\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\kappa_{4,Y} \\
&\quad + (\alpha^{x+u} + 2\alpha^{x+v} + 4\alpha^{x+w} - 3\alpha^{x+v+u} - 6\alpha^{x+w+u} - 9\alpha^{x+w+v} + 11\alpha^{x+w+v+u})\kappa_{3,Y} \\
&\quad + (\alpha^x - \alpha^{x+u} - 2\alpha^{x+v} - 4\alpha^{x+w} + 2\alpha^{x+v+u} + 4\alpha^{x+w+u} + 6\alpha^{x+w+v} - 6\alpha^{x+w+v+u})\sigma_Y^2 \\
&= \alpha^{x+w+v+u}(\kappa_{5,Y} - 6\kappa_{4,Y} + 11\kappa_{3,Y} - 6\sigma_Y^2) \\
&\quad + (\alpha^{x+v+u} + 2\alpha^{x+w+u} + 3\alpha^{x+w+v})(\kappa_{4,Y} - 3\kappa_{3,Y} + 2\sigma_Y^2)
\end{aligned}
$$

$$+ (\alpha^{x+u} + 2\alpha^{x+v} + 4\alpha^{x+w})(\kappa_{3,Y} - \sigma_Y^2)$$
$$+ \alpha^x \sigma_Y^2.$$

Sixth autocumulant function

We have, for $y \geq x \geq w \geq v \geq u \geq 0$,

$$\Big[Cov(Y_t, \alpha^u \circ Y_t)Cov(\alpha^v \circ Y_t, \alpha^w \circ Y_t)Cov(\alpha^x \circ Y_t, \alpha^y \circ Y_t)\langle 15\rangle\Big]$$
$$= \Big[\alpha^u \sigma_Y^2 \{\alpha^{x+v}(\sigma_Y^2 - \mu_Y) + \alpha^x \mu_Y\}\{\alpha^{y+w}(\sigma_Y^2 - \mu_Y) + \alpha^y \mu_Y\}\langle 15\rangle\Big]$$
$$= 15\alpha^{y+x+w+v+u}(\sigma_Y^2 - \mu_Y)^2 \sigma_Y^2$$
$$+ 3(\alpha^{y+w+v+u} + 2\alpha^{y+x+v+u} + 3\alpha^{y+x+w+u} + 4\alpha^{y+x+w+v})(\sigma_Y^2 - \mu_Y)\mu_Y \sigma_Y^2$$
$$+ (\alpha^{y+w+u} + 2\alpha^{y+w+v} + 2\alpha^{y+x+u} + 4\alpha^{y+x+v} + 6\alpha^{y+x+w})\sigma_Y^2 \mu_Y^2$$
$$= 15\alpha^{y+x+w+v+u}\sigma_Y^6$$
$$+ 3(\alpha^{y+w+v+u} + 2\alpha^{y+x+v+u} + 3\alpha^{y+x+w+u} + 4\alpha^{y+x+w+v} - 10\alpha^{y+x+w+v+u})\sigma_Y^4 \mu_Y$$
$$+ (\alpha^{y+w+u} + 2\alpha^{y+w+v} + 2\alpha^{y+x+u} + 4\alpha^{y+x+v} + 6\alpha^{y+x+w} - 3\alpha^{y+w+v+u} - 6\alpha^{y+x+v+u}$$
$$- 9\alpha^{y+x+w+u} - 12\alpha^{y+x+w+v} + 15\alpha^{y+x+w+v+u})\sigma_Y^2 \mu_Y^2,$$
$$\Big[Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t)Cum(\alpha^w \circ Y_t, \alpha^x \circ Y_t, \alpha^y \circ Y_t)\langle 10\rangle\Big]$$
$$= \Big[\{\alpha^{v+u}\kappa_{3,Y} + \alpha^v(1 - \alpha^u)\sigma_Y^2\}$$
$$\times \{\alpha^{y+x+w}\kappa_{3,Y} + (\alpha^{y+w} + 2\alpha^{y+x} - 3\alpha^{y+x+w})\sigma_Y^2 + (\alpha^y - \alpha^{y+w} - 2\alpha^{y+x} + 2\alpha^{y+x+w})\mu_Y\}\langle 10\rangle\Big]$$
$$= 10\alpha^{y+x+w+v+u}\kappa_{3,Y}^2$$
$$+ (4\alpha^{y+w+v+u} + 8\alpha^{y+x+v+u} + 12\alpha^{y+x+w+u} + 16\alpha^{y+x+w+v} - 40\alpha^{y+x+w+v+u})\kappa_{3,Y}\sigma_Y^2$$
$$+ (2\alpha^{y+w+u} + 4\alpha^{y+w+v} + 4\alpha^{y+x+u} + 8\alpha^{y+x+v} + 12\alpha^{y+x+w}$$
$$- 6\alpha^{y+w+v+u} - 12\alpha^{y+x+v+u} - 18\alpha^{y+x+w+u} - 24\alpha^{y+x+w+v} + 30\alpha^{y+x+w+v+u})\sigma_Y^4$$
$$+ (\alpha^{y+v+u} + \alpha^{y+w+u} + \alpha^{y+w+v} + 2\alpha^{y+x+u} + 2\alpha^{y+x+v} + 3\alpha^{y+x+w}$$
$$- 3\alpha^{y+w+v+u} - 6\alpha^{y+x+v+u} - 9\alpha^{y+x+w+u} - 12\alpha^{y+x+w+v} + 20\alpha^{y+x+w+v+u})\kappa_{3,Y}\mu_Y$$
$$+ (\alpha^{y+v} + 3\alpha^{y+w} + 6\alpha^{y+x} - \alpha^{y+v+u} - 3\alpha^{y+w+u} - 5\alpha^{y+w+v} - 6\alpha^{y+x+u} - 10\alpha^{y+x+v} - 15\alpha^{y+x+w}$$
$$+ 5\alpha^{y+w+v+u} + 10\alpha^{y+x+v+u} + 15\alpha^{y+x+w+u} + 20\alpha^{y+x+w+v} - 20\alpha^{y+x+w+v+u})\sigma_Y^2 \mu_Y,$$
$$\Big[Cum(Y_t, \alpha^w \circ Y_t, \alpha^x \circ Y_t, \alpha^y \circ Y_t)Cov(\alpha^u \circ Y_t, \alpha^v \circ Y_t)\langle 10\rangle\Big]$$
$$= \Big[\{\alpha^{y+x+w}\kappa_{4,Y} + (\alpha^{y+w} + 2\alpha^{y+x} - 3\alpha^{y+x+w})\kappa_{3,Y} + (\alpha^y - \alpha^{y+w} - 2\alpha^{y+x} + 2\alpha^{y+x+w})\sigma_Y^2\}$$
$$\times \{\alpha^{v+u}\sigma_Y^2 + \alpha^v(1 - \alpha^u)\mu_Y\}\langle 10\rangle\Big]$$
$$= 10\alpha^{y+x+w+v+u}\kappa_{4,Y}(\sigma_Y^2 - \mu_Y)$$
$$+ (\alpha^{y+w+v+u} + 2\alpha^{y+x+v+u} + 3\alpha^{y+x+w+u} + 4\alpha^{y+x+w+v})\kappa_{4,Y}\mu_Y$$

$$
\begin{aligned}
&+ (3\alpha^{y+w+v+u} + 6\alpha^{y+x+v+u} + 9\alpha^{y+x+w+u} + 12\alpha^{y+x+w+v} - 30\alpha^{y+x+w+v+u})\kappa_{3,Y}(\sigma_Y^2 - \mu_Y) \\
&+ (2\alpha^{y+w+u} + 4\alpha^{y+w+v} + 4\alpha^{y+x+u} + 8\alpha^{y+x+v} + 12\alpha^{y+x+w} - 3\alpha^{y+w+v+u} - 6\alpha^{y+x+v+u} - 9\alpha^{y+x+w+u} \\
&\quad - 12\alpha^{y+x+w+v})\kappa_{3,Y}\mu_Y \\
&+ (\alpha^{y+v+u} + \alpha^{y+w+u} + \alpha^{y+w+v} + 2\alpha^{y+x+u} + 2\alpha^{y+x+v} + 3\alpha^{y+x+w} - 3\alpha^{y+w+v+u} - 6\alpha^{y+x+v+u} \\
&\quad - 9\alpha^{y+x+w+u} - 12\alpha^{y+x+w+v} + 20\alpha^{y+x+w+v+u})\sigma_Y^2(\sigma_Y^2 - \mu_Y) \\
&+ (\alpha^{y+v} + 3\alpha^{y+w} + 6\alpha^{y+x} - 2\alpha^{y+w+u} - 4\alpha^{y+w+v} - 4\alpha^{y+x+u} - 8\alpha^{y+x+v} - 12\alpha^{y+x+w} \\
&\quad + 2\alpha^{y+w+v+u} + 4\alpha^{y+x+v+u} + 6\alpha^{y+x+w+u} + 8\alpha^{y+x+w+v})\sigma_Y^2\mu_Y \\
= {}& 10\alpha^{y+x+w+v+u}\kappa_{4,Y}\sigma_Y^2 \\
&+ (\alpha^{y+w+v+u} + 2\alpha^{y+x+v+u} + 3\alpha^{y+x+w+u} + 4\alpha^{y+x+w+v} - 10\alpha^{y+x+w+v+u})\kappa_{4,Y}\mu_Y \\
&+ (3\alpha^{y+w+v+u} + 6\alpha^{y+x+v+u} + 9\alpha^{y+x+w+u} + 12\alpha^{y+x+w+v} - 30\alpha^{y+x+w+v+u})\kappa_{3,Y}\sigma_Y^2 \\
&+ (2\alpha^{y+w+u} + 4\alpha^{y+w+v} + 4\alpha^{y+x+u} + 8\alpha^{y+x+v} + 12\alpha^{y+x+w} - 6\alpha^{y+w+v+u} - 12\alpha^{y+x+v+u} - 18\alpha^{y+x+w+u} \\
&\quad - 24\alpha^{y+x+w+v} + 30\alpha^{y+x+w+v+u})\kappa_{3,Y}\mu_Y \\
&+ (\alpha^{y+v+u} + \alpha^{y+w+u} + \alpha^{y+w+v} + 2\alpha^{y+x+u} + 2\alpha^{y+x+v} + 3\alpha^{y+x+w} - 3\alpha^{y+w+v+u} - 6\alpha^{y+x+v+u} \\
&\quad - 9\alpha^{y+x+w+u} - 12\alpha^{y+x+w+v} + 20\alpha^{y+x+w+v+u})\sigma_Y^4 \\
&+ (\alpha^{y+v} + 3\alpha^{y+w} + 6\alpha^{y+x} - \alpha^{y+v+u} - 3\alpha^{y+w+u} - 5\alpha^{y+w+v} - 6\alpha^{y+x+u} - 10\alpha^{y+x+v} - 15\alpha^{y+x+w} \\
&\quad + 5\alpha^{y+w+v+u} + 10\alpha^{y+x+v+u} + 15\alpha^{y+x+w+u} + 20\alpha^{y+x+w+v} - 20\alpha^{y+x+w+v+u})\sigma_Y^2\mu_Y.
\end{aligned}
$$

Further, using

$$
\begin{aligned}
&Cum[\beta \circ Y, \gamma \circ (\beta \circ Y), \delta \circ (\gamma \circ (\beta \circ Y)), \eta \circ (\delta \circ (\gamma \circ (\beta \circ Y)))] \\
&= \eta\delta^2\gamma^3 Cum_4(\beta \circ Y) + \eta\delta\gamma^2(1 + 2\delta - 3\delta\gamma)Cum_3(\beta \circ Y) + \eta\delta\gamma(1-\gamma)(1-2\delta\gamma)V(\beta \circ Y) \\
&\quad \text{(by Lemma 2.1(ii) and (B))} \\
&= \eta\delta^2\gamma^3\{\beta^4 Cum_4(Y) + 6\beta^3(1-\beta)Cum_3(Y) + \beta^2(1-\beta)(7-11\beta)V(Y) + \beta(1-\beta)(1-6\beta+6\beta^2)E(Y)\} \\
&\quad + \eta\delta\gamma^2(1+2\delta-3\delta\gamma)\{\beta^3 Cum_3(Y) + 3\beta^2(1-\beta)V(Y) + \beta(1-\beta)(1-2\beta)E(Y)\} \\
&\quad + \eta\delta\gamma(1-\gamma)(1-2\delta\gamma)\{\beta^2 V(Y) + \beta(1-\beta)E(Y)\} \quad \text{(by Corollary 2.2)} \\
&= \eta\delta^2\gamma^3\beta^4 Cum_4(Y) \\
&\quad + (\eta\delta\gamma^2\beta^3 + 2\eta\delta^2\gamma^2\beta^3 + 3\eta\delta^2\gamma^3\beta^3 - 6\eta\delta^2\gamma^3\beta^4)Cum_3(Y) \\
&\quad + (\eta\delta\gamma\beta^2 + 2\eta\delta\gamma^2\beta^2 + 4\eta\delta^2\gamma^2\beta^2 - 3\iota\eta\delta^2\gamma^3 - 6\eta\delta^2\gamma^2\beta^3 - 9\eta\delta^2\gamma^3\beta^3 + 11\eta\delta^3\gamma^3\beta^4)V(Y) \\
&\quad + (\eta\delta\gamma\beta - \eta\delta\gamma\beta^2 - 2\eta\delta\gamma^2\beta^2 - 4\eta\delta^2\gamma^2\beta^2 + 2\eta\delta\gamma^2\beta^3 + 4\eta\delta^2\gamma^2\beta^3 + 6\eta\delta^2\gamma^2\beta^3 - 6\eta\delta^2\gamma^3\beta^4)E(Y),
\end{aligned}
$$

we have

$$
\Big[Cov(Y_t, \alpha^u \circ Y_t)Cum(\alpha^v \circ Y_t, \alpha^w \circ Y_t, \alpha^x \circ Y_t, \alpha^y \circ Y_t)\langle 5\rangle\Big]
$$

$$= \Big[\alpha^u\sigma_Y^2\{\alpha^{y+x+w+v}\kappa_{4,Y} + (\alpha^{y+w+v} + 2\alpha^{y+x+v} + 3\alpha^{y+x+w} - 6\alpha^{y+x+w+v})\kappa_{3,Y}$$

$$+ (\alpha^{y+v} + 2\alpha^{y+w} - 3\alpha^{y+w+v} + 4\alpha^{y+x} - 6\alpha^{y+x+v} - 9\alpha^{y+x+w} + 11\alpha^{y+x+w+v})\sigma_Y^2$$

$$+ (\alpha^{y} - \alpha^{y+v} - 2\alpha^{y+w} + 2\alpha^{y+w+v} - 4\alpha^{y+x} + 4\alpha^{y+x+v} + 6\alpha^{y+x+w} - 6\alpha^{y+x+w+v})\mu_Y\}\langle 5\rangle\Big]$$

$$= 5\alpha^{y+x+w+v+u}\kappa_{4,Y}\sigma_Y^2$$

$$+ (3\alpha^{y+w+v+u} + 6\alpha^{y+x+v+u} + 9\alpha^{y+x+w+u} + 12\alpha^{y+x+w+v} - 30\alpha^{y+x+w+v+u})\kappa_{3,Y}\sigma_Y^2$$

$$+ (2\alpha^{y+v+u} + 3\alpha^{y+w+u} + 4\alpha^{y+w+v} + 6\alpha^{y+x+u} + 8\alpha^{y+x+v} + 12\alpha^{y+x+w}$$

$$- 9\alpha^{y+w+v+u} - 18\alpha^{y+x+v+u} - 27\alpha^{y+x+w+u} - 36\alpha^{y+x+w+v} + 55\alpha^{y+x+w+v+u})\sigma_Y^4$$

$$+ (\alpha^{y+u} + \alpha^{y+v} + \alpha^{y+w} + 2\alpha^{y+x} - 2\alpha^{y+v+u} - 3\alpha^{y+w+u} - 4\alpha^{y+w+v} - 6\alpha^{y+x+u} - 8\alpha^{y+x+v} - 12\alpha^{y+x+w}$$

$$+ 6\alpha^{y+w+v+u} + 12\alpha^{y+x+v+u} + 18\alpha^{y+x+w+u} + 24\alpha^{y+x+w+v} - 30\alpha^{y+x+w+v+u})\sigma_Y^2\mu_Y.$$

Then, we have, for $y \geq x \geq w \geq v \geq u \geq 0$,

$$Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t, \alpha^w \circ Y_t, \alpha^x \circ Y_t, \alpha^y \circ Y_t)$$

$$= E[(Y_t - \mu_Y)(\alpha^u \circ Y_t - \alpha^u\mu_Y)(\alpha^v \circ Y_t - \alpha^v\mu_Y)(\alpha^w \circ Y_t - \alpha^w\mu_Y)(\alpha^x \circ Y_t - \alpha^x\mu_Y)]$$

$$- \Big[Cov(Y_t, \alpha^u \circ Y_t)Cov(\alpha^v \circ Y_t, \alpha^w \circ Y_t)Cov(\alpha^x \circ Y_t, \alpha^y \circ Y_t)\langle 15\rangle\Big]$$

$$- \Big[Cum(Y_t, \alpha^u \circ Y_t, \alpha^v \circ Y_t)Cum(\alpha^w \circ Y_t, \alpha^x \circ Y_t, \alpha^y \circ Y_t)\langle 10\rangle\Big]$$

$$- \Big[Cum(Y_t, \alpha^w \circ Y_t, \alpha^x \circ Y_t, \alpha^y \circ Y_t)Cov(\alpha^u \circ Y_t, \alpha^v \circ Y_t)\langle 10\rangle\Big]$$

$$- \Big[Cov(Y_t, \alpha^u \circ Y_t)Cum(\alpha^v \circ Y_t, \alpha^w \circ Y_t, \alpha^x \circ Y_t, \alpha^y \circ Y_t)\langle 5\rangle\Big]$$

$$= \alpha^{y+x+w+v+u}\{E[(Y_t - \mu_Y)^6] - 15\kappa_{4,Y}\sigma_Y^2 - 10\kappa_{3,Y}^2 - 15\sigma_Y^6\}$$

$$+ (\alpha^{y+w+v+u} + 2\alpha^{y+x+v+u} + 3\alpha^{y+x+w+u} + 4\alpha^{y+x+w+v} - 10\alpha^{y+x+w+v+u})\kappa_{5,Y}$$

$$+ (\alpha^{y+v+u} + 2\alpha^{y+w+u} + 3\alpha^{y+w+v} + 4\alpha^{y+x+u} + 6\alpha^{y+x+v} + 9\alpha^{y+x+w} - 6\alpha^{y+w+v+u} - 12\alpha^{y+x+v+u}$$

$$- 18\alpha^{y+x+w+u} - 24\alpha^{y+x+w+v} + 35\alpha^{y+x+w+v+u})\kappa_{4,Y}$$

$$+ (\alpha^{y+u} + 2\alpha^{y+v} + 4\alpha^{y+w} + 8\alpha^{y+x} - 3\alpha^{y+v+u} - 6\alpha^{y+w+u} - 9\alpha^{y+w+v} - 12\alpha^{y+x+u} - 18\alpha^{y+x+v}$$

$$- 27\alpha^{y+x+w} + 11\alpha^{y+w+v+u} + 22\alpha^{y+x+v+u} + 33\alpha^{y+x+w+u} + 44\alpha^{y+x+w+v} - 50\alpha^{y+x+w+v+u})\kappa_{3,Y}$$

$$+ (\alpha^{y} - \alpha^{y+u} - 2\alpha^{y+v} - 4\alpha^{y+w} - 8\alpha^{y+x} + 2\alpha^{y+v+u} + 4\alpha^{y+w+u} + 6\alpha^{y+w+v} + 8\alpha^{y+x+u} + 12\alpha^{y+x+v}$$

$$+ 18\alpha^{y+x+w} - 6\alpha^{y+w+v+u} - 12\alpha^{y+x+v+u} - 18\alpha^{y+x+w+u} - 24\alpha^{y+x+w+v} + 24\alpha^{y+x+w+v+u})\sigma_Y^2$$

(by Lemma 2.6)

$$= \alpha^{y+x+w+v+u}(\kappa_{6,Y} - 10\kappa_{5,Y} + 35\kappa_{4,Y} - 50\kappa_{3,Y} + 24\sigma_Y^2)$$

$$+ (\alpha^{y+w+v+u} + 2\alpha^{y+x+v+u} + 3\alpha^{y+x+w+u} + 4\alpha^{y+x+w+v})(\kappa_{5,Y} - 6\kappa_{4,Y} + 11\kappa_{3,Y} - 6\sigma_Y^2)$$

$$+ (\alpha^{y+v+u} + 2\alpha^{y+w+u} + 3\alpha^{y+w+v} + 4\alpha^{y+x+u} + 6\alpha^{y+x+v} + 9\alpha^{y+x+w})(\kappa_{4,Y} - 3\kappa_{3,Y} + 2\sigma_Y^2)$$

$$+ (\alpha^{y+u} + 2\alpha^{y+v} + 4\alpha^{y+w} + 8\alpha^{y+x})(\kappa_{3,Y} - \sigma_Y^2)$$

$+\alpha^y \sigma_Y^2.$

Proof of Lemmas 2.3–2.6

After some tedious algebra, we obtain

$$
\begin{aligned}
&E[G(Y)(\beta\circ Y-\beta E(Y))(\gamma\circ(\beta\circ Y)-\gamma\beta E(Y))]\\
&=\gamma E[G(Y)(\beta\circ Y-\beta E(Y))^2] \quad \text{(use the conditional expectation)}\\
&=\gamma\Big[\beta^2 E[G(Y)(Y-E(Y))^2]+\beta(1-\beta)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}\Big] \quad \text{(by Lemma 2.2(iii))},\\
&E[G(Y)(\beta\circ Y-\beta E(Y))(\gamma\circ(\beta\circ Y)-\gamma\beta E(Y))(\delta\circ(\gamma\circ(\beta\circ Y))-\delta\gamma\beta E(Y))]\\
&=\delta\Big[\gamma^2 E[G(Y)(\beta\circ Y-\beta E(Y))^3]\\
&\qquad+\gamma(1-\gamma)\Big\{E[G(Y)(\beta\circ Y-\beta E(Y))^2]+\beta E[G(Y)(\beta\circ Y-\beta E(Y))]E(Y)\Big\}\Big]\\
&\quad \text{(use the conditional expectation in relation to Lemma 2.3)}\\
&=\delta\gamma^2\Big[\beta^3 E[G(Y)(Y-E(Y))^3]+3\beta^2(1-\beta)\Big\{E[G(Y)(Y-E(Y))^2]+E[G(Y)(Y-E(Y))]E(Y)\Big\}\\
&\qquad\quad+\beta(1-\beta)(1-2\beta)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}\Big]\\
&\quad+\delta\gamma(1-\gamma)\Big[\beta^2 E[G(Y)(Y-E(Y))^2]+\beta(1-\beta)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}\\
&\qquad\qquad+\beta^2 E[G(Y)(Y-E(Y))]E(Y)\Big] \quad \text{(by Lemma 2.2(ii,iii))}\\
&=\delta\gamma^2\beta^3 E[G(Y)(Y-E(Y))^3]\\
&\quad+\delta\gamma\beta^2(1+2\gamma-3\gamma\beta)\Big\{E[G(Y)(Y-E(Y))^2]+E[G(Y)(Y-E(Y))]E(Y)\Big\}\\
&\quad+\delta\gamma\beta(1-\beta)(1-2\gamma\beta)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\},\\
&E[G(Y)(\beta\circ Y-\beta E(Y))(\gamma\circ(\beta\circ Y)-\gamma\beta E(Y))(\delta\circ(\gamma\circ(\beta\circ Y))-\delta\gamma\beta E(Y))\\
&\quad\times(\eta\circ(\delta\circ(\gamma\circ(\beta\circ Y)))-\eta\delta\gamma\beta E(Y))]\\
&=\eta\delta^2\gamma^3 E[G(Y)(\beta\circ Y-\beta E(Y))^4]\\
&\quad+\eta\delta\gamma^2(1+2\delta-3\delta\gamma)\Big\{E[G(Y)(\beta\circ Y-\beta E(Y))^3]+\beta E[G(Y)(\beta\circ Y-\beta E(Y))^2]E(Y)\Big\}\\
&\quad+\eta\delta\gamma^2(1-\gamma)(1-2\delta\gamma)\Big\{E[G(Y)(\beta\circ Y-\beta E(Y))^2]+\beta E[G(Y)(\beta\circ Y-\beta E(Y))]E(Y)\Big\}\\
&\quad \text{(use the conditional expectation in relation to Lemma 2.4)}\\
&=\eta\delta^2\gamma^3\Big[\beta^4 E[G(Y)(Y-E(Y))^4]+6\beta^3(1-\beta)\Big\{E[G(Y)(Y-E(Y))^3]+E[G(Y)(Y-E(Y))^2]E(Y)\Big\}\\
&\qquad\quad+\beta^2(1-\beta)(7-11\beta)E[G(Y)(Y-E(Y))^2]+2\beta^2(1-\beta)(5-7\beta)E[G(Y)(Y-E(Y))]E(Y)\\
&\qquad\quad+3\beta^2(1-\beta)^2 E[G(Y)][E(Y)]^2\Big]\\
&\quad+\eta\delta\gamma^2(1+2\delta-3\delta\gamma)\Big\{\beta^3 E[G(Y)(Y-E(Y))^3]\\
&\qquad\qquad+3\beta^2(1-\beta)\Big\{E[G(Y)(Y-E(Y))^2]+E[G(Y)(Y-E(Y))]E(Y)\Big\}
\end{aligned}
$$

$$+\beta\Big[\beta^2E[G(Y)(Y-E(Y))^2]$$
$$+\beta(1-\beta)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}\Big]E(Y)\Big\}$$
$$+\eta\delta\gamma^2(1-\gamma)(1-2\delta\gamma)\Big[\beta^2E[G(Y)(Y-E(Y))^2]+\beta(1-\beta)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}$$
$$+\beta^2E[G(Y)(\beta Y-E(Y))]E(Y)\Big]\quad\text{(by Lemma 2.2(ii,iii))}$$
$$=\eta\delta^2\gamma^3\beta^4E[G(Y)(Y-E(Y))^4]$$
$$+\eta\delta\gamma\beta(\gamma\beta^2+2\delta\gamma\beta^2+3\delta\gamma^2\beta^2-6\delta\gamma^2\beta^3)\{E[G(Y)(Y-E(Y))^3]+E[G(Y)(Y-E(Y))^2]E(Y)\}$$
$$+\eta\delta\gamma\beta(\beta+2\gamma\beta+4\delta\gamma\beta-3\gamma\beta^2-6\delta\gamma\beta^2-9\delta\gamma^2\beta^2+11\delta\gamma^2\beta^3)E[G(Y)(Y-E(Y))^2]$$
$$+\eta\delta\gamma\beta(\beta+3\gamma\beta+6\delta\gamma\beta-4\gamma\beta^2-8\delta\gamma\beta^2-12\delta\gamma^2\beta^2+14\delta\gamma^2\beta^3)E[G(Y)(Y-E(Y))]E(Y)$$
$$+\eta\delta\gamma\beta(\gamma\beta+2\delta\gamma\beta-\gamma\beta^2-2\delta\gamma\beta^2-3\delta\gamma^2\beta^2+3\delta\gamma^2\beta^3)E[G(Y)][E(Y)]^2$$
$$+\eta\delta\gamma\beta(1-\beta-2\gamma\beta-4\delta\gamma\beta+2\gamma\beta^2+4\delta\gamma\beta^2+6\delta\gamma^2\beta^2-6\delta\gamma^2\beta^3)$$
$$\times\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\},$$
$$E[(Y-E(Y))(\beta\circ Y-\beta E(Y))(\gamma\circ(\beta\circ Y)-\gamma\beta E(Y))(\delta\circ(\gamma\circ(\beta\circ Y))-\delta\gamma\beta E(Y))$$
$$\times(\eta\circ(\delta\circ(\gamma\circ(\beta\circ Y)))-\eta\delta\gamma\beta E(Y))(\iota\circ(\eta\circ(\delta\circ(\gamma\circ(\beta\circ Y))))-\iota\eta\delta\gamma\beta E(Y))]$$
$$=\iota\eta^2\delta^3\gamma^4E[(Y-E(Y))(\beta\circ Y-\beta E(Y))^5]$$
$$+\iota\eta\delta\gamma(\delta\gamma^2+2\eta\delta\gamma^2+3\eta\delta^2\gamma^2-6\eta\delta^2\gamma^3)$$
$$\times\Big\{E[(Y-E(Y))(\beta\circ Y-\beta E(Y))^4]+E[(Y-E(Y))(\beta\circ Y-\beta E(Y))^3]\beta E(Y)\Big\}$$
$$+\iota\eta\delta\gamma(\gamma+2\delta\gamma+4\eta\delta\gamma-3\delta\gamma^2-6\eta\delta\gamma^2-9\eta\delta^2\gamma^2+11\eta\delta^2\gamma^3)E[(Y-E(Y))(\beta\circ Y-\beta E(Y))^3]$$
$$+\iota\eta\delta\gamma(\gamma+3\delta\gamma+6\eta\delta\gamma-4\delta\gamma^2-8\eta\delta\gamma^2-12\eta\delta^2\gamma^2+14\eta\delta^2\gamma^3)\beta E[(Y-E(Y))(\beta\circ Y-\beta E(Y))^2]E(Y)$$
$$+\iota\eta\delta\gamma(\delta\gamma+2\eta\delta\gamma-\delta\gamma^2-2\eta\delta\gamma^2-3\eta\delta^2\gamma^2+3\eta\delta^2\gamma^3)\beta^2E[(Y-E(Y))(\beta\circ Y-\beta E(Y))][E(Y)]^2$$
$$+\iota\eta\delta\gamma(1-\gamma-2\delta\gamma-4\eta\delta\gamma+2\delta\gamma^2+4\eta\delta\gamma^2+6\eta\delta^2\gamma^2-6\eta\delta^2\gamma^3)$$
$$\times\Big\{E[(Y-E(Y))(\beta\circ Y-\beta E(Y))^2]+\beta E[(Y-E(Y))(\beta\circ Y-\beta E(Y))]E(Y)\Big\}$$

(use the conditional expectation in relation to Lemma 2.5)

$$=\iota\eta^2\delta^3\gamma^4\Big[\beta^5E[(Y-E(Y))^6]+10(\beta^4-\beta^5)\Big\{E[(Y-E(Y))^5]+E[(Y-E(Y))^4]E(Y)\Big\}$$
$$+5(5\beta^3-12\beta^4+7\beta^5)E[(Y-E(Y))^4]+10(4\beta^3-9\beta^4+5\beta^5)E[(Y-E(Y))^3]E(Y)$$
$$+5(3\beta^2-15\beta^3+22\beta^4-10\beta^5)E[(Y-E(Y))^3]+15(\beta^3-2\beta^4+\beta^5)V(Y)[E(Y)]^2$$
$$+5(5\beta^2-23\beta^3+32\beta^4-14\beta^5)V(Y)E(Y)+\iota\eta^2\delta^3\gamma^4(\beta-15\beta^2+50\beta^3-60\beta^4+24\beta^5)V(Y)\Big]$$
$$+\iota\eta\delta\gamma(\delta\gamma^2+2\eta\delta\gamma^2+3\eta\delta^2\gamma^2-6\eta\delta^2\gamma^3)$$
$$\times\Big[\beta^4E[(Y-E(Y))^5]+6(\beta^3-\beta^4)\Big\{E[(Y-E(Y))^4]+E[(Y-E(Y))^3]E(Y)\Big\}$$
$$+(7\beta^2-18\beta^3+11\beta^4)E[(Y-E(Y))^3]$$
$$+2(5\beta^2-12\beta^3+7\beta^4)V(Y)E(Y)+(\beta-7\beta^2+12\beta^3-6\beta^4)V(Y)$$

$$
\begin{aligned}
&+ \Big\{\beta^3 E[(Y-E(Y))^4] + 3\beta^2(1-\beta)\{E[(Y-E(Y))^3] + V(Y)E(Y)\} \\
&\quad + \beta(1-\beta)(1-2\beta)V(Y)\Big\}\beta E(Y)\Big] \\
&+ \iota\eta\delta\gamma(\gamma + 2\delta\gamma + 4\eta\delta\gamma - 3\delta\gamma^2 - 6\eta\delta\gamma^2 - 9\eta\delta^2\gamma^2 + 11\eta\delta^2\gamma^3) \\
&\quad \times \Big\{\beta^3 E[(Y-E(Y))^4] + 3\beta^2(1-\beta)\{E[(Y-E(Y))^3] + V(Y)E(Y)\} + \beta(1-\beta)(1-2\beta)V(Y)\Big\} \\
&+ \iota\eta\delta\gamma(\gamma + 3\delta\gamma + 6\eta\delta\gamma - 4\delta\gamma^2 - 8\eta\delta\gamma^2 - 12\eta\delta^2\gamma^2 + 14\eta\delta^2\gamma^3)\beta \\
&\quad \times \Big\{\beta^2 E[(Y-E(Y))^3] + \beta(1-\beta)V(Y)\Big\}E(Y) \\
&+ \iota\eta\delta\gamma(\delta\gamma + 2\eta\delta\gamma - \delta\gamma^2 - 2\eta\delta\gamma^2 - 3\eta\delta^2\gamma^2 + 3\eta\delta^2\gamma^3)\beta^3 V(Y)[E(Y)]^2 \\
&+ \iota\eta\delta\gamma(1 - \gamma - 2\delta\gamma - 4\eta\delta\gamma + 2\delta\gamma^2 + 4\eta\delta\gamma^2 + 6\eta\delta^2\gamma^2 - 6\eta\delta^2\gamma^3) \\
&\quad \times \Big\{\beta^2 E[(Y-E(Y))^3] + \beta(1-\beta)V(Y) + \beta^2 V(Y)E(Y)\Big\} \quad \text{(by Lemma 2.2(ii,iii))} \\
&= \iota\eta^2\delta^3\gamma^4\beta^5 E[(Y-E(Y))^6] \\
&+ \iota\eta\delta^2\gamma^3\beta^4(1 + 2\eta + 3\eta\delta + 4\eta\delta\gamma - 10\eta\delta\gamma\beta)\Big\{E[(Y-E(Y))^5] + E[(Y-E(Y))^4]E(Y)\Big\} \\
&+ \iota\eta\delta\gamma\beta(\gamma\beta^2 + 2\delta\gamma\beta^2 + 3\delta\gamma^2\beta^2 + 4\eta\delta\gamma\beta^2 + 6\eta\delta\gamma^2\beta^2 + 9\eta\delta^2\gamma^2\beta^2 - 6\delta\gamma^2\beta^3 - 12\eta\delta\gamma^2\beta^3 \\
&\qquad - 18\eta\delta^2\gamma^2\beta^3 - 24\eta\delta^2\gamma^3\beta^3 + 35\eta\delta^2\gamma^3\beta^4)E[(Y-E(Y))^4] \\
&+ \iota\eta\delta\gamma\beta(\gamma\beta^2 + 3\delta\gamma\beta^2 + 5\delta\gamma^2\beta^2 + 6\eta\delta\gamma\beta^2 + 10\eta\delta\gamma^2\beta^2 + 15\eta\delta^2\gamma^2\beta^2 - 9\delta\gamma^2\beta^3 - 18\eta\delta\gamma^2\beta^3 \\
&\qquad - 27\eta\delta^2\gamma^2\beta^3 - 36\eta\delta^2\gamma^3\beta^3 + 50\eta\delta^2\gamma^3\beta^4)E[(Y-E(Y))^3]E(Y) \\
&+ \iota\eta\delta\gamma\beta(\beta + 2\gamma\beta + 4\delta\gamma\beta + 8\eta\delta\gamma\beta - 3\gamma\beta^2 - 6\delta\gamma\beta^2 - 9\delta\gamma^2\beta^2 - 12\eta\delta\gamma\beta^2 - 18\eta\delta\gamma^2\beta^2 - 27\eta\delta^2\gamma^2\beta^2 \\
&\qquad + 11\delta\gamma^2\beta^3 + 22\eta\delta\gamma^2\beta^3 + 33\eta\delta^2\gamma^2\beta^3 + 44\eta\delta^2\gamma^3\beta^3 - 50\eta\delta^2\gamma^3\beta^4)E[(Y-E(Y))^3] \\
&+ \iota\eta\delta\gamma\beta(\beta + 3\gamma\beta + 7\delta\gamma\beta + 14\eta\delta\gamma\beta - 4\gamma\beta^2 - 9\delta\gamma\beta^2 - 14\delta\gamma^2\beta^2 - 18\eta\delta\gamma\beta^2 - 28\eta\delta\gamma^2\beta^2 - 42\eta\delta^2\gamma^2\beta^2 \\
&\qquad + 16\delta\gamma^2\beta^3 + 32\eta\delta\gamma^2\beta^3 + 48\eta\delta^2\gamma^2\beta^3 + 64\eta\delta^2\gamma^3\beta^3 - 70\eta\delta^2\gamma^3\beta^4)V(Y)E(Y) \\
&+ \iota\eta\delta\gamma\beta(\delta\gamma\beta^2 + 2\delta\gamma^2\beta^2 + 2\eta\delta\gamma\beta^2 + 4\eta\delta\gamma^2\beta^2 + 6\eta\delta^2\gamma^2\beta^2 - 3\delta\gamma^2\beta^3 - 6\eta\delta\gamma^2\beta^3 - 9\eta\delta^2\gamma^2\beta^3 \\
&\qquad - 12\eta\delta^2\gamma^3\beta^3 + 15\eta\delta^2\gamma^3\beta^4)V(Y)[E(Y)]^2 \\
&+ \iota\eta\delta\gamma\beta(1 - \beta - 2\gamma\beta - 4\delta\gamma\beta - 8\eta\delta\gamma\beta + 2\gamma\beta^2 + 4\delta\gamma\beta^2 + 6\delta\gamma^2\beta^2 + 8\eta\delta\gamma\beta^2 + 12\eta\delta\gamma^2\beta^2 + 18\eta\delta^2\gamma^2\beta^2 \\
&\qquad - 6\delta\gamma^2\beta^3 - 12\eta\delta\gamma^2\beta^3 - 18\eta\delta^2\gamma^2\beta^3 - 24\eta\delta^2\gamma^3\beta^3 + 24\eta\delta^2\gamma^3\beta^4)V(Y).
\end{aligned}
$$

Chapter 3

Some estimators in INAR(1) process

3.1 Introduction

Al-Osh and Alzaid (1987) mainly considered the nonnegative integer-valued autoregressive process of the first-order (INAR(1)) with the Poisson marginals, i.e., $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$, where $\alpha\circ$ is the thinning operator with $\alpha \in [0, 1)$, and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables according to the Poisson distribution $\text{Po}((1 - \alpha)\mu)$ for $\mu > 0$. They studied the Yule–Walker (YW), conditional least squares (CLS), and (conditional or full) maximum likelihood (ML) estimators for the parameter α. Park and Oh (1997) additionally established the asymptotic normality of the YW and CLS estimators, and Freeland and McCabe (2005) pointed out that Park and Oh's asymptotic variance of the CLS estimator is incorrect and then showed that the CLS estimator is asymptotically equivalent to the YW estimator. Bourguignon and Vasconcellos (2015a) considered the stationary INAR(1) process under the power series innovation, and conducted the simulations for the YW, CLS, and conditional ML estimators for the parameter α.

On the other hand, these estimators are biased in a finite-sample. Some authors thus studied the bias-correction in the stationary INAR(1) process, whose autocorrelation at lag 1 is the same as the usual stationary autoregressive process of the first-order (AR(1)). Unlike Fujikoshi and Ochi (1984) and Kakizawa (1996), the higher-order comparison of the mean squared errors (MSEs) in the stationary INAR(1) process is, however, not fully discussed. Indeed, Bourguignon and Vasconcellos (2015b) and Weiß and Schweer (2016) studied the INAR(1) with the Poisson marginals only. Jung et al. (2006) did not derive the variance and MSE after the bias-correction. This chapter attempts to fill these gaps.

The rest of this chapter is organized as follows. In Section 3.2, we consider a class of estimators for the parameter α in the stationary INAR(1) process under a general innovation. In Section 3.3, we derive, explicitly, the asymptotic expansion for the bias of such a general estimator. We construct an analytical bias-corrected estimator for the parameter α and examine the MSE after the bias-correction. In Section 3.4, we conduct the simulation experiments to demonstrate the finite-sample performances of the estimators in the stationary INAR(1) process under the Poisson or negative binomial (NB) innovation.

Section 3.5 concludes this chapter. The technical proofs are postponed to Sections 3.6 and 3.7.

3.2 A class of estimators for the parameter α

Suppose that the observation $\{Y_1, \ldots, Y_n\}$ of length n is generated by $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$, where we assume that $\alpha \in [0, 1)$ (i.e., the INAR(1) process is strictly stationary and ergodic; see Du and Li (1991)) and that $E(\varepsilon_t^J)$ exists for some $J \geq 2$ (see Remark 2.1). Let $\overline{Y} = (1/n)\sum_{t=1}^{n} Y_t$. Also, I_S stands for the indicator of the set S.

Perhaps, the YW estimator

$$\widehat{\alpha}_{YW} = \frac{\sum_{t=2}^{n}(Y_t - \overline{Y})(Y_{t-1} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2} \tag{3.1}$$

is the simplest estimator for the parameter α, since the autocorrelation of the stationary INAR(1) process is equal to α, as in the usual stationary AR(1) process (e.g., Brockwell and Davis (1987)). On the other hand, the CLS method due to Klimko and Nelson (1978) is quite standard, based on the criterion

$$J(\alpha, \mu_\varepsilon) = \sum_{t=2}^{n}(Y_t - E[Y_t|Y_{t-1}])^2 = \sum_{t=2}^{n}(Y_t - \alpha Y_{t-1} - \mu_\varepsilon)^2, \tag{3.2}$$

hence, the CLS estimator for the parameter α is given by

$$\widehat{\alpha}_{CLS} = \frac{\sum_{t=2}^{n} Y_t Y_{t-1} - (n-1)^{-1}\sum_{t=2}^{n} Y_t \sum_{t=2}^{n} Y_{t-1}}{\sum_{t=2}^{n} Y_{t-1}^2 - (n-1)^{-1}\Big(\sum_{t=2}^{n} Y_{t-1}\Big)^2}. \tag{3.3}$$

The method of moment (MM) estimator from the estimating equation

$$\sum_{t=2}^{n}[(Y_t - \overline{Y})(Y_{t-1} - \overline{Y}) - \alpha(Y_{t-1} - \overline{Y})^2] = 0$$

is given by

$$\widehat{\alpha}_{MM} = \frac{\sum_{t=2}^{n}(Y_t - \overline{Y})(Y_{t-1} - \overline{Y})}{\sum_{t=2}^{n}(Y_{t-1} - \overline{Y})^2}. \tag{3.4}$$

We define a more general estimator $\widehat{\alpha}_{c_1,c_2}$ for the parameter α in the stationary INAR(1) process, given by

$$\widehat{\alpha}_{c_1,c_2} = \frac{\sum_{t=2}^{n}(Y_t - \overline{Y})(Y_{t-1} - \overline{Y})}{c_1(Y_1 - \overline{Y})^2 + \sum_{t=2}^{n-1}(Y_t - \overline{Y})^2 + c_2(Y_n - \overline{Y})^2}, \quad c_1, c_2 \geq 0. \tag{3.5}$$

Note that, in addition to the YW estimator $\widehat{\alpha}_{YW} = \widehat{\alpha}_{1,1}$ and the MM estimator $\widehat{\alpha}_{MM} = \widehat{\alpha}_{1,0}$, the Burg estimator $\widehat{\alpha}_{Burg} = \widehat{\alpha}_{1/2,1/2}$ is widely used in the usual stationary AR(1) process. We consider an analytical bias-correction of such a general estimator (3.5) and study the higher-order MSE comparison in the stationary INAR(1) process. It should be remarked that

$$\sum_{t=2}^{n}(Y_t - \overline{Y})(Y_{t-1} - \overline{Y}) - \Big[\sum_{t=2}^{n} Y_t Y_{t-1} - \frac{1}{n-1}\sum_{t=2}^{n} Y_t \sum_{t=2}^{n} Y_{t-1}\Big] = \frac{(\overline{Y} - Y_1)(\overline{Y} - Y_n)}{n-1} = O_p(n^{-1}),$$

$$\sum_{t=2}^{n}(Y_{t-1} - \overline{Y})^2 - \Big[\sum_{t=2}^{n} Y_{t-1}^2 - \frac{1}{n-1}\Big(\sum_{t=2}^{n} Y_{t-1}\Big)^2\Big] = \frac{(\overline{Y} - Y_1)^2}{n-1} = O_p(n^{-1}),$$

hence, all results of $\widehat{\alpha}_{MM}$ described below remain valid for $\widehat{\alpha}_{CLS}$, if $(c_1, c_2) = (1, 0)$.

Remark 3.1. The conditional ML method is available, when the distribution of the innovation $\{\varepsilon_t\}$ is known. However, we here do not assume the parametric family of the distribution of the innovation, except for some simulation experiments. We emphasize that, in this sense, the study of this chapter can be regarded as to be semi-parametric inference.

3.3 Analytical bias-correction

First of all, we prove the strong consistency and asymptotic normality of the general estimator $\widehat{\alpha}_{c_1,c_2}$, and derive the asymptotic expansion for the bias $E(\widehat{\alpha}_{c_1,c_2} - \alpha)$.

Proposition 3.1. *(i)* $\widehat{\alpha}_{c_1,c_2} \xrightarrow{a.s.} \alpha$.

(ii) $\sqrt{n}(\widehat{\alpha}_{c_1,c_2} - \alpha) \xrightarrow{d} \mathrm{N}(0, v/\sigma_Y^4)$, *where*

$$v = (1 - \alpha)\{\alpha Q_{2:3,Y} + \alpha\sigma_Y^2 + (1 + \alpha)\sigma_Y^4\}.$$

Therefore, the general estimator $\widehat{\alpha}_{c_1,c_2}$ *is asymptotically equivalent, regardless of the choice of* (c_1, c_2)*.*

Proposition 3.2. *The* n^{-1} *bias of* $\widehat{\alpha}_{c_1,c_2}$ *is given by*

$$E(\widehat{\alpha}_{c_1,c_2} - \alpha) = -\frac{1}{n}\Big\{1 + (2 + c)\alpha + \frac{2\alpha^2 Q_{2:3,Y}}{(1 + \alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2}\Big\} + o(n^{-1}),$$

where $c = c_1 + c_2$*.*

Secondly, we construct an analytical bias-corrected estimator

$$\widetilde{\alpha}_{c_1,c_2} = \widehat{\alpha}_{c_1,c_2} + \frac{1}{n}\Big\{1 + (2+c)\widehat{\alpha}_{c_1,c_2} + \frac{2\widehat{\alpha}^2_{c_1,c_2}\widehat{Q}_{2:3,Y}}{(1+\widehat{\alpha}_{c_1,c_2})\widehat{\sigma}^4_Y} + \frac{\widehat{\alpha}_{c_1,c_2}}{\widehat{\sigma}^2_Y}\Big\} \tag{3.6}$$

(it is also consistent and asymptotically normal, as in Proposition 3.1; the detail is omitted here), where

$$\widehat{\sigma}^2_Y = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \overline{Y})^2,\ \ \widehat{Q}_{2:3,Y} = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \overline{Y})^3 - \frac{1}{n}\sum_{t=1}^{n}(Y_t - \overline{Y})^2$$

converge to σ^2_Y and $Q_{2:3,Y}$ a.s., respectively, by strictly stationarity and ergodicity of $\{Y_t\}$.

Our interest here is to examine the effect of the parameter (c_1, c_2) in terms of the MSE.

Proposition 3.3. *(i) The n^{-1} bias of $\widehat{\alpha}_{c_1,c_2}$ is removed, i.e., $E(\widetilde{\alpha}_{c_1,c_2} - \alpha) = o(n^{-1})$.*

(ii) Also,

$$V(\widetilde{\alpha}_{c_1,c_2}) - V(\widetilde{\alpha}_{0,0}) = \frac{A_{c_1,c_2}}{n^2} + o(n^{-2}),$$
$$MSE(\widetilde{\alpha}_{c_1,c_2}) - MSE(\widetilde{\alpha}_{0,0}) = \frac{A_{c_1,c_2}}{n^2} + o(n^{-2}),$$

where

$$A_{c_1,c_2} = \alpha^2\Big[(c_1^2 - 2c_1 + c_2^2)\frac{Q_{2:4,Y}}{\sigma^4_Y} + \{3(c_1^2 - 2c_1 + c_2^2) + (3c_1^2 - 2c_1 + 3c_2^2)\alpha\}\frac{Q_{2:3,Y}}{(1+\alpha)\sigma^4_Y} + (c_1^2 + c_2^2)\Big(\frac{1}{\sigma^2_Y} + 2\Big)\Big].$$

For the INAR(1) process $\{Y_t\}$ with the Poisson marginals ($Q_{2:4,Y} = Q_{2:3,Y} = 0$), the choice $(c_1, c_2) = (0, 0)$ is optimal, as in the usual stationary Gaussian AR(1) process (unfortunately, such a result, independent of α, does not hold for a general innovation).

The following two propositions reveal that, even under the Poisson marginals, there is no uniform results on the choice (c_1, c_2) of the estimator (without bias-correction) and the MSE comparison between $\widehat{\alpha}_{c_1,c_2}$ and $\widetilde{\alpha}_{c_1,c_2}$.

Proposition 3.4. *We have*

$$V(\widehat{\alpha}_{c_1,c_2}) - V(\widehat{\alpha}_{1,0}) = \frac{1}{n^2}\Big[\{(c_1-1)^2 + c_2^2\}\alpha^2\frac{Q_{2:4,Y}}{\sigma^4_Y} + A_{1,(c_1,c_2)}\frac{Q_{2:3,Y}}{(1+\alpha)\sigma^4_Y} + A_{2,(c_1,c_2)}\frac{1}{\sigma^2_Y} + A_{3,(c_1,c_2)}\Big] + o(n^{-2}),$$

$$MSE(\widehat{\alpha}_{c_1,c_2}) - MSE(\widehat{\alpha}_{1,0}) = \frac{1}{n^2}\Big[\{(c_1-1)^2 + c_2^2\}\alpha^2\frac{Q_{2:4,Y}}{\sigma^4_Y} + \{A_{1,(c_1,c_2)} + 4(c_1 + c_2 - 1)\alpha^3\}\frac{Q_{2:3,Y}}{(1+\alpha)\sigma^4_Y} + \{A_{2,(c_1,c_2)} + 2(c_1 + c_2 - 1)\alpha^2\}\frac{1}{\sigma^2_Y}$$

$$+ A_{3,(c_1,c_2)} + 2(c_1+c_2-1)\alpha + (c_1+c_2-1)(c_1+c_2+5)\alpha^2\Big] + o(n^{-2}),$$

where

$$\begin{aligned}
A_{1,(c_1,c_2)} &= \alpha[2(1-c_1-c_2) + 3\{(c_1-1)^2 + c_2^2\}\alpha + \{3(c_1-1)^2 + 6(c_1-1) + 3c_2^2 + 2c_2\}\alpha^2],\\
A_{2,(c_1,c_2)} &= \alpha[2(1-c_1-c_2) + \{(c_1-1)^2 + 4(c_1-1) + c_2^2 + 2c_2\}\alpha],\\
A_{3,(c_1,c_2)} &= 2(1-c_1-c_2) + 2\{(c_1-1)^2 + 3(c_1-1) + c_2^2 + c_2\}\alpha^2.
\end{aligned}$$

Proposition 3.5. *We have*

$$\begin{aligned}
V(\widehat{\alpha}_{c_1,c_2}) - V(\widetilde{\alpha}_{c_1,c_2}) &= \frac{1}{n^2}\Big[2c(\alpha-1)\Big(1+\alpha+\frac{\alpha Q_{2:3,Y}}{\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}\Big)+A\Big]+o(n^{-2}),\\
MSE(\widehat{\alpha}_{c_1,c_2}) - MSE(\widetilde{\alpha}_{c_1,c_2}) &= \frac{1}{n^2}\Big[c^2\alpha^2 + 2c\Big\{3\alpha^2+\alpha-1+\frac{(3\alpha^3-\alpha)Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{2\alpha^2-\alpha}{\sigma_Y^2}\Big\}\\
&\qquad + \Big\{1+2\alpha+\frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}\Big\}^2 + A\Big] + o(n^{-2}),
\end{aligned}$$

where

$$\begin{aligned}
A = 2\Big[&-\Big\{\frac{(4\alpha+2\alpha^2)Q_{2:3,Y}}{(1+\alpha)^2\sigma_Y^4}+\frac{1}{\sigma_Y^2}+2\Big\}(1-\alpha)\Big(\frac{\alpha Q_{2:3,Y}}{\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+1+\alpha\Big)\\
&+\Big\{\frac{4\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha+3\alpha^2}{(1+\alpha)\sigma_Y^2}\Big\}\Big\{\frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+2\alpha\Big\}\\
&-\frac{6\alpha^2}{(1+\alpha)\sigma_Y^2}\Big\{\frac{\alpha^3 Q_{2:4,Y}}{(1+\alpha+\alpha^2)\sigma_Y^4}+\alpha^2\Big(\frac{2}{1+\alpha}+\frac{1+\alpha}{1+\alpha+\alpha^2}+1\Big)\frac{Q_{2:3,Y}}{\sigma_Y^4}+\frac{2\alpha}{\sigma_Y^2}+\alpha\Big\}\Big].
\end{aligned}$$

3.4 Simulation results

We conduct the simulations ($\alpha = 0.2, 0.5, 0.8$ and $n = 100, 200, 300$, with 2000 replications) about the estimators $\widehat{\alpha}_{c_1,c_2}$ and $\widetilde{\alpha}_{c_1,c_2}$ (see (3.5) and (3.6) with $(c_1, c_2) = (0,0), (1/2,1/2), (1,0), (1,1)$), under the set-up that the innovation $\{\varepsilon_t\}$ follows the Poisson distribution $\mathrm{Po}((1-\alpha)\mu)$ or negative binomial (NB) distribution $\mathrm{NB}(r, r/\{r+(1-\alpha)\mu\})$ (we set $\mu = 10$ and $r = 10$).

We observe from Tables 3.1 and 3.2 that the biases, variances, and MSEs of all estimators decrease as the sample size n increases (the estimators have the downward biases) and that, as expected from Proposition 3.3(i), the proposed analytical bias-correction works well. Under the Poisson innovation, the bias-corrected estimator $\widetilde{\alpha}_{0,0}$ has, overall, the smallest variance and MSE (see Table 3.1), which is in agreement with Proposition 3.3(ii).

It may be true that $\widetilde{\alpha}_{c_1,c_2}$ suffers from the inflation of the variance, because of the trade-off between the bias and the variance. But, Proposition 3.5 tells us that the uniform inferiority of $\mathrm{MSE}(\widetilde{\alpha}_{c_1,c_2})$ over $\mathrm{MSE}(\widehat{\alpha}_{c_1,c_2})$, with respect to α, is not asserted, at least asymptotically. Indeed, comparing the upper and lower panels in Tables 3.1 and 3.2, it is revealed that, if $\alpha = 0.8$ ($\alpha = 0.2$) is large, then, the bias-corrected estimator $\widetilde{\alpha}_{c_1,c_2}$ is likely to outperform (underperform) the bias-uncorrected estimator $\widehat{\alpha}_{c_1,c_2}$ in terms of the MSE.

3.5 Concluding remark

After deriving the asymptotic expansion for the bias of the general estimator $\widehat{\alpha}_{c_1,c_2}$, we have constructed an analytical bias-corrected estimator $\widetilde{\alpha}_{c_1,c_2}$ and examined the effect of the parameter (c_1, c_2) in terms of the MSE. It turns out that, for the INAR(1) process with the Poisson marginals, the choice $(c_1, c_2) = (0, 0)$ is optimal, as in the usual stationary Gaussian AR(1) process.

In Fujikoshi and Ochi (1984), Kakizawa (1996), and references therein, the so-called Edgeworth expansion yields a refinement of normal approximation for the distributions of some estimators in the usual stationary AR(1) process. Unfortunately, it can not be directly applied in the stationary INAR(1) process, since the INAR(1) process is discrete.

3.6 Proof of Proposition 3.1

Before proving Proposition 3.1, we prepare two lemmas (Lemmas 3.1 and 3.2).

Lemma 3.1. *Suppose that* $\alpha \in [0, 1)$ *and that* $E(\varepsilon_t^2)$ *exists. The following hold.*

(i) $\sqrt{n}(\overline{Y} - \mu_Y) = O_p(1)$.

(ii) $\overline{Y} \xrightarrow{a.s.} \mu_Y$. *Moreover, for* $k = 0, 1$, $\{1/(n-1)\} \sum_{t=2}^n (Y_{t-k} - \overline{Y})(Y_{t-1} - \overline{Y}) \xrightarrow{a.s.} \alpha^{1-k}\sigma_Y^2$.

Proof (i) Noting that

$$\begin{aligned} nV(\overline{Y}) &= \frac{1}{n}\Big\{\sum_{t=1}^n V(Y_t) + 2\sum_{s,t=1}^n I_{\{s<t\}} Cov(Y_s, Y_t)\Big\} \\ &= \sigma_Y^2 + \frac{2}{n}\sum_{s,t=1}^n I_{\{s<t\}}\alpha^{t-s}\sigma_Y^2 \\ &= \sigma_Y^2 + 2\sigma_Y^2\Big\{\frac{\alpha}{1-\alpha} - \frac{\alpha(1-\alpha^n)}{n(1-\alpha)^2}\Big\} \to \frac{1+\alpha}{1-\alpha}\sigma_Y^2, \end{aligned}$$

the result follows by Brockwell and Davies (1987).

(ii) For $k = 0, 1$, $\{1/(n-1)\}\sum_{t=2}^n Y_{t-k}Y_{t-1} \xrightarrow{a.s.} E(Y_{1-k}Y_0)$ and $\{1/(n-1)\}\sum_{t=2}^n, Y_{t-k} \xrightarrow{a.s.} \mu_Y$, as well as $\overline{Y} \xrightarrow{a.s.} \mu_Y$, by strictly stationarity and ergodicity of $\{Y_t\}$, hence,

$$\begin{aligned} \frac{1}{n-1}\sum_{t=2}^n (Y_{t-k} - \overline{Y})(Y_{t-1} - \overline{Y}) &= \frac{1}{n-1}\sum_{t=2}^n Y_{t-k}Y_{t-1} - \frac{1}{n-1}\sum_{t=2}^n (Y_{t-k} + Y_{t-1})\overline{Y} + \overline{Y}^2 \\ &\xrightarrow{a.s.} E(Y_{1-k}Y_0) - \mu_Y^2 = \gamma_Y(1-k). \quad \square \end{aligned}$$

Table 3.1: Biases, variances, MSEs of the estimators $\widehat{\alpha}_{c_1,c_2}$ (without the bias-correction) and $\widetilde{\alpha}_{c_1,c_2}$ (with the bias-correction) in the stationary INAR(1) process unde the Poisson innovation.

		Biases (×10)				Variances (×100)				MSEs (×100)			
α	n	$\widehat{\alpha}_{0,0}$	$\widehat{\alpha}_{1/2,1/2}$	$\widehat{\alpha}_{1,0}$	$\widehat{\alpha}_{1,1}$	$\widehat{\alpha}_{0,0}$	$\widehat{\alpha}_{1/2,1/2}$	$\widehat{\alpha}_{1,0}$	$\widehat{\alpha}_{1,1}$	$\widehat{\alpha}_{0,0}$	$\widehat{\alpha}_{1/2,1/2}$	$\widehat{\alpha}_{1,0}$	$\widehat{\alpha}_{1,1}$
0.2	100	-0.164	-0.183	-0.183	-0.200	1.002	0.984	0.984	0.968	1.029	1.017	1.018	1.045
	200	-0.053	-0.063	-0.063	-0.072	0.481	0.477	0.476	0.472	0.484	0.481	0.480	0.492
	300	-0.047	-0.054	-0.054	-0.061	0.319	0.317	0.317	0.315	0.321	0.320	0.320	0.323
0.5	100	-0.227	-0.277	-0.275	-0.325	0.811	0.793	0.795	0.781	0.862	0.870	0.871	0.845
	200	-0.094	-0.119	-0.119	-0.143	0.413	0.408	0.408	0.405	0.422	0.422	0.422	0.422
	300	-0.069	-0.087	-0.087	-0.104	0.271	0.269	0.270	0.269	0.275	0.277	0.278	0.274
0.8	100	-0.280	-0.368	-0.369	-0.452	0.508	0.500	0.506	0.507	0.587	0.635	0.642	0.530
	200	-0.133	-0.175	-0.173	-0.217	0.202	0.202	0.203	0.206	0.220	0.233	0.233	0.207
	300	-0.088	-0.115	-0.115	-0.142	0.129	0.128	0.129	0.129	0.136	0.141	0.142	0.130
α	n	$\widetilde{\alpha}_{0,0}$	$\widetilde{\alpha}_{1/2,1/2}$	$\widetilde{\alpha}_{1,0}$	$\widetilde{\alpha}_{1,1}$	$\widetilde{\alpha}_{0,0}$	$\widetilde{\alpha}_{1/2,1/2}$	$\widetilde{\alpha}_{1,0}$	$\widetilde{\alpha}_{1,1}$	$\widetilde{\alpha}_{0,0}$	$\widetilde{\alpha}_{1/2,1/2}$	$\widetilde{\alpha}_{1,0}$	$\widetilde{\alpha}_{1,1}$
0.2	100	-0.025	-0.026	-0.027	-0.027	1.044	1.046	1.046	1.049	1.045	1.047	1.047	1.050
	200	0.017	0.017	0.017	0.018	0.492	0.491	0.491	0.492	0.492	0.492	0.491	0.492
	300	-0.0004	-0.001	-0.001	-0.001	0.323	0.324	0.324	0.324	0.323	0.324	0.324	0.324
0.5	100	-0.027	-0.031	-0.029	-0.033	0.845	0.843	0.845	0.847	0.845	0.843	0.846	0.848
	200	0.007	0.007	0.006	0.006	0.422	0.421	0.421	0.422	0.422	0.421	0.421	0.422
	300	-0.001	-0.002	-0.003	-0.004	0.274	0.275	0.276	0.276	0.274	0.275	0.276	0.276
0.8	100	-0.019	-0.033	-0.033	-0.044	0.530	0.531	0.537	0.549	0.530	0.532	0.539	0.551
	200	-0.001	-0.004	-0.002	-0.008	0.207	0.209	0.210	0.214	0.207	0.209	0.210	0.214
	300	0.001	-0.0003	-0.001	-0.001	0.130	0.131	0.131	0.132	0.130	0.131	0.131	0.132

Table 3.2: Biases, variances, MSEs of the estimators $\widehat{\alpha}_{c_1,c_2}$ (without the bias-correction) and $\widetilde{\alpha}_{c_1,c_2}$ (with the bias-correction) in the stationary INAR(1) process unde the NB innovation.

		Biases (×10)				Variances (×100)				MSEs (×100)			
α	n	$\widehat{\alpha}_{0,0}$	$\widehat{\alpha}_{1/2,1/2}$	$\widehat{\alpha}_{1,0}$	$\widehat{\alpha}_{1,1}$	$\widehat{\alpha}_{0,0}$	$\widehat{\alpha}_{1/2,1/2}$	$\widehat{\alpha}_{1,0}$	$\widehat{\alpha}_{1,1}$	$\widehat{\alpha}_{0,0}$	$\widehat{\alpha}_{1/2,1/2}$	$\widehat{\alpha}_{1,0}$	$\widehat{\alpha}_{1,1}$
0.2	100	-0.160	-0.179	-0.178	-0.197	0.999	0.979	0.979	0.960	1.025	1.011	1.011	1.042
	200	-0.077	-0.086	-0.086	-0.096	0.494	0.490	0.490	0.486	0.500	0.497	0.497	0.505
	300	-0.017	-0.023	-0.023	-0.029	0.333	0.331	0.331	0.329	0.333	0.331	0.331	0.339
0.5	100	-0.194	-0.250	-0.254	-0.303	0.849	0.830	0.830	0.818	0.887	0.892	0.894	0.886
	200	-0.096	-0.123	-0.124	-0.149	0.388	0.384	0.383	0.381	0.397	0.399	0.399	0.397
	300	-0.064	-0.083	-0.083	-0.101	0.261	0.259	0.260	0.258	0.265	0.266	0.267	0.265
0.8	100	-0.229	-0.335	-0.353	-0.435	0.497	0.480	0.483	0.485	0.549	0.592	0.608	0.519
	200	-0.102	-0.155	-0.163	-0.206	0.199	0.195	0.196	0.197	0.209	0.219	0.223	0.204
	300	-0.071	-0.105	-0.111	-0.138	0.133	0.132	0.132	0.133	0.138	0.143	0.144	0.136
α	n	$\widetilde{\alpha}_{0,0}$	$\widetilde{\alpha}_{1/2,1/2}$	$\widetilde{\alpha}_{1,0}$	$\widetilde{\alpha}_{1,1}$	$\widetilde{\alpha}_{0,0}$	$\widetilde{\alpha}_{1/2,1/2}$	$\widetilde{\alpha}_{1,0}$	$\widetilde{\alpha}_{1,1}$	$\widetilde{\alpha}_{0,0}$	$\widetilde{\alpha}_{1/2,1/2}$	$\widetilde{\alpha}_{1,0}$	$\widetilde{\alpha}_{1,1}$
0.2	100	-0.021	-0.022	-0.022	-0.023	1.041	1.040	1.041	1.041	1.042	1.041	1.041	1.041
	200	-0.006	-0.007	-0.007	-0.007	0.505	0.505	0.505	0.506	0.505	0.505	0.505	0.506
	300	0.030	0.031	0.031	0.031	0.338	0.338	0.338	0.338	0.339	0.339	0.339	0.339
0.5	100	0.007	-0.003	-0.006	-0.011	0.886	0.882	0.882	0.887	0.886	0.882	0.882	0.887
	200	0.006	0.003	0.002	0.001	0.397	0.396	0.396	0.397	0.397	0.396	0.396	0.397
	300	0.004	0.002	0.001	-0.0002	0.265	0.265	0.266	0.265	0.265	0.265	0.266	0.265
0.8	100	0.034	0.002	-0.017	-0.025	0.518	0.510	0.513	0.525	0.519	0.510	0.513	0.526
	200	0.031	0.017	0.008	0.004	0.203	0.201	0.202	0.205	0.204	0.201	0.202	0.205
	300	0.018	0.010	0.004	0.003	0.135	0.135	0.135	0.137	0.136	0.135	0.135	0.137

Let

$$M_t = \{Y_t - \alpha Y_{t-1} - (1-\alpha)\mu_Y\}(Y_{t-1} - \mu_Y), \ \ t = 2, \dots, n.$$

Lemma 3.2. *Suppose that* $\alpha \in [0,1)$ *and that* $E(\varepsilon_t^4)$ *exists. The following hold.*

$$(i) \ \ \frac{1}{\sqrt{n}} \sum_{t=2}^{n} M_t \xrightarrow{d} \mathrm{N}(0, v), \ \ \textit{where } v = (1-\alpha)[\alpha Q_{2:3,Y} + \alpha\sigma_Y^2 + (1+\alpha)\sigma_Y^4].$$

$$(ii) \ \ \frac{1}{\sqrt{n}} \sum_{t=2}^{n} [(Y_t - \overline{Y})(Y_{t-1} - \overline{Y}) - \alpha(Y_{t-1} - \overline{Y})^2] \xrightarrow{d} \mathrm{N}(0, v).$$

Proof (i) Let $\mathscr{F}_t = \sigma\{Y_1, \dots, Y_t\}$ be a sigma-field, and let $n \geq 2$. Note that $\{n^{-1/2}M_t, t = 2, \dots, n\}$ is a martingale difference array, since, for $t = 2, \dots, n$,

$$E(Y_t|\mathscr{F}_{t-1}) = E(\alpha \circ Y_{t-1} + \varepsilon_t|\mathscr{F}_{t-1}) = \alpha Y_{t-1} + (1-\alpha)\mu_Y, \ \text{ i.e., } \ E(n^{-1/2}M_t|\mathscr{F}_{t-1}) = 0.$$

Now, we know that $E(M_t^2) = E(M_2^2)(< \infty)$, by strictly stationarity of $\{Y_t\}$. Then, for any $\varepsilon > 0$, we have

$$\begin{aligned} P\Big(\max_{t=2,\dots,n} |n^{-1/2}M_t| > \varepsilon\Big) &\leq \sum_{t=2}^{n} P(|M_t| > \sqrt{n}\varepsilon) \\ &= (n-1)P(|M_2| > \sqrt{n}\varepsilon) \quad \text{(by strictly stationarity of } \{Y_t\}) \\ &\leq \frac{1}{\varepsilon^2} E[I_{\{|M_2| > \sqrt{n}\varepsilon\}} M_2^2] \\ &\to 0 \quad \text{(by Lebesgue's dominated convergence theorem),} \end{aligned}$$

hence,

$$\max_{t=2,\dots,n} |n^{-1/2}M_t| = o_p(1). \tag{3.7}$$

Also, we have

$$\frac{1}{n} E\Big[\max_{t=2,\dots,n} M_t^2\Big] \leq \frac{1}{n} \sum_{t=2}^{n} E(M_t^2) \leq E(M_2^2), \tag{3.8}$$

$$\frac{1}{n} \sum_{t=2}^{n} M_t^2 \xrightarrow{a.s.} E(M_2^2) \quad \text{(by strictly stationarity and ergodicity of } \{Y_t\}). \tag{3.9}$$

The result is shown by martingale central limit theorem (see McLeish (1974)) and (3.7)–(3.9).

(ii) Noting that

$$\frac{1}{\sqrt{n}} \sum_{t=2}^{n} [(Y_t - \overline{Y})(Y_{t-1} - \overline{Y}) - \alpha(Y_{t-1} - \overline{Y})^2] - \frac{1}{\sqrt{n}} \sum_{t=2}^{n} [(Y_t - \mu_Y)(Y_{t-1} - \mu_Y) - \alpha(Y_{t-1} - \mu_Y)^2]$$

$$= \frac{1}{\sqrt{n}}\sum_{t=2}^{n}[-\{(Y_t-\mu_Y)+(Y_{t-1}-\mu_Y)\}(\overline{Y}-\mu_Y)+(\overline{Y}-\mu_Y)^2-\alpha\{-2(Y_{t-1}-\mu_Y)(\overline{Y}-\mu_Y)+(\overline{Y}-\mu_Y)^2\}]$$

$$= O_p(1/\sqrt{n}) \quad \text{(by Lemma 3.1(i), together with } \textstyle\sum_{t=2}^{n}(Y_{t-j}-\mu_Y)=n(\overline{Y}-\mu_Y)+O_p(1) \text{ for } j=0,1)$$

$$= o_p(1),$$

the result is shown by Slutsky's theorem and (i). □

We are ready to prove Proposition 3.1.

Proof of Proposition 3.1**.** The strong consistency is shown by Lemma 3.1(ii) and $n^{-1}(Y_\# - \overline{Y})^2 \xrightarrow{a.s.} 0$ for # = 0, 1. Also, it is easy to show that

$$\sqrt{n}(\widehat{\alpha}_{c_1,c_2}-\alpha) = \frac{\dfrac{1}{\sqrt{n}}\displaystyle\sum_{t=2}^{n}\{(Y_t-\overline{Y})(Y_{t-1}-\overline{Y})-\alpha(Y_{t-1}-\overline{Y})^2\}-\dfrac{\alpha}{\sqrt{n}}\{(c_1-1)(Y_1-\overline{Y})^2+c_2(Y_n-\overline{Y})^2\}}{\dfrac{1}{n}\Big\{\displaystyle\sum_{t=2}^{n}(Y_{t-1}-\overline{Y})^2+(c_1-1)(Y_1-\overline{Y})^2+c_2(Y_n-\overline{Y})^2\Big\}}$$

$$\xrightarrow{d} \mathrm{N}(0, v/\sigma_Y^4) \quad \text{(use Lemmas 3.1(ii) and 3.2(ii) and Slutsky's theorem).} \quad \square$$

3.7 Proofs of Propositions 3.2–3.5

Let $\overline{X}=(1/n)\sum_{t=1}^{n}X_t$, where $X_t=(Y_t-\mu_Y)/\sigma_Y$. One can rewrite $U_1=(n\sigma_Y^2)^{-1}\sum_{t=2}^{n}(Y_t-\overline{Y})(Y_{t-1}-\overline{Y})$ and $U_2^{c_1,c_2}=(n\sigma_Y^2)^{-1}[c_1(Y_1-\overline{Y})^2+\sum_{t=2}^{n-1}(Y_t-\overline{Y})^2+c_2(Y_n-\overline{Y})^2]$ as

$$U_1=\frac{1}{n}\sum_{t=2}^{n}X_tX_{t-1}-\Big(1+\frac{1}{n}\Big)\overline{X}^2+\frac{1}{n}(X_1\overline{X}+X_n\overline{X}),$$

$$U_2^{c_1,c_2}=\frac{1}{n}\Big[c_1X_1^2+\sum_{t=2}^{n-1}X_t^2+c_2X_n^2+2(1-c_1)X_1\overline{X}+2(1-c_2)X_n\overline{X}+(c_1+c_2-n-2)\overline{X}^2\Big]$$

$$=\frac{1}{n}\sum_{t=2}^{n}X_{t-1}^2+\Big(-1+\frac{c_1+c_2-2}{n}\Big)\overline{X}^2+\frac{1}{n}[(c_1-1)X_1^2+c_2X_n^2-2(c_1-1)X_1\overline{X}-2(c_2-1)X_n\overline{X}],$$

respectively, hence,

$$U_1-\alpha U_2^{c_1,c_2}=\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1}-\Big[1-\alpha+\frac{1}{n}\{1+(c_1+c_2-2)\alpha\}\Big]\overline{X}^2$$
$$+\frac{1}{n}[\alpha\{-(c_1-1)X_1^2-c_2X_n^2\}+\{1+2(c_1-1)\alpha\}X_1\overline{X}+\{1+2(c_2-1)\alpha\}X_n\overline{X}].$$

Before proving Propositions 3.2–3.5, we prepare the stochastic expansions of $\widehat{\alpha}_{c_1,c_2}$ and $\widetilde{\alpha}_{c_1,c_2}$. Using

$$E(X_tX_{t-1})=\alpha, E(X_t^2)=1 \quad \text{(note that } E[(X_t-\alpha X_{t-1})X_{t-1}]=0),$$

$$E(X_1\overline{X})=\frac{1}{n}\sum_{t=1}^{n}E(X_1X_t)=\frac{1}{n}\sum_{t=1}^{n}\alpha^{t-1}=\frac{1-\alpha^n}{n(1-\alpha)},$$

$$E(X_n\overline{X}) = \frac{1}{n}\sum_{t=1}^{n} E(X_tX_n) = \frac{1}{n}\sum_{t=1}^{n}\alpha^{n-t} = \frac{1-\alpha^n}{n(1-\alpha)},$$
$$E(\overline{X}^2) = \frac{1}{n^2}\sum_{t=1}^{n} E(X_t^2) + \frac{2}{n^2}\sum_{s,t=1}^{n} I_{\{s<t\}}E(X_sX_t) = \frac{1}{n} + \frac{2}{n^2}\sum_{s,t=1}^{n} I_{\{s<t\}}\alpha^{t-s} = \frac{1+\alpha}{n(1-\alpha)} - \frac{2\alpha(1-\alpha^n)}{n^2(1-\alpha)^2},$$

where

$$\sum_{s,t=1}^{n} I_{\{s<t\}}\alpha^{t-s} = \frac{\alpha}{1-\alpha}\sum_{s=1}^{n-1}(1-\alpha^{n-s}) = \frac{\alpha}{1-\alpha}\Big(n - \frac{1-\alpha^n}{1-\alpha}\Big),$$

we have

$$\begin{aligned}
E(U_2^{c_1,c_2}) &= \frac{1}{n}(n+c_1+c_2-2) - E(\overline{X}^2) + O(n^{-2}) \\
&= 1 + \frac{1}{n}\Big(-\frac{1+\alpha}{1-\alpha} + c_1 + c_2 - 2\Big) + O(n^{-2}) \\
&= 1 + \frac{1}{n}\delta^{c_1,c_2} + O(n^{-2}) \quad \text{(say)}, \\
E(U_1 - \alpha U_2^{c_1,c_2}) &= -(1-\alpha)E(\overline{X}^2) + O(n^{-2}) \\
&\quad + \frac{1}{n}[-(c_1+c_2-1)\alpha + \{1+2(c_1-1)\alpha\}E(X_1\overline{X}) + \{1+2(c_2-1)\alpha\}E(X_n\overline{X})] \\
&= -\frac{1}{n}\{1+(c_1+c_2)\alpha\} + O(n^{-2}) \\
&= \frac{1}{n}\Delta^{c_1,c_2} + O(n^{-2}) \quad \text{(say)}.
\end{aligned}$$

On the other hand, for $\# = 1, n$,

$$\begin{aligned}
V\Big(\frac{1}{n}X_\#\overline{X}\Big) &= \frac{1}{n^2}Cum(\overline{X},\overline{X},X_\#,X_\#) + \frac{1}{n^2}V(X_\#)V(\overline{X}) + \frac{1}{n^2}[Cov(X_\#,\overline{X})]^2 \\
&= \frac{1}{n^4}\sum_{t=1}^{n} Cum(X_t,X_t,X_\#,X_\#) + \frac{2}{n^4}\sum_{s,t=1}^{n} I_{\{s<t\}}Cum(X_s,X_t,X_\#,X_\#) \\
&\quad + \frac{1}{n^2}V(\overline{X}) + \frac{1}{n^4}\Big[\sum_{t=1}^{n} Cov(X_\#,X_t)\Big]^2 \\
&= O(n^{-3}),
\end{aligned}$$

since there exists a constant $C > 0$ (independent of $t_1, t_2 \in \{1, \ldots, n\}$) such that

$$|Cum(X_{t_1}, X_{t_2}, X_1, X_1)| \le C\alpha^{\max(t_1,t_2)-1},$$
$$|Cum(X_{t_1}, X_{t_2}, X_n, X_n)| \le C\alpha^{n-\min(t_1,t_2)}.$$

Also,

$$V(\overline{X}^2) = Cum(\overline{X},\overline{X},\overline{X},\overline{X}) + 2[V(\overline{X})]^2$$

$$
\begin{aligned}
&= \frac{1}{n^4}\sum_{t=1}^{n} Cum(X_t, X_t, X_t, X_t) + \frac{4}{n^4}\sum_{s,t=1}^{n} I_{\{s<t\}}[Cum(X_s, X_s, X_s, X_t) + Cum(X_s, X_t, X_t, X_t)] \\
&\quad + \frac{6}{n^4}\sum_{s,t=1}^{n} I_{\{s<t\}} Cum(X_s, X_s, X_t, X_t) \\
&\quad + \frac{12}{n^4}\sum_{s,t,u=1}^{n} I_{\{s<t<u\}}[Cum(X_s, X_t, X_u, X_u) + Cum(X_s, X_t, X_t, X_u) + Cum(X_s, X_s, X_t, X_u)] \\
&\quad + \frac{24}{n^4}\sum_{s,t,u,v=1}^{n} I_{\{s<t<u<v\}} Cum(X_s, X_t, X_u, X_v) \\
&\quad + 2[V(\overline{X})]^2 \\
&= \frac{2(1+\alpha)^2}{n^2(1-\alpha)^2} + O(n^{-3}),
\end{aligned}
$$

since there exists a constant $C > 0$ (independent of $t_1, t_2, t_3, t_4 \in \{1, \ldots, n\}$) such that

$$
|Cum(X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4})| \leq C\alpha^{\max(t_1,t_2,t_3,t_4) - \min(t_1,t_2,t_3,t_4)}.
$$

It is shown that

$$
\begin{aligned}
\frac{1}{n}X_\#\overline{X} &= \frac{1}{n}E(X_\#\overline{X}) + O_p(n^{-3/2}) = O_p(n^{-3/2}), \quad \# = 1, n, \\
\frac{1}{n}\overline{X}^2 &= \frac{1}{n}[E(\overline{X}^2) + O_p(n^{-1})] = O_p(n^{-2}).
\end{aligned}
$$

Then,

$$
\begin{aligned}
&U_2^{c_1,c_2} - E(U_2^{c_1,c_2}) \\
&= \frac{1}{n}\sum_{t=2}^{n}(X_{t-1}^2 - 1) + \Big[-\{\overline{X}^2 - E(\overline{X}^2)\} + \frac{1}{n}\{(c_1 - 1)(X_1^2 - 1) + c_2(X_n^2 - 1)\}\Big] + o_p(n^{-1}) \\
&= D_{-1/2} + D_{-1}^{c_1,c_2} + o_p(n^{-1})
\end{aligned}
$$

and

$$
\begin{aligned}
&U_1 - \alpha U_2^{c_1,c_2} - E(U_1 - \alpha U_2^{c_1,c_2}) \\
&= \frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1} - \Big[(1-\alpha)\{\overline{X}^2 - E(\overline{X}^2)\} + \frac{\alpha}{n}\{(c_1 - 1)(X_1^2 - 1) + c_2(X_n^2 - 1)\}\Big] \\
&\quad + \frac{1}{n}\Big[\{1 + 2(c_1 - 1)\alpha\}\{X_1\overline{X} - E(X_1\overline{X})\} + \{1 + 2(c_2 - 1)\alpha\}\{X_n\overline{X} - E(X_n\overline{X})\}\Big] + o_p(n^{-3/2}) \\
&= \Psi_{-1/2} + \Psi_{-1}^{c_1,c_2} + \Psi_{-3/2}^{c_1,c_2} + o_p(n^{-3/2}),
\end{aligned}
$$

where

$$
\begin{aligned}
D_{-1/2} &= \frac{1}{n}\sum_{t=2}^{n}(X_{t-1}^2 - 1),\\
D_{-1}^{c_1,c_2} &= -\{\overline{X}^2 - E(\overline{X}^2)\} + \frac{1}{n}\{(c_1-1)(X_1^2-1) + c_2(X_n^2-1)\},\\
\Psi_{-1/2} &= \frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1},\\
\Psi_{-1}^{c_1,c_2} &= -(1-\alpha)\{\overline{X}^2 - E(\overline{X}^2)\} - \frac{\alpha}{n}\{(c_1-1)(X_1^2-1) + c_2(X_n^2-1)\},\\
\Psi_{-3/2}^{c_1,c_2} &= \frac{1}{n}\Big[\{1+2(c_1-1)\alpha\}\{X_1\overline{X} - E(X_1\overline{X})\} + \{1+2(c_2-1)\alpha\}\{X_n\overline{X} - E(X_n\overline{X})\}\Big].
\end{aligned}
$$

In this way, we obtain the stochastic expansions

$$
\begin{aligned}
\widehat{\alpha}_{c_1,c_2} - \alpha &= \frac{E(U_1 - \alpha U_2^{c_1,c_2}) + \{U_1 - \alpha U_2^{c_1,c_2} - E(U_1 - \alpha U_2^{c_1,c_2})\}}{E(U_2^{c_1,c_2}) + \{U_2^{c_1,c_2} - E(U_2^{c_1,c_2})\}}\\
&= \frac{\Psi_{-1/2} + (n^{-1}\Delta^{c_1,c_2} + \Psi_{-1}^{c_1,c_2}) + \Psi_{-3/2}^{c_1,c_2} + o_p(n^{-3/2})}{1 + D_{-1/2} + (n^{-1}\delta^{c_1,c_2} + D_{-1}^{c_1,c_2}) + o_p(n^{-1})}\\
&= \Psi_{-1/2} + (n^{-1}\Delta^{c_1,c_2} + \Psi_{-1}^{c_1,c_2}) + \Psi_{-3/2}^{c_1,c_2} - \{\Psi_{-1/2} + (n^{-1}\Delta^{c_1,c_2} + \Psi_{-1}^{c_1,c_2})\}D_{-1/2}\\
&\quad - \Psi_{-1/2}(n^{-1}\delta^{c_1,c_2} + D_{-1}^{c_1,c_2}) + \Psi_{-1/2}D_{-1/2}^2 + o_p(n^{-3/2})\\
&= \Psi_{-1/2} + (n^{-1}\Delta^{c_1,c_2} + S_{-1}^{c_1,c_2}) + S_{-3/2}^{c_1,c_2} + o_p(n^{-3/2}),\\
\widetilde{\alpha}_{c_1,c_2} - \alpha &= \Psi_{-1/2} + (n^{-1}\dot{\Delta}^{c_1,c_2} + S_{-1}^{c_1,c_2}) + \dot{S}_{-3/2}^{c_1,c_2} + o_p(n^{-3/2}),
\end{aligned}
$$

where

$$
\begin{aligned}
S_{-1}^{c_1,c_2} &= \Psi_{-1}^{c_1,c_2} - \Psi_{-1/2}D_{-1/2},\\
S_{-3/2}^{c_1,c_2} &= \Psi_{-3/2}^{c_1,c_2} - (n^{-1}\Delta^{c_1,c_2} + \Psi_{-1}^{c_1,c_2})D_{-1/2} - \Psi_{-1/2}(n^{-1}\delta^{c_1,c_2} + D_{-1}^{c_1,c_2}) + \Psi_{-1/2}D_{-1/2}^2,\\
\dot{\Delta}^{c_1,c_2} &= \Delta^{c_1,c_2} + 1 + (2 + c_1 + c_2)\alpha + \frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2} = \frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2} + 2\alpha,\\
\dot{S}_{-3/2}^{c_1,c_2} &= S_{-3/2}^{c_1,c_2} + \frac{1}{n}\Big\{2 + c_1 + c_2 + \frac{(4\alpha + 2\alpha^2)Q_{2:3,Y}}{(1+\alpha)^2\sigma_Y^4} + \frac{1}{\sigma_Y^2}\Big\}\Psi_{-1/2}\\
&\quad + \frac{2\alpha^2}{n(1+\alpha)\sigma_Y^2}\Big[\frac{\sigma_Y}{n}\sum_{t=2}^{n}\{X_{t-1}^3 - E(X_{t-1}^3)\} - D_{-1/2}\Big] - \frac{1}{n}\Big\{\frac{4\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2}\Big\}D_{-1/2}.
\end{aligned}
$$

Proof of Propositions 3.2 **and** 3.3**(i).** It is easy to see that

$$
\begin{aligned}
&Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1}, \frac{1}{n}\sum_{t=2}^{n}X_{t-1}^2\Big]\\
&= \frac{1}{n^2}\sum_{s,t=2}^{n} I_{\{s<t\}} Cov[(X_s - \alpha X_{s-1})X_{s-1}, X_{t-1}^2]\\
&= \frac{1}{n^2}\sum_{s,t=2}^{n} I_{\{s<t\}}\Big[2\alpha^{2(t-s)}\frac{Q_{2:3,Y}}{\sigma_Y^4} + \frac{\alpha^{t-s}}{\sigma_Y^2} + 2\alpha^{2(t-s)-1} - \alpha\Big\{2\alpha^{2(t-s)}\frac{Q_{2:3,Y}}{\sigma_Y^4} + \frac{\alpha^{t-s}}{\sigma_Y^2} + 2\alpha^{2(t-s)}\Big\}\Big]
\end{aligned}
$$

$$= \frac{1}{n}\Big\{2\alpha + \frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2}\Big\} + O(n^{-2}).$$

Then, we have

$$\begin{aligned}
E(\widehat{\alpha}_{c_1,c_2} - \alpha) &= n^{-1}\Delta^{c_1,c_2} + E(S_{-1}^{c_1,c_2}) + o(n^{-1}) \\
&= -\frac{1}{n}\{1 + (c_1+c_2)\alpha\} - Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2\Big] + o(n^{-1}) \\
&= -\frac{1}{n}\Big\{1 + (2 + c_1 + c_2)\alpha + \frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2}\Big\} + o(n^{-1}), \\
E(\widetilde{\alpha}_{c_1,c_2} - \alpha) &= n^{-1}\dot{\Delta}^{c_1,c_2} + E(S_{-1}^{c_1,c_2}) + o(n^{-1}) \\
&= \frac{1}{n}\Big\{\frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2} + 2\alpha\Big\} - Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2\Big] + o(n^{-1}) \\
&= o(n^{-1}). \quad \square
\end{aligned}$$

Proof of Propositions 3.3**(ii),** 3.4**, and** 3.5**.** We notice that

$$\begin{aligned}
V(\widehat{\alpha}_{c_1,c_2}) &= V(\Psi_{-1/2}) + 2Cov(\Psi_{-1/2}, S_{-1}^{c_1,c_2}) + V(S_{-1}^{c_1,c_2}) + 2Cov(\Psi_{-1/2}, S_{-3/2}^{c_1,c_2}) + o(n^{-2}), \\
V(\widetilde{\alpha}_{c_1,c_2}) &= V(\Psi_{-1/2}) + 2Cov(\Psi_{-1/2}, S_{-1}^{c_1,c_2}) + V(S_{-1}^{c_1,c_2}) + 2Cov(\Psi_{-1/2}, \dot{S}_{-3/2}^{c_1,c_2}) + o(n^{-2}).
\end{aligned}$$

Now, it is not difficult to see that, for $\# = 1, n$,

$$\begin{aligned}
&Cum\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1}, \frac{1}{n}\sum_{t=2}^{n}(X_t - 3X_{t-1})X_{t-1}, \frac{1}{n}X_\#^2\Big] \\
&= \frac{1}{n^3}\sum_{s,t=2}^{n} I_{\{s\le t\}} Cum\{X_{s-1}(X_s - \alpha X_{s-1}), X_{t-1}(X_t - 3X_{t-1}), X_\#^2\} \\
&\quad + \frac{1}{n^3}\sum_{s,t=2}^{n} I_{\{s<t\}} Cum\{X_{s-1}(X_s - 3X_{s-1}), X_{t-1}(X_t - \alpha X_{t-1}), X_\#^2\} \\
&= O(n^{-3}),
\end{aligned}$$

since there exists a constant $C > 0$ (independent of $t_1, t_2, t_3, t_4 \in \{1, \ldots, n\}$) such that

$$\begin{aligned}
|Cum(X_{t_1}X_{t_2}, X_{t_3}X_{t_4}, X_1^2)| &\le C\alpha^{\max(t_1,t_2,t_3,t_4)-1}, \\
|Cum(X_{t_1}X_{t_2}, X_{t_3}X_{t_4}, X_n^2)| &\le C\alpha^{n-\min(t_1,t_2,t_3,t_4)}.
\end{aligned}$$

Similarly,

$$Cov(X_1^2, X_n^2) = O(\alpha^{n-1}),$$

and, for # = 0, 1,

$$
\begin{aligned}
Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{1}{n}X_{\#}\overline{X}\Big] &= \frac{1}{n^3}\sum_{s=1}^{n}\sum_{t=2}^{n}I_{\{s\le t\}}Cov\{(X_t-\alpha X_{t-1})X_{t-1},X_sX_{\#}\}\\
&\quad+\frac{1}{n^3}\sum_{s,t=2}^{n}I_{\{s<t\}}Cov\{(X_s-\alpha X_{s-1})X_{s-1},X_tX_{\#}\}\\
&= O(n^{-3}),\\
Cov\Big(\overline{X}^2,\frac{1}{n}X_{\#}^2\Big) &= \frac{1}{n^3}\sum_{s,t=1}^{n}I_{\{s\le t\}}Cov(X_tX_s,X_{\#}^2)+\frac{1}{n^3}\sum_{s,t=1}^{n}I_{\{s<t\}}Cum(X_sX_t,X_{\#}^2)\\
&= O(n^{-3}),
\end{aligned}
$$

since there exists a constant $C>0$ (independent of $t_1,t_2,t_3\in\{1,\ldots,n\}$) such that

$$
\begin{aligned}
|Cov(X_{t_1}X_{t_2},X_{t_3}X_1)| &\le C\alpha^{\max(t_1,t_2,t_3)-1},\\
|Cov(X_{t_1}X_{t_2},X_{t_3}X_n)| &\le C\alpha^{n-\min(t_1,t_2,t_3)}.
\end{aligned}
$$

Furthermore, we have

$$
\begin{aligned}
V\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1}\Big] &= \frac{1}{n^2}\sum_{t=2}^{n}V[(X_t-\alpha X_{t-1})X_{t-1}] = \Big(\frac{1}{n}-\frac{1}{n^2}\Big)\frac{v}{\sigma_Y^4},\\
V(X_{\#}^2) &= \frac{Q_{2:4,Y}}{\sigma_Y^4}+\frac{3Q_{2:3,Y}}{\sigma_Y^4}+\frac{1}{\sigma_Y^2}+2,\quad \# = 1,n,\\
Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{1}{n}X_1^2\Big] &= 0,\\
Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{1}{n}X_n^2\Big] &= \frac{1}{n^2}\Big\{\frac{2\alpha^2Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+2\alpha\Big\}+O(n^{-3}).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&V(\widehat{\alpha}_{c_1,c_2})-V(\widehat{\alpha}_{1,0})\\
&= 2Cov(\Psi_{-1/2},\Psi_{-1}^{c_1,c_2}-\Psi_{-1}^{1,0})+V(\Psi_{-1}^{c_1,c_2})-V(\Psi_{-1}^{1,0})\\
&\quad-2Cov(\Psi_{-1/2}D_{-1/2},\Psi_{-1}^{c_1,c_2}-\Psi_{-1}^{1,0})+2Cov(\Psi_{-1/2},\Psi_{-3/2}^{c_1,c_2}-\Psi_{-3/2}^{1,0})\\
&\quad-\frac{2}{n}(\Delta^{c_1,c_2}-\Delta^{1,0})Cov(\Psi_{-1/2},D_{-1/2})-2Cov\{\Psi_{-1/2},(\Psi_{-1}^{c_1,c_2}-\Psi_{-1}^{1,0})D_{-1/2}\}\\
&\quad-\frac{2}{n}(\delta^{c_1,c_2}-\delta^{1,0})V(\Psi_{-1/2})-2Cov\{\Psi_{-1/2},\Psi_{-1/2}(D_{-1}^{c_1,c_2}-D_{-1}^{1,0})\}+o(n^{-2})\\
&= 2Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},-\frac{\alpha}{n}\{(c_1-1)X_1^2+c_2X_n^2\}\Big]\\
&\quad+V\Big[\frac{\alpha}{n}\{(c_1-1)X_1^2+c_2X_n^2\}\Big]+2Cov\Big[(1-\alpha)\overline{X}^2,\frac{\alpha}{n}\{(c_1-1)X_1^2+c_2X_n^2\}\Big]\\
&\quad+2Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{2\alpha}{n}\{(c_1-1)X_1\overline{X}+c_2X_n\overline{X}\}\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{2}{n}(c_1+c_2-1)\alpha Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{1}{n}\sum_{t=2}^{n}X_{t-1}^2\Big]\\
&-\frac{2}{n}(c_1+c_2-1)V\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1}\Big]\\
&-2Cum\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{1}{n}\sum_{t=2}^{n}(X_t-3\alpha X_{t-1})X_{t-1},\frac{1}{n}\{(c_1-1)X_1^2+c_2X_n^2\}\Big]+o(n^{-2})\\
=&-\frac{2c_2\alpha}{n^2}\Big\{\frac{2\alpha^2Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+2\alpha\Big\}+\{(c_1-1)^2+c_2^2\}\frac{\alpha^2}{n^2}\Big(\frac{Q_{2:4,Y}}{\sigma_Y^4}+\frac{3Q_{2:3,Y}}{\sigma_Y^4}+\frac{1}{\sigma_Y^2}+2\Big)\\
&+\frac{2}{n^2}(c_1+c_2-1)\alpha\Big\{\frac{2\alpha^2Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+2\alpha\Big\}-\frac{2}{n^2}(c_1+c_2-1)(1-\alpha)\Big(\frac{\alpha Q_{2:3,Y}}{\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+1+\alpha\Big)+o(n^{-2})\\
=&\frac{1}{n^2}\Big[\{(c_1-1)^2+c_2^2\}\alpha^2\frac{Q_{2:4,Y}}{\sigma_Y^4}\\
&\quad+[2(1-c_1-c_2)\alpha+3\{(c_1-1)^2+c_2^2\}\alpha^2+\{3(c_1-1)^2+6(c_1-1)+3c_2^2+2c_2\}\alpha^3]\frac{Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}\\
&\quad+[2(1-c_1-c_2)\alpha+\{(c_1-1)^2+4(c_1-1)+c_2^2+2c_2\}\alpha^2]\frac{1}{\sigma_Y^2}\\
&\quad+2(1-c_1-c_2)+2\{(c_1-1)^2+3(c_1-1)+c_2^2+c_2\}\alpha^2\Big]\\
&+o(n^{-2}).
\end{aligned}
$$

Further,

$$
[E(\widehat{\alpha}_{c_1,c_2}-\alpha)]^2-[E(\widehat{\alpha}_{1,0}-\alpha)]^2=\frac{1}{n^2}(c_1+c_2-1)\alpha\Big\{2+(5+c_1+c_2)\alpha+\frac{4\alpha^2Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{2\alpha}{\sigma_Y^2}\Big\}+o(n^{-2}).
$$

Similarly, we have

$$
\begin{aligned}
&V(\widetilde{\alpha}_{c_1,c_2})-V(\widetilde{\alpha}_{0,0})\\
=&2Cov(\Psi_{-1/2},\Psi_{-1}^{c_1,c_2}-\Psi_{-1}^{0,0})+V(\Psi_{-1}^{c_1,c_2})-V(\Psi_{-1}^{0,0})-2Cov(\Psi_{-1/2}D_{-1/2},\Psi_{-1}^{c_1,c_2}-\Psi_{-1}^{0,0})\\
&+2Cov(\Psi_{-1/2},\Psi_{-3/2}^{c_1,c_2}-\Psi_{-3/2}^{0,0})+\frac{2}{n}(c_1+c_2)V(\Psi_{-1/2})\\
&-\frac{2}{n}(\Delta^{c_1,c_2}-\Delta^{0,0})Cov(\Psi_{-1/2},D_{-1/2})-2Cov\{\Psi_{-1/2},(\Psi_{-1}^{c_1,c_2}-\Psi_{-1}^{0,0})D_{-1/2}\}\\
&-\frac{2}{n}(\delta^{c_1,c_2}-\delta^{0,0})V(\Psi_{-1/2})-2Cov\{\Psi_{-1/2},\Psi_{-1/2}(D_{-1}^{c_1,c_2}-D_{-1}^{0,0})\}+o(n^{-2})\\
=&2Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},-\frac{\alpha}{n}(c_1X_1^2+c_2X_n^2)\Big]\\
&+V\Big[\frac{\alpha}{n}\{(c_1-1)X_1^2+c_2X_n^2\}\Big]-V\Big(\frac{\alpha}{n}X_1^2\Big)+2Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{2\alpha}{n}(c_1X_1\overline{X}+c_2X_n\overline{X})\Big]\\
&+\frac{2}{n}(c_1+c_2)\alpha Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{1}{n}\sum_{t=2}^{n}X_{t-1}^2\Big]\\
&-2Cum\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t-\alpha X_{t-1})X_{t-1},\frac{1}{n}\sum_{t=2}^{n}(X_t-3\alpha X_{t-1})X_{t-1},\frac{1}{n}(c_1X_1^2+c_2X_n^2)\Big]+o(n^{-2})
\end{aligned}
$$

$$= -\frac{2c_2\alpha}{n^2}\Big\{\frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2} + 2\alpha\Big\} + \{(c_1-1)^2 + c_2^2 - 1\}\frac{\alpha^2}{n^2}\Big(\frac{Q_{2:4,Y}}{\sigma_Y^4} + \frac{3Q_{2:3,Y}}{\sigma_Y^4} + \frac{1}{\sigma_Y^2} + 2\Big)$$

$$+ \frac{2}{n^2}(c_1+c_2)\alpha\Big\{\frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha}{\sigma_Y^2} + 2\alpha\Big\} + o(n^{-2})$$

$$= \frac{\alpha^2}{n^2}\Big[(c_1^2 - 2c_1 + c_2^2)\frac{Q_{2:4,Y}}{\sigma_Y^4} + \{3(c_1^2 - 2c_1 + c_2^2) + (3c_1^2 - 2c_1 + 3c_2^2)\alpha\}\frac{Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + (c_1^2 + c_2^2)\Big(\frac{1}{\sigma_Y^2} + 2\Big)\Big]$$

$$+ o(n^{-2}).$$

On the other hand, we additionally obtain

$$Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^3\Big]$$

$$= \frac{1}{n^2}\sum_{s,t=2}^{n} I_{\{s<t\}} Cov[(X_s - \alpha X_{s-1})X_{s-1}, X_{t-1}^3]$$

$$= \frac{1}{n^2}\sum_{s,t=2}^{n} I_{\{s<t\}}\Big[Cum(X_{s-1}, X_s - \alpha X_{s-1}, X_{t-1}, X_{t-1}, X_{t-1})$$

$$+ 3Cum(X_{s-1}, X_s - \alpha X_{s-1}, X_{t-1})V(X_{t-1})$$

$$+ 3Cum(X_s - \alpha X_{s-1}, X_{t-1}, X_{t-1})Cov(X_{t-1}, X_{s-1})$$

$$+ 3Cum(X_{s-1}, X_{t-1}, X_{t-1})Cov(X_s - \alpha X_{s-1}, X_{t-1})\Big]$$

$$= \frac{1}{n^2\sigma_Y}\sum_{s,t=2}^{n} I_{\{s<t\}}\Big[(1-\alpha)\Big\{3\alpha^{3(t-s)}\frac{Q_{2:4,Y}}{\sigma_Y^4} + 6\alpha^{2(t-s)}\frac{Q_{2:3,Y}}{\sigma_Y^4} + \frac{\alpha^{t-s}}{\sigma_Y^2}\Big\} + 3(1-\alpha)\alpha^{t-s}$$

$$+ 3\Big\{(1-\alpha^3)\alpha^{2(t-s-1)}\frac{Q_{2:3,Y}}{\sigma_Y^4} + (1-\alpha^2)\frac{\alpha^{t-s-1}}{\sigma_Y^2}\Big\}\alpha^{t-s}$$

$$+ 3\Big\{\alpha^{2(t-s)}\frac{Q_{2:3,Y}}{\sigma_Y^4} + \frac{\alpha^{t-s}}{\sigma_Y^2}\Big\}(1-\alpha^2)\alpha^{t-s-1}\Big]$$

$$= \frac{3}{n\sigma_Y}\Big\{\frac{\alpha^3 Q_{2:4,Y}}{(1+\alpha+\alpha^2)\sigma_Y^4} + \alpha^2\Big(\frac{2}{1+\alpha} + \frac{1+\alpha}{1+\alpha+\alpha^2} + 1\Big)\frac{Q_{2:3,Y}}{\sigma_Y^4} + \frac{2\alpha}{\sigma_Y^2} + \alpha\Big\} + O(n^{-2}),$$

where

$$\sum_{s,t=2}^{n} I_{\{s<t\}}\alpha^{j(t-s)} = \frac{\alpha^j}{1-\alpha^j}\sum_{s=2}^{n-1}(1-\alpha^{j(n-s)}) = \frac{\alpha^j}{1-\alpha^j}\Big(n - 1 - \frac{1-\alpha^{j(n-1)}}{1-\alpha^j}\Big), j = 1, 2, 3.$$

It follows that

$$V(\widehat{\alpha}_{c_1,c_2}) - V(\widetilde{\alpha}_{c_1,c_2})$$

$$= 2Cov(\Psi_{-1/2}, S_{-3/2}^{c_1,c_2} - \dot{S}_{-3/2}^{c_1,c_2}) + o(n^{-2})$$

$$= -\frac{2}{n}\Big\{\frac{(4\alpha + 2\alpha^2)Q_{2:3,Y}}{(1+\alpha)^2\sigma_Y^4} + \frac{1}{\sigma_Y^2} + 2 + c\Big\}V(\Psi_{-1/2})$$

$$+ \frac{2}{n}\Big\{\frac{4\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{\alpha + 3\alpha^2}{(1+\alpha)\sigma_Y^2}\Big\}Cov(\Psi_{-1/2}, D_{-1/2}) - \frac{4\alpha^2}{n(1+\alpha)\sigma_Y}Cov\Big[\Psi_{-1/2}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^3\Big] + o(n^{-2})$$

$$
\begin{aligned}
&= \frac{2}{n^2}\Big[c(\alpha-1)\Big(\frac{\alpha Q_{2:3,Y}}{\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+1+\alpha\Big)-\Big\{\frac{(4\alpha+2\alpha^2)Q_{2:3,Y}}{(1+\alpha)^2\sigma_Y^4}+\frac{1}{\sigma_Y^2}+2\Big\}(1-\alpha)\Big(\frac{\alpha Q_{2:3,Y}}{\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+1+\alpha\Big)\\
&\quad+\Big\{\frac{4\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha+3\alpha^2}{(1+\alpha)\sigma_Y^2}\Big\}\Big\{\frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}+2\alpha\Big\}\\
&\quad-\frac{6\alpha^2}{(1+\alpha)\sigma_Y^2}\Big\{\frac{\alpha^3 Q_{2:4,Y}}{(1+\alpha+\alpha^2)\sigma_Y^4}+\alpha^2\Big(\frac{2}{1+\alpha}+\frac{1+\alpha}{1+\alpha+\alpha^2}+1\Big)\frac{Q_{2:3,Y}}{\sigma_Y^4}+\frac{2\alpha}{\sigma_Y^2}+\alpha\Big\}\Big]+o(n^{-2}).
\end{aligned}
$$

Further,

$$
[E(\widehat{\alpha}_{c_1,c_2}-\alpha)]^2-[E(\widetilde{\alpha}_{c_1,c_2}-\alpha)]^2=\frac{1}{n^2}\Big\{1+(2+c)\alpha+\frac{2\alpha^2 Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4}+\frac{\alpha}{\sigma_Y^2}\Big\}^2+o(n^{-2}). \quad \square
$$

Chapter 4

Whittle estimation in INAR(1) process and test of equidispersion

4.1 Introduction

Recall that a nonnegative integer-valued autoregressive process of the first-order (INAR(1)) is defined by $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$. Al-Osh and Alzaid (1987) and Park and Oh (1997) studied the Yule–Walker (YW), conditional least squares (CLS), and (conditional or full) maximum likelihood (ML) estimators for the parameter α in the INAR(1) process with the Poisson marginals. Freeland and McCabe (2005) pointed out that Park and Oh's asymptotic variance of the CLS estimator is incorrect and showed that the CLS estimator is asymptotically equivalent to the YW estimator. Weiß (2012) also proposed a squared difference estimator for the innovation mean.

Without assuming the Poisson innovation, Drost et al. (2009) developed a semi-parametrically efficient estimation in the stationary INAR(p) process. Zeng and Kakizawa (2022) recently considered a class of estimators for the parameter α in the stationary INAR(1) process under a general innovation, which includes the YW, Burg, and method of moment estimators as special cases. See Chapter 3.

There are some reasons to apply the Whittle likelihood method. Nowadays, the frequency domain analysis is a standard tool in the stationary processes (see, e.g., Brockwell and Davis (1987)). Some simulation results about the Whittle estimation in the INAR(1) process with the Poisson marginals were given by da Silva and Oliveira (2004) (see also Zhang and Wang (2015) for random coefficient INAR(1) process). However, to the best of our knowledge, there is no theoretical results about the frequency domain analysis for the stationary INAR(1) process. A contribution of this chapter is to study the large sample theory of the Whittle estimation in the stationary INAR(1) process under a general innovation.

Although the Poisson distribution is fundamental for the count data analysis and yields a mathematical elegance due to its equidispersion property (i.e., the mean is equal to the variance), the mean and variance of other distributions may be not the same, so that, in practice, the overdispersed (or underdispersed) case is more important. Schweer and Weiß (2014,2016) studied some tests for the INAR(1) process with the Poisson marginals. Another contribution of this chapter is to propose the Wald-type test about the equidispersion, which is a competitor of the test considered by Schweer and Weiß (2014).

The rest of this chapter is organized as follows. In Section 4.2, after a brief introduction of the stationary INAR(1) process, the Whittle likelihood is given and the strong consistency and asymptotic normality of the Whittle estimator for the parameter α and the innovation mean and variance (denoted by μ_ε and σ_ε^2) in the stationary INAR(1) process. The Wald-type tests about the equidispersion are also considered in Section 4.3, on the basis of the estimators for μ_ε and σ_ε^2. Section 4.4 assesses the finite-sample performances of (i) the Whittle, YW, and CLS estimators and (ii) the empirical sizes (i.e., type I errors) and powers of the proposed equidispersion tests, through the simulations. Section 4.5 concludes this chapter. The proofs of Propositions 4.1 and 4.2 are postponed to Section 4.6.

4.2 Whittle estimator for the parameter α and the innovation mean and variance

We consider the INAR(1) process (Al-Osh and Alzaid (1987)), defined by

$$Y_t = \alpha \circ Y_{t-1} + \varepsilon_t, \quad t = 0, \pm 1, \dots, \tag{4.1}$$

where $\alpha\circ$ is the binomial thinning operator with $\alpha \in [0, 1)$, and $\{\varepsilon_t\}$, referred to as an innovation, is a sequence of independent and identically distributed nonnegative integer-valued random variables, such that ε_t and Y_{t-i} are independent for all $i \geq 1$. Let $\kappa_{j,\varepsilon} = Cum_j(\varepsilon_t)$ be the jth cumulant of ε_t, where $j = 3, 4, \dots$.

We emphasize that, throughout this thesis, except for some simulation experiments, we do not assume any distributional form about the innovation $\{\varepsilon_t\}$. In this sense, we treat (4.1) to be semi-parametric.

When $\alpha \in [0, 1)$, the process (4.1) is strictly stationary and ergodic (Du and Li (1991)), whose mean μ_Y, variance σ_Y^2, and autocovariance function at lag $u \geq 0$ are given by

$$\mu_Y = \frac{\mu_\varepsilon}{1-\alpha}, \quad \sigma_Y^2 = \frac{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}{1-\alpha^2}, \quad \text{and} \quad Cov(Y_t, Y_{t+u}) = \alpha^u \sigma_Y^2 = \gamma_Y(u) \quad \text{(say)}, \tag{4.2}$$

respectively (see, e.g., Al-Osh and Alzaid (1987)). We derived, in Chapter 2, explicit expressions of the third and fourth cumulant of Y_t;

$$\kappa_{3,Y} = \frac{1}{1-\alpha^3}\{\kappa_{3,\varepsilon} + 3\alpha^2(1-\alpha)\sigma_Y^2 + \alpha(1-\alpha)(1-2\alpha)\mu_Y\},$$
$$\kappa_{4,Y} = \frac{1}{1-\alpha^4}\{\kappa_{4,\varepsilon} + 6\alpha^3(1-\alpha)\kappa_{3,Y} + \alpha^2(1-\alpha)(7-11\alpha)\sigma_Y^2 + \alpha(1-\alpha)(1-6\alpha+6\alpha^2)\mu_Y\},$$

respectively. Let

$$Q_{2:3,Y} = \kappa_{3,Y} - \sigma_Y^2 \text{ and } Q_{2:4,Y} = \kappa_{4,Y} - 3\kappa_{3,Y} + 2\sigma_Y^2.$$

Noting that all cumulants of a Poisson distribution Po(λ) are equal to λ, we see that $Q_{2:3,Y} = Q_{2:4,Y} = 0$ for the INAR(1) process $\{Y_t\}$ with the Poisson marginals (it is a particular case where the innovation $\{\varepsilon_t\}$ follows the Poisson distribution Po($(1-\alpha)\mu$) or an alternative parameterization Po(η), where $\mu(>0)$ and $\eta(>0)$ are independent of α).

Using (4.2), the spectral density of the stationary INAR(1) process $\{Y_t\}$ is given by

$$f_Y(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma_Y(|h|)e^{-ih\lambda} = \frac{\sigma_Y^2}{2\pi}\Big[1 + \sum_{h=1}^{\infty}\alpha^h(e^{-i\lambda h} + e^{i\lambda h})\Big] = \frac{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}{2\pi[1-\alpha(e^{-i\lambda}+e^{i\lambda})+\alpha^2]},$$

as in the usual stationary autoregressive process of the first-order (see, e.g., Brockwell and Davis (1987)).

From now on, suppose that the observation $\{Y_1, \ldots, Y_n\}$ of length n is generated by (4.1). On the basis of the frequency domain approach, we consider simultaneous estimation of $(\alpha, \mu_\varepsilon, \sigma_\varepsilon^2)^T = \theta$ (say). Let $\overline{Y} = (1/n)\sum_{t=1}^n Y_t$. We define the periodogram for the demeaned data by

$$\widetilde{I}_Y(\lambda) = \frac{1}{2\pi n}\Big|\sum_{t=1}^{n}(Y_t - \overline{Y})e^{-it\lambda}\Big|^2 = \frac{1}{2\pi}\Big[\widehat{\gamma}(0) + \sum_{h=1}^{n-1}\widehat{\gamma}(h)(e^{-ih\lambda} + e^{ih\lambda})\Big],$$

where $\widehat{\gamma}(h) = (1/n)\sum_{t=1}^{n-h}(Y_t - \overline{Y})(Y_{t+h} - \overline{Y})$, $h = 0, 1, \ldots, n-1$ (note that $\widehat{\sigma}_Y^2 = \widehat{\gamma}(0)$).

It is easy to see that

$$\begin{aligned}\int_{-\pi}^{\pi}\frac{\widetilde{I}_Y(\lambda)}{f_Y(\lambda)}d\lambda &= \int_{-\pi}^{\pi}\Big[\widehat{\gamma}(0) + \sum_{h=1}^{n-1}\widehat{\gamma}(h)(e^{-ih\lambda} + e^{ih\lambda})\Big]\frac{1-\alpha(e^{-i\lambda}+e^{i\lambda})+\alpha^2}{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}d\lambda \\ &= \frac{2\pi[(1+\alpha^2)\widehat{\gamma}(0) - 2\alpha\widehat{\gamma}(1)]}{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}.\end{aligned}$$

Also,

$$\begin{aligned}\int_{-\pi}^{\pi}\log f_Y(\lambda)d\lambda &= \int_{-\pi}^{\pi}[\log(\alpha\mu_\varepsilon + \sigma_\varepsilon^2) - \log 2\pi - \log\{1-\alpha(e^{-i\lambda}+e^{i\lambda})+\alpha^2\}]d\lambda \\ &= 2\pi[\log(\alpha\mu_\varepsilon + \sigma_\varepsilon^2) - \log 2\pi].\end{aligned}$$

The Whittle (W) likelihood for the demeaned data is defined by

$$\begin{aligned}\widetilde{L}_W(\theta) &= -\frac{1}{4\pi}\int_{-\pi}^{\pi}\Big[\log f_Y(\lambda) + \frac{\widetilde{I}_Y(\lambda)}{f_Y(\lambda)}\Big]d\lambda \\ &= -\frac{1}{2}\Big[\log(\alpha\mu_\varepsilon + \sigma_\varepsilon^2) - \log 2\pi + \frac{(1+\alpha^2)\widehat{\gamma}(0) - 2\alpha\widehat{\gamma}(1)}{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}\Big] \qquad (4.3) \\ &\le -\frac{1}{2}[\log\{(1+\alpha^2)\widehat{\gamma}(0) - 2\alpha\widehat{\gamma}(1)\} - \log 2\pi + 1]\end{aligned}$$

(the equality holds iff $\sigma_\varepsilon^2 + \alpha\mu_\varepsilon = (1+\alpha^2)\widehat{\gamma}(0) - 2\alpha\widehat{\gamma}(1)$), so that the Whittle estimator for the parameter α, which corresponds to the minimizer of $(1+\alpha^2)\widehat{\gamma}(0) - 2\alpha\widehat{\gamma}(1)$, is given by $\widehat{\gamma}(1)/\widehat{\gamma}(0) = \widehat{\alpha}_{YW}$ (the YW estimator). Although the criterion (4.3) is not identifiable with respect to $(\mu_\varepsilon, \sigma_\varepsilon^2)$, the formula (4.2) enables us to see that the innovation mean $\mu_\varepsilon = (1-\alpha)\mu_Y$ can be naturally estimated as $\widehat{\mu}_\varepsilon = (1-\widehat{\alpha}_{YW})\overline{Y}$, and then the estimator for the innovation variance σ_ε^2 is given by

$$\widehat{\sigma_\varepsilon^2} = (1+\widehat{\alpha}_{YW}^2)\widehat{\gamma}(0) - 2\widehat{\alpha}_{YW}\widehat{\gamma}(1) - \widehat{\alpha}_{YW}\widehat{\mu}_\varepsilon = (1-\widehat{\alpha}_{YW}^2)\widehat{\gamma}(0) - \widehat{\alpha}_{YW}\widehat{\mu}_\varepsilon.$$

Due to the non-identifiability of (4.3) using the demeaned periodogram, we consider

$$I_Y^\bullet(\lambda) = \frac{1}{2\pi n}\Big|\sum_{t=1}^n (Y_t - \mu_Y)e^{-it\lambda}\Big|^2 = \frac{1}{2\pi}\Big[\widehat{\gamma}^\bullet(0) + \sum_{h=1}^{n-1}\widehat{\gamma}^\bullet(h)(e^{-ih\lambda} + e^{ih\lambda})\Big],$$

where $\widehat{\gamma}^\bullet(h) = (1/n)\sum_{t=1}^{n-h}(Y_t - \mu_Y)(Y_{t+h} - \mu_Y)$, $h = 0, 1, \ldots, n-1$, and

$$\int_{-\pi}^{\pi} \frac{I_Y^\bullet(\lambda)}{f_Y(\lambda)} d\lambda = \frac{2\pi}{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}[(1+\alpha^2)\widehat{\gamma}^\bullet(0) - 2\alpha\widehat{\gamma}^\bullet(1)].$$

We emphasize that the Whittle likelihood, defined by

$$\begin{aligned} L_W^\bullet(\theta) &= -\frac{1}{4\pi}\int_{-\pi}^{\pi}\Big[\log f_Y(\lambda) + \frac{I_Y^\bullet(\lambda)}{f_Y(\lambda)}\Big]d\lambda \\ &= -\frac{1}{2}\Big[\log(\alpha\mu_\varepsilon + \sigma_\varepsilon^2) - \log 2\pi + \frac{(1+\alpha^2)\widehat{\gamma}^\bullet(0) - 2\alpha\widehat{\gamma}^\bullet(1)}{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}\Big] \\ &\leq -\frac{1}{2}[\log\{(1+\alpha^2)\widehat{\gamma}^\bullet(0) - 2\alpha\widehat{\gamma}^\bullet(1)\} - \log 2\pi + 1] \end{aligned} \tag{4.4}$$

(the equality holds iff $\sigma_\varepsilon^2 = (1+\alpha^2)\widehat{\gamma}^\bullet(0) - 2\alpha\widehat{\gamma}^\bullet(1) - \alpha\mu_\varepsilon$), turns out to be identifiable, unlike (4.3). That is, the Whittle estimator for α and μ_ε, denoted by $\widehat{\alpha}_W$ and $\widehat{\mu}_{\varepsilon;W}$, is defined as the minimizer of $(1+\alpha^2)\widehat{\gamma}^\bullet(0) - 2\alpha\widehat{\gamma}^\bullet(1) = \mathcal{J}_Y(\alpha, \mu_\varepsilon)$ (say), and the Whittle estimator for σ_ε^2 is given by

$$\begin{aligned} \widehat{\sigma_{\varepsilon;W}^2} &= (1+\widehat{\alpha}_W^2)\Big[\widehat{\gamma}(0) + \Big(\overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)^2\Big] \\ &\quad - 2\widehat{\alpha}_W\Big[\widehat{\gamma}(1) + \frac{n-1}{n}\Big(\overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)^2 + \frac{1}{n}\Big(\overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)\{(\overline{Y} - Y_n) + (\overline{Y} - Y_1)\}\Big] - \widehat{\alpha}_W\widehat{\mu}_{\varepsilon;W} \\ &= (1+\widehat{\alpha}_W^2)\widehat{\gamma}(0) - 2\widehat{\alpha}_W\widehat{\gamma}(1) - \widehat{\alpha}_W\widehat{\mu}_{\varepsilon;W} + \Big\{(1-\widehat{\alpha}_W)^2 + \frac{2\widehat{\alpha}_W}{n}\Big\}\Big(\overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)^2 \\ &\quad - \frac{2\widehat{\alpha}_W}{n}\Big(\overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)\{(\overline{Y} - Y_n) + (\overline{Y} - Y_1)\}. \end{aligned}$$

To establish the strong consistency and asymptotic normality of $\widehat{\theta}_W = (\widehat{\alpha}_W, \widehat{\mu}_{\varepsilon;W}, \widehat{\sigma_{\varepsilon;W}^2})^T$, we make the following assumptions:

(A1) $\{Y_t\}$ is INAR(1) process with $\alpha \in [0, 1)$.

(A2) $E(Y_t^J)$ exists for some integer $J \geq 4$.

(A3) The true parameter $\theta_0' = (\alpha_0, \mu_{\varepsilon,0})^T$ is an interior point of the parameter space Θ'. Also, $\sigma^2_{\varepsilon,0}(>0)$ is the true innovation variance.

(A4) $\Theta' = [\alpha_L, \alpha_U] \times [\mu_L, \mu_U]$, where $0 \le \alpha_L < \alpha_U < 1$ and $0 \le \mu_L < \mu_U < \infty$.

The first assumption is our statistical setting, without any distributional form of the innovation $\{\varepsilon_t\}$, which ensures that the INAR(1) process $\{Y_t\}$ is strictly stationary and ergodic (Du and Li (1991)). The second assumption is implied by the existence of $E(\varepsilon_t^J)$. The third and fourth assumptions are standard for the asymptotic theory.

In what follows, the true parameter of θ is denoted by $\theta_0 = (\alpha_0, \mu_{\varepsilon,0}, \sigma^2_{\varepsilon,0})^T$. Let

$$\mu_{Y,0} = \frac{\mu_{\varepsilon,0}}{1-\alpha_0}, \quad \sigma^2_{Y,0} = \frac{\alpha_0\mu_{\varepsilon,0} + \sigma^2_{\varepsilon,0}}{1-\alpha_0^2}(>0),$$
$$Q_{2:3,Y,0} = \kappa_{3,Y,0} - \sigma^2_{Y,0}, \text{ and } Q_{2:4,Y,0} = \kappa_{4,Y,0} - 3\kappa_{3,Y,0} + 2\sigma^2_{Y,0},$$

where

$$\kappa_{3,Y,0} = \frac{1}{1-\alpha_0^3}\{\kappa_{3,\varepsilon,0} + 3\alpha_0^2(1-\alpha_0)\sigma^2_{Y,0} + \alpha_0(1-\alpha_0)(1-2\alpha_0)\mu_{Y,0}\},$$
$$\kappa_{4,Y,0} = \frac{1}{1-\alpha_0^4}\{\kappa_{4,\varepsilon,0} + 6\alpha_0^3(1-\alpha_0)\kappa_{3,Y,0} + \alpha_0^2(1-\alpha_0)(7-11\alpha_0)\sigma^2_{Y,0}$$
$$+ \alpha_0(1-\alpha_0)(1-6\alpha_0+6\alpha_0^2)\mu_{Y,0}\}$$

($\kappa_{3,\varepsilon,0}$ and $\kappa_{4,\varepsilon,0}$ are the true third and fourth cumulants of ε_t, respectively).

We state the strong consistency and asymptotic normality of the Whittle estimator $\widehat{\theta}_W$.

Proposition 4.1. *Suppose that (A1)–(A4) hold. The following hold.*

(i) $\widehat{\theta}_W \xrightarrow{a.s.} \theta_0$.

(ii) $\sqrt{n}(\widehat{\theta}_W - \theta_0) \xrightarrow{d} \mathrm{N}(0, V_0)$, *where*

$$V_0 = (1-\alpha_0)\begin{pmatrix} \omega_{\alpha\alpha,0} & \omega_{\alpha\mu_\varepsilon,0} & \omega_{\alpha\sigma^2_\varepsilon,0} \\ \omega_{\mu_\varepsilon\alpha,0} & \omega_{\mu_\varepsilon\mu_\varepsilon,0} & \omega_{\mu_\varepsilon\sigma^2_\varepsilon,0} \\ \omega_{\sigma^2_\varepsilon\alpha,0} & \omega_{\sigma^2_\varepsilon\mu_\varepsilon,0} & \omega_{\sigma^2_\varepsilon\sigma^2_\varepsilon,0} \end{pmatrix}$$

is 3×3 *symmetric matrix, with*

$$\omega_{\alpha\alpha,0} = \Big(\frac{\alpha_0 Q_{2:3,Y,0}}{\sigma^4_{Y,0}} + \frac{\alpha_0}{\sigma^2_{Y,0}} + 1 + \alpha_0\Big),$$
$$\omega_{\mu_\varepsilon\alpha,0} = \alpha_0 - \Big(\frac{\alpha_0 Q_{2:3,Y,0}}{\sigma^4_{Y,0}} + \frac{\alpha_0}{\sigma^2_{Y,0}} + 1 + \alpha_0\Big)\mu_{Y,0},$$
$$\omega_{\sigma^2_\varepsilon\alpha,0} = (1-2\alpha_0)\Big[\alpha_0 - \Big(\frac{\alpha_0 Q_{2:3,Y,0}}{\sigma^4_{Y,0}} + \frac{\alpha_0}{\sigma^2_{Y,0}} + 1 + \alpha_0\Big)\mu_{Y,0}\Big],$$

$$\omega_{\mu_\varepsilon\mu_\varepsilon,0} = \Big(\frac{\alpha_0 Q_{2:3,Y,0}}{\sigma_{Y,0}^4} + \frac{\alpha_0}{\sigma_{Y,0}^2} + 1 + \alpha_0\Big)\mu_{Y,0}^2 + (1+\alpha_0)\sigma_{Y,0}^2 - 2\alpha_0\mu_{Y,0},$$

$$\omega_{\sigma_\varepsilon^2\mu_\varepsilon,0} = (1+\alpha_0+\alpha_0^2)Q_{2:3,Y,0} + (1-2\alpha_0)\Big(\frac{\alpha_0 Q_{2:3,Y,0}}{\sigma_{Y,0}^4} + \frac{\alpha_0}{\sigma_{Y,0}^2} + 1 + \alpha_0\Big)\mu_{Y,0}^2$$
$$+ (1+\alpha_0-2\alpha_0^2)\sigma_{Y,0}^2 - 2\alpha_0(1-2\alpha_0)\mu_{Y,0},$$

$$\omega_{\sigma_\varepsilon^2\sigma_\varepsilon^2,0} = (1+\alpha_0)(1-\alpha_0^2)(Q_{2:4,Y,0} + 2\sigma_{Y,0}^4) + 3(1+\alpha_0+\alpha_0^2-\alpha_0^3)Q_{2:3,Y,0}$$
$$+ (1-2\alpha_0)^2\Big(\frac{\alpha_0 Q_{2:3,Y,0}}{\sigma_{Y,0}^4} + \frac{\alpha_0}{\sigma_{Y,0}^2} + 1 + \alpha_0\Big)\mu_{Y,0}^2$$
$$+ (1+\alpha_0-4\alpha_0^2+4\alpha_0^3)\sigma_{Y,0}^2 - 2\alpha_0(1-2\alpha_0)^2\mu_{Y,0}.$$

The CLS method is widely used in the INAR-type processes. Minimizing the criterion

$$J(\alpha,\mu_\varepsilon) = \sum_{t=2}^n (Y_t - E[Y_t|Y_{t-1}])^2 = \sum_{t=2}^n (Y_t - \alpha Y_{t-1} - \mu_\varepsilon)^2, \tag{4.5}$$

the CLS estimator for α and μ_ε is defined by

$$\widehat{\alpha}_{CLS} = \frac{\sum_{t=2}^n Y_t Y_{t-1} - \frac{1}{n-1}\sum_{t=2}^n Y_t \sum_{t=2}^n Y_{t-1}}{\sum_{t=2}^n Y_{t-1}^2 - \frac{1}{n-1}\Big(\sum_{t=2}^n Y_{t-1}\Big)^2} \quad \text{and} \quad \widehat{\mu}_{\varepsilon;CLS} = \frac{1}{n-1}\Big(\sum_{t=2}^n Y_t - \widehat{\alpha}_{CLS}\sum_{t=2}^n Y_{t-1}\Big)$$

(see also Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)). Using (4.2), the innovation variance $\sigma_\varepsilon^2 = (1-\alpha^2)\sigma_Y^2 - \alpha\mu_\varepsilon$ can be estimated as

$$\widehat{\sigma^2_{\varepsilon;CLS}} = (1-\widehat{\alpha}_{CLS}^2)\widehat{\gamma}(0) - \widehat{\alpha}_{CLS}\widehat{\mu}_{\varepsilon;CLS}.$$

On the other hand, Zeng and Kakizawa (2022) studied a class of estimators for the parameter α in the statioinary INAR(1) process under a general innovation, given by

$$\widehat{\alpha}_{(c_1,c_2)} = \frac{\sum_{t=2}^n (Y_t - \overline{Y})(Y_{t-1} - \overline{Y})}{c_1(Y_1-\overline{Y})^2 + \sum_{t=2}^{n-1}(Y_t-\overline{Y})^2 + c_2(Y_n-\overline{Y})^2}, \quad c_1, c_2 \geq 0$$

(the YW estimator $\widehat{\alpha}_{YW}$ is a special case of $c_1 = c_2 = 1$). Then, using (4.2), the innovation mean μ_ε and variance σ_ε^2 can be estimated as

$$\widehat{\mu}_{\varepsilon;(c_1,c_2)} = (1-\widehat{\alpha}_{(c_1,c_2)})\overline{Y} \ \text{ and } \ \widehat{\sigma^2_{\varepsilon;(c_1,c_2)}} = (1-\widehat{\alpha}_{(c_1,c_2)}^2)\widehat{\gamma}(0) - \widehat{\alpha}_{(c_1,c_2)}\widehat{\mu}_{\varepsilon;(c_1,c_2)},$$

respectively.

The strong consistency and asymptotic normality of the estimators $\widehat{\theta}_{CLS} = (\widehat{\alpha}_{CLS}, \widehat{\mu}_{\varepsilon;CLS}, \widehat{\sigma^2_{\varepsilon;CLS}})^T$ and $\widehat{\theta}_{(c_1,c_2)} = (\widehat{\alpha}_{(c_1,c_2)}, \widehat{\mu}_{\varepsilon;(c_1,c_2)}, \widehat{\sigma^2_{\varepsilon;(c_1,c_2)}})^T$ can be established similarly.

Proposition 4.2. *Suppose that (A1) and (A2) hold. The following hold.*

(i) $\widehat{\theta}_{CLS} \xrightarrow{a.s.} \theta_0$ *and* $\widehat{\theta}_{(c_1,c_2)} \xrightarrow{a.s.} \theta_0$.

(ii) $\sqrt{n}(\widehat{\theta}_{CLS} - \theta_0) \xrightarrow{d} \mathrm{N}(0, V_0)$ *and* $\sqrt{n}(\widehat{\theta}_{(c_1,c_2)} - \theta_0) \xrightarrow{d} \mathrm{N}(0, V_0)$, *i.e., the estimators* $\widehat{\theta}_{CLS}$ *and* $\widehat{\theta}_{(c_1,c_2)}$ *are asymptotically equivalent to the Whittle estimator* $\widehat{\theta}_W$.

As usual, we define $\widehat{\kappa}_{3,Y} = (1/n)\sum_{t=1}^{n}(Y_t - \overline{Y})^3$ and $\widehat{\kappa}_{4,Y} = (1/n)\sum_{t=1}^{n}(Y_t - \overline{Y})^4 - 3\widehat{\sigma}_Y^4$ (we set $\widehat{Q}_{2:3,Y} = \widehat{\kappa}_{3,Y} - \widehat{\sigma}_Y^2$ and $\widehat{Q}_{2:4,Y} = \widehat{\kappa}_{4,Y} - 3\widehat{\kappa}_{3,Y} + 2\widehat{\sigma}_Y^2$). With the replacement of $\alpha, \mu_Y, \sigma_Y^2$, and $Q_{2:j,Y}$ by $\widehat{\alpha}_W$ (or $\widehat{\alpha}_{CLS}$, $\widehat{\alpha}_{(c_1,c_2)}$), $\overline{Y}$, $\widehat{\sigma}_Y^2$, and $\widehat{Q}_{2:j,Y}$, where $j = 3, 4$, we naturally have a consistent estimator for V_0, denoted by $\widehat{V}_{0;W}$ (or $\widehat{V}_{0;CLS}$, $\widehat{V}_{0;(c_1,c_2)}$). The standard errors (SEs) of the estimators are given by, e.g., $\mathrm{SE} = \sqrt{(1 - \widehat{\alpha}_W)\widehat{\omega}_{\alpha\alpha,0;W}/n}$ for $\widehat{\alpha}_W$.

4.3 Test of equidispersion

It is easy to see that the dispersion index of the process $\{Y_t\}$ is given by

$$\frac{\sigma_{Y,0}^2}{\mu_{Y,0}} = \frac{\alpha_0\mu_{\varepsilon,0} + \sigma_{\varepsilon,0}^2}{(1+\alpha_0)\mu_{\varepsilon,0}} = 1 + \frac{1}{1+\alpha_0}\Big(\frac{\sigma_{\varepsilon,0}^2}{\mu_{\varepsilon,0}} - 1\Big),$$

where the left-hand-side is equal to 1 when $\sigma_{\varepsilon,0}^2 = \mu_{\varepsilon,0}$ (the equidispersion property of the innovation $\{\varepsilon_t\}$). That is, the INAR(1) process (4.1) is shown to be overdispersed (underdispersed) when $\sigma_{\varepsilon,0}^2 > \mu_{\varepsilon,0}$ ($\sigma_{\varepsilon,0}^2 < \mu_{\varepsilon,0}$), so that the testing problem about the equidispersion of the stationary INAR(1) process can be formulated, as follows.

- Two-sided testing problem is given by

$$\text{(I) } \mathrm{H}_0\text{: } \sigma_{\varepsilon,0}^2 - \mu_{\varepsilon,0} = 0 \text{ against } \mathrm{H}_1\text{: } \sigma_{\varepsilon,0}^2 - \mu_{\varepsilon,0} \neq 0.$$

- One-sided testing problem for the overdispersed (or underdispersed) alternative is given by

$$\text{(II) } \mathrm{H}_0\text{: } \sigma_{\varepsilon,0}^2 - \mu_{\varepsilon,0} = 0 \text{ against } \mathrm{H}_1\text{: } \sigma_{\varepsilon,0}^2 - \mu_{\varepsilon,0} > 0,$$

$$\text{(III) } \mathrm{H}_0\text{: } \sigma_{\varepsilon,0}^2 - \mu_{\varepsilon,0} = 0 \text{ against } \mathrm{H}_1\text{: } \sigma_{\varepsilon,0}^2 - \mu_{\varepsilon,0} < 0.$$

On the basis of the asymptotic normality of the estimators for μ_ε and σ_ε^2, we can construct the Wald-type test for the linear constraint H_0.

By Propositions 4.1 and 4.2, we state the following corollaries (without proof).

Corollary 4.1. *Suppose that (A1)–(A4) hold. Then,*

$$\frac{\sqrt{n}}{\sqrt{v_0}}[(\widehat{\sigma^2_{\varepsilon;W}} - \widehat{\mu}_{\varepsilon;W}) - (\sigma^2_{\varepsilon,0} - \mu_{\varepsilon,0})] \xrightarrow{d} \mathrm{N}(0,1),$$

where

$$\begin{aligned} v_0 &= (1-\alpha_0)(\omega_{\sigma^2_\varepsilon\sigma^2_\varepsilon,0} - 2\omega_{\sigma^2_\varepsilon\mu_\varepsilon,0} + \omega_{\mu_\varepsilon\mu_\varepsilon,0}) \\ &= (1-\alpha_0)\Big[(1+\alpha_0)(1-\alpha_0^2)(Q_{2:4,Y,0} + 2\sigma^4_{Y,0}) + (1+\alpha_0+\alpha_0^2-3\alpha_0^3)Q_{2:3,Y,0} \\ &\quad + 4\alpha_0^2\Big\{\Big(\frac{\alpha_0 Q_{2:3,Y,0}}{\sigma^4_{Y,0}} + \frac{\alpha_0}{\sigma^2_{Y,0}} + 1 + \alpha_0\Big)\mu^2_{Y,0} + \alpha_0\sigma^2_{Y,0} - 2\alpha_0\mu_{Y,0}\Big\}\Big]. \end{aligned}$$

Under (A1) and (A2), the same results hold for $\widehat{\sigma^2_{\varepsilon;CLS}} - \widehat{\mu}_{\varepsilon;CLS}$ *and* $\widehat{\sigma^2_{\varepsilon;(c_1,c_2)}} - \widehat{\mu}_{\varepsilon;(c_1,c_2)}$.

With the replacement of $\alpha, \mu_Y, \sigma^2_Y$, and $Q_{2:j,Y}$ by $\widehat{\alpha}_W$ (or $\widehat{\alpha}_{CLS}$, $\widehat{\alpha}_{(c_1,c_2)}$), $\overline{Y}$, $\widehat{\sigma}^2_Y$, and $\widehat{Q}_{2:j,Y}$, where $j = 3, 4$, we naturally have a consistent estimator for v_0, denoted by $\widehat{v}_{0;W}$ (or $\widehat{v}_{0;CLS}$, $\widehat{v}_{0;(c_1,c_2)}$). By Corollary 4.1 and Slutsky's theorem, we immediately have the following corollary, which is the key result for the above-mentioned test of the equidispersion.

Corollary 4.2. *Suppose that (A1)–(A4) hold. Then,*

$$\frac{\sqrt{n}}{\sqrt{\widehat{v}_{0;W}}}[(\widehat{\sigma^2_{\varepsilon;W}} - \widehat{\mu}_{\varepsilon;W}) - (\sigma^2_{\varepsilon,0} - \mu_{\varepsilon,0})] \xrightarrow{d} \mathrm{N}(0,1).$$

Under (A1) and (A2), the same results hold for $\widehat{\sigma^2_{\varepsilon;CLS}} - \widehat{\mu}_{\varepsilon;CLS}$ *and* $\widehat{\sigma^2_{\varepsilon;(c_1,c_2)}} - \widehat{\mu}_{\varepsilon;(c_1,c_2)}$, *if* $\widehat{v}_{0;W}$ *is replaced by* $\widehat{v}_{0;CLS}$ *and* $\widehat{v}_{0;(c_1,c_2)}$, *respectively.*

The Wald-type statistics for testing the hypothesis $\mathrm{H}_0 : \sigma^2_{\varepsilon,0} - \mu_{\varepsilon,0} = 0$ are defined by

$$\begin{aligned} Z_W &= \frac{\sqrt{n}(\widehat{\sigma^2_{\varepsilon;W}} - \widehat{\mu}_{\varepsilon;W})}{\sqrt{\widehat{v}_{0;W}}}, \\ Z_{CLS} &= \frac{\sqrt{n}(\widehat{\sigma^2_{\varepsilon;CLS}} - \widehat{\mu}_{\varepsilon;CLS})}{\sqrt{\widehat{v}_{0;CLS}}}, \\ Z_{(c_1,c_2)} &= \frac{\sqrt{n}(\widehat{\sigma^2_{\varepsilon;(c_1,c_2)}} - \widehat{\mu}_{\varepsilon;(c_1,c_2)})}{\sqrt{\widehat{v}_{0;(c_1,c_2)}}} \end{aligned}$$

(especially, let $Z_{YW} = Z_{(1,0)}$). For a given significance level a, the notation $z(a)$ stands for the upper a percentile of the standard normal distribution N(0,1). After computing the z-value z_W (say), the null hypothesis H_0 is rejected if $|z_W| > z(a/2)$ for the two-sided problem (I). Similarly, reject H_0 if $z_W > z(a)$ for the one-sided problem (II) (reject H_0 if $z_W < -z(a)$ for the one-sided problem (III)).

Remark 4.1. For the INAR(1) process with the Poisson marginals, Schweer and Weiß (2014) proved that

$$\left\{\frac{n(1-\alpha^2)}{2(1+\alpha^2)}\right\}^{1/2}\left(\frac{\widehat{\sigma}_Y^2}{\overline{Y}}-1\right) \xrightarrow{d} \mathrm{N}(0,1).$$

Thus, other statistics

$$Z_{SW,\#} = \left\{\frac{n(1-\widehat{\alpha}_\#^2)}{2(1+\widehat{\alpha}_\#^2)}\right\}^{1/2}\left(\frac{\widehat{\sigma}_Y^2}{\overline{Y}}-1\right), \quad \# = W, CLS, YW, \tag{4.6}$$

are available to test the equidispersion hypothesis $\sigma_{\mathrm{Y}}^2/\mu_{\mathrm{Y}} = 1$.

4.4 Simulation results

Performances of the estimators for α, μ_ε, and σ_ε^2

We first report the simulation results ($n = 100, 200, 300$, with 2000 replications) about the Whittle (W), YW, and CLS estimators for α, μ_ε, and σ_ε^2, under the set-up that the innovation $\{\varepsilon_t\}$ follows the Poisson distribution $\mathrm{Po}((1-\alpha)\mu)$ or negative binomial (NB) distribution $\mathrm{NB}(r, r/\{r+(1-\alpha)\mu\})$, where the second parameter $r/\{r+(1-\alpha)\mu\}$ is the success probability in each trial (we set $r = 10$). In the simulations, we fix $(1-\alpha)\mu = 5$. The parameter α to be estimated is $\alpha = 0.2, 0.5, 0.8$, and the innovation mean and variance to be estimated are given, as follows.

- For the Poisson case, $\mu_\varepsilon = \sigma_\varepsilon^2 = 5$.
- For the NB case, $\mu_\varepsilon = 5$ and $\sigma_\varepsilon^2 = 5(1+5/r)$.

In view of Tables 4.1 and 4.2, the biases, variances, and mean squared errors (MSEs) of the respective estimators decrease as the sample size n increases. The biases of $\widehat{\alpha}_{CLS}$ and $\widehat{\mu}_{\varepsilon;CLS}$ are overall smallest, whereas the bias of $\widehat{\sigma^2_{\varepsilon;W}}$ is smallest when $\alpha = 0.8$. On the other hand, if α is small, the variance and MSE of $\widehat{\theta}_W$ are smallest. In summary, $\widehat{\alpha}_W$ outperforms $\widehat{\alpha}_{YW}$ and $\widehat{\alpha}_{CLS}$ in terms of the variance and MSE when $\alpha = 0.2$ and $n = 100$, whereas the performances of $\widehat{\alpha}_W$, $\widehat{\alpha}_{YW}$, and $\widehat{\alpha}_{CLS}$ are almost the same when $n = 200, 300$ (this finding is in agreement with Proposition 4.2).

Performances of the equidispersion tests

We next report the empirical type I errors and powers about two-sided tests of the equidispersion having the rejection regions $|z_\#| > z(a/2)$ (thick lines) or $|z_{SW,\#}| > z(a/2)$ (thin lines) for $\# = W, YW, CLS$, with the significance level $a = 0.1, 0.05, 0.01$ (10000 replications).

For the type I error ($n = 100, 200, 300, 400$), we generate the innovation $\{\varepsilon_t\}$ according to the Poisson distribution $\mathrm{Po}((1-\alpha)\mu)$, under the set-up that $(1-\alpha)\mu = 5$ and $\alpha = 0.2, 0.5, 0.8$. Figure 4.1 indicates that, as the sample size n increases, the empirical type I errors of the tests proposed tend to the significance level (these tests are liberal), except that the Whittle-based test is oversized when $\alpha = 0.8$. The tests using (4.6) are conservative (liberal) when $\alpha = 0.2$ ($\alpha = 0.8$).

For the power (we consider $n = 300$ only), we generate the innovation $\{\varepsilon_t\}$ according to the NB distribution $\mathrm{NB}(r, r/\{r + (1-\alpha)\mu\}$, where $(1-\alpha)\mu = 5$, $\alpha = 0.2, 0.5, 0.8$ and $r = 2^j$, for $j = 1, \ldots, 7$. Figure 4.2 indicates that smaller r becomes (this case is far away from the null hypothesis of equidispersion, since $\sigma_\varepsilon^2 - \mu_\varepsilon = 25/r$), higher the power of the test is. The proposed tests are reasonable. The simulation reveals, however, that the tests using (4.6) are likely to be more powerful.

4.5 Concluding remarks

We have mainly considered asymptotic theory about the frequency domain analysis in the stationary INAR(1) process. We have pointed out that the Whittle likelihood from the demeaned data (see (4.3)) is not identifiable with respect to $(\mu_\varepsilon, \sigma_\varepsilon^2)$, whereas a variant of the Whittle likelihood (4.4) is identifiable. We have shown that the Whittle estimator $\widehat{\theta}_W$ maximizing the criterion (4.4) is strongly consistent and asymptotically normal, together with asymptotic equivalence among the estimators $\widehat{\theta}_W$, $\widehat{\theta}_{CLS}$, and $\widehat{\theta}_{(c_1,c_2)}$. Furthermore, we have proposed the Wald-type tests about the equidispersion, on the basis of the estimators $(\widehat{\mu}_{\varepsilon;W}, \widehat{\sigma^2_{\varepsilon;W}})^T$, $(\widehat{\mu}_{\varepsilon;CLS}, \widehat{\sigma^2_{\varepsilon;CLS}})^T$, and $(\widehat{\mu}_{\varepsilon;(c_1,c_2)}, \widehat{\sigma^2_{\varepsilon;(c_1,c_2)}})^T$.

Although the CLS method is fundamental and widely used in the time series analysis, the criterion (4.5) is free of the innovation variance σ_ε^2 so that additional estimation tool is needed (e.g., Karlsen and Tjøstheim (1988)). The YW estimator $\widehat{\alpha}_{YW}$ (more generally, a class of estimators $\widehat{\alpha}_{(c_1,c_2)}$) is a kind of the method of moment estimator but the construction of the estimator $(\widehat{\mu}_{\varepsilon;(c_1,c_2)}, \widehat{\sigma^2_{\varepsilon;(c_1,c_2)}})^T$, being the sample analogues of (4.2), is ad-hoc. In conclusion, the use of the identifiable Whittle likelihood (4.4) makes simultaneous inference of $(\alpha, \mu_\varepsilon, \sigma_\varepsilon^2)$ possible; besides, this approach is applicable even if any parametric assumption about the innovation $\{\varepsilon_t\}$ is not assumed, unlike the ML method.

4.6 Proofs of Propositions 4.1 and 4.2

Before proving Proposition 4.1, we prepare the following lemma.

Lemma 4.1. *Suppose that (A1)–(A4) hold. Then,*

$$\sup_{\theta' \in \Theta'} |\mathcal{J}_Y(\alpha, \mu_\varepsilon) - q(\alpha, \mu_\varepsilon)| \xrightarrow{a.s.} 0,$$

where

$$q(\alpha, \mu_\varepsilon) = (1 - 2\alpha\alpha_0 + \alpha^2)\sigma_{Y,0}^2 + (1-\alpha)^2(\mu_{Y,0} - \mu_Y)^2.$$

Table 4.1: Biases, variances, and MSEs (2000 replications) of $\widehat{\alpha}_{\#}$, $\widehat{\mu}_{\varepsilon;\#}$, and $\widehat{\sigma^2_{\varepsilon;\#}}$, $\# = W, YW, CLS$, in the stationary INAR(1) process under the Poisson innovation.

	n	Biases			Variances			MSEs		
		$\widehat{\alpha}_W$	$\widehat{\alpha}_{YW}$	$\widehat{\alpha}_{CLS}$	$\widehat{\alpha}_W$	$\widehat{\alpha}_{YW}$	$\widehat{\alpha}_{CLS}$	$\widehat{\alpha}_W$	$\widehat{\alpha}_{YW}$	$\widehat{\alpha}_{CLS}$
$\alpha = 0.2$	100	-0.0191	-0.0205	-0.0187	0.0089	0.0095	0.0097	0.0093	0.0099	0.0100
	200	-0.0095	-0.0096	-0.0088	0.0049	0.0050	0.0050	0.0050	0.0051	0.0051
	300	-0.0071	-0.0071	-0.0065	0.0032	0.0032	0.0032	0.0033	0.0033	0.0033
0.5	100	-0.0318	-0.0325	-0.0274	0.0079	0.0078	0.0080	0.0089	0.0089	0.0087
	200	-0.0149	-0.0151	-0.0127	0.0040	0.0040	0.0040	0.0042	0.0042	0.0042
	300	-0.0122	-0.0123	-0.0105	0.0027	0.0027	0.0027	0.0028	0.0028	0.0028
0.8	100	-0.0328	-0.0451	-0.0365	0.0061	0.0049	0.0050	0.0072	0.0070	0.0063
	200	-0.0165	-0.0203	-0.0162	0.0022	0.0020	0.0020	0.0024	0.0024	0.0022
	300	-0.0126	-0.0143	-0.0116	0.0015	0.0014	0.0014	0.0016	0.0016	0.0015
$(\mu_\varepsilon = 5)$	n	$\widehat{\mu}_{\varepsilon;W}$	$\widehat{\mu}_{\varepsilon;YW}$	$\widehat{\mu}_{\varepsilon;CLS}$	$\widehat{\mu}_{\varepsilon;W}$	$\widehat{\mu}_{\varepsilon;YW}$	$\widehat{\mu}_{\varepsilon;CLS}$	$\widehat{\mu}_{\varepsilon;W}$	$\widehat{\mu}_{\varepsilon;YW}$	$\widehat{\mu}_{\varepsilon;CLS}$
$\alpha = 0.2$	100	0.1167	0.1260	0.1157	0.3970	0.4177	0.4238	0.4107	0.4336	0.4371
	200	0.0587	0.0592	0.0537	0.2153	0.2166	0.2188	0.2188	0.2201	0.2217
	300	0.0417	0.0419	0.0381	0.1416	0.1417	0.1422	0.1433	0.1435	0.1436
0.5	100	0.3187	0.3237	0.2730	0.8084	0.8045	0.8177	0.9100	0.9093	0.8922
	200	0.1533	0.1552	0.1315	0.4092	0.4075	0.4094	0.4327	0.4316	0.4267
	300	0.1270	0.1283	0.1110	0.2719	0.2717	0.2732	0.2880	0.2881	0.2855
0.8	100	0.7934	1.1086	0.8959	3.8760	3.0621	3.0918	4.5055	4.2911	3.8944
	200	0.4019	0.5010	0.3981	1.3664	1.2299	1.2247	1.5279	1.4809	1.3832
	300	0.3129	0.3540	0.2869	0.9033	0.8651	0.8626	1.0012	0.9904	0.9449
$(\sigma^2_\varepsilon = 5)$	n	$\widehat{\sigma^2_{\varepsilon;W}}$	$\widehat{\sigma^2_{\varepsilon;YW}}$	$\widehat{\sigma^2_{\varepsilon;CLS}}$	$\widehat{\sigma^2_{\varepsilon;W}}$	$\widehat{\sigma^2_{\varepsilon;YW}}$	$\widehat{\sigma^2_{\varepsilon;CLS}}$	$\widehat{\sigma^2_{\varepsilon;W}}$	$\widehat{\sigma^2_{\varepsilon;YW}}$	$\widehat{\sigma^2_{\varepsilon;CLS}}$
$\alpha = 0.2$	100	0.0299	0.0384	0.0268	0.8506	0.8640	0.8683	0.8514	0.8655	0.8690
	200	0.0147	0.0153	0.0098	0.4429	0.4436	0.4441	0.4431	0.4439	0.4442
	300	0.0113	0.0114	0.0073	0.2914	0.2915	0.2914	0.2915	0.2916	0.2915
0.5	100	0.0039	0.0065	-0.0440	1.1600	1.1618	1.1523	1.1600	1.1618	1.1543
	200	-0.0049	-0.0045	-0.0287	0.5528	0.5532	0.5507	0.5528	0.5532	0.5515
	300	0.0217	0.0217	0.0040	0.3993	0.3994	0.3970	0.3997	0.3999	0.3970
0.8	100	-0.2711	-0.4098	-0.6075	3.1594	2.7902	2.5255	3.2329	2.9582	2.8946
	200	-0.1624	-0.2078	-0.3059	1.3262	1.2575	1.2005	1.3526	1.3007	1.2941
	300	-0.1199	-0.1373	-0.2036	0.9184	0.9020	0.8707	0.9328	0.9209	0.9122

Table 4.2: Biases, variances, and MSEs (2000 replications) of $\widehat{\alpha}_{\#}$, $\widehat{\mu}_{\varepsilon;\#}$, and $\widehat{\sigma^2_{\varepsilon;\#}}$, $\# = W, YW, CLS$, in the stationary INAR(1) process under the NB innovation.

		Biases			Variances			MSEs		
	n	$\widehat{\alpha}_W$	$\widehat{\alpha}_{YW}$	$\widehat{\alpha}_{CLS}$	$\widehat{\alpha}_W$	$\widehat{\alpha}_{YW}$	$\widehat{\alpha}_{CLS}$	$\widehat{\alpha}_W$	$\widehat{\alpha}_{YW}$	$\widehat{\alpha}_{CLS}$
$\alpha = 0.2$	100	-0.0197	-0.0216	-0.0199	0.0086	0.0094	0.0096	0.0090	0.0098	0.0099
	200	-0.0112	-0.0113	-0.0104	0.0045	0.0045	0.0046	0.0046	0.0046	0.0047
	300	-0.0062	-0.0062	-0.0055	0.0032	0.0032	0.0033	0.0033	0.0033	0.0033
0.5	100	-0.0287	-0.0308	-0.0260	0.0077	0.0074	0.0076	0.0085	0.0084	0.0083
	200	-0.0146	-0.0152	-0.0127	0.0039	0.0039	0.0039	0.0041	0.0041	0.0041
	300	-0.0095	-0.0098	-0.0081	0.0025	0.0025	0.0025	0.0026	0.0026	0.0026
0.8	100	-0.0356	-0.0440	-0.0352	0.0055	0.0047	0.0048	0.0067	0.0067	0.0060
	200	-0.0200	-0.0224	-0.0183	0.0022	0.0020	0.0020	0.0026	0.0025	0.0024
	300	-0.0126	-0.0137	-0.0111	0.0013	0.0013	0.0013	0.0015	0.0015	0.0014
$(\mu_\varepsilon = 5)$	n	$\widehat{\mu}_{\varepsilon;W}$	$\widehat{\mu}_{\varepsilon;YW}$	$\widehat{\mu}_{\varepsilon;CLS}$	$\widehat{\mu}_{\varepsilon;W}$	$\widehat{\mu}_{\varepsilon;YW}$	$\widehat{\mu}_{\varepsilon;CLS}$	$\widehat{\mu}_{\varepsilon;W}$	$\widehat{\mu}_{\varepsilon;YW}$	$\widehat{\mu}_{\varepsilon;CLS}$
$\alpha = 0.2$	100	0.1688	0.1723	0.1237	0.4137	0.4506	0.4578	0.4422	0.4803	0.4731
	200	0.0903	0.0866	0.0617	0.2107	0.2132	0.2154	0.2189	0.2207	0.2192
	300	0.0516	0.0487	0.0322	0.1492	0.1502	0.1509	0.1519	0.1525	0.1520
0.5	100	0.4335	0.3891	0.2657	0.8330	0.8523	0.8675	1.0210	1.0036	0.9381
	200	0.2232	0.1940	0.1313	0.4217	0.4287	0.4320	0.4715	0.4664	0.4492
	300	0.1411	0.1199	0.0790	0.2679	0.2715	0.2730	0.2878	0.2859	0.2792
0.8	100	0.8792	1.0908	0.8693	3.4608	3.0198	3.0392	4.2339	4.2095	3.7949
	200	0.4993	0.5624	0.4590	1.3533	1.2841	1.2897	1.6027	1.6005	1.5003
	300	0.3163	0.3458	0.2803	0.8380	0.8148	0.8110	0.9381	0.9344	0.8895
$(\sigma^2_\varepsilon = 7.5)$	n	$\widehat{\sigma^2_{\varepsilon;W}}$	$\widehat{\sigma^2_{\varepsilon;YW}}$	$\widehat{\sigma^2_{\varepsilon;CLS}}$	$\widehat{\sigma^2_{\varepsilon;W}}$	$\widehat{\sigma^2_{\varepsilon;YW}}$	$\widehat{\sigma^2_{\varepsilon;CLS}}$	$\widehat{\sigma^2_{\varepsilon;W}}$	$\widehat{\sigma^2_{\varepsilon;YW}}$	$\widehat{\sigma^2_{\varepsilon;CLS}}$
$\alpha = 0.2$	100	0.1157	0.1294	0.1234	2.0510	2.0647	2.0686	2.0643	2.0814	2.0838
	200	0.0780	0.0797	0.0764	0.9940	0.9946	0.9951	1.0001	1.0010	1.0010
	300	0.0440	0.0448	0.0423	0.6413	0.6414	0.6422	0.6433	0.6434	0.6440
0.5	100	0.4884	0.5285	0.4958	2.6552	2.6770	2.6656	2.8938	2.9564	2.9113
	200	0.2443	0.2637	0.2490	1.2515	1.2560	1.2549	1.3112	1.3255	1.3169
	300	0.1725	0.1852	0.1756	0.8432	0.8450	0.8457	0.8729	0.8793	0.8766
0.8	100	-0.4097	-0.4857	-0.7736	4.4315	4.2408	3.7012	4.5993	4.4767	4.2998
	200	-0.2158	-0.2412	-0.3813	2.0397	2.0043	1.8866	2.0863	2.0624	2.0320
	300	-0.1401	-0.1519	-0.2418	1.2644	1.2539	1.2075	1.2840	1.2769	1.2660

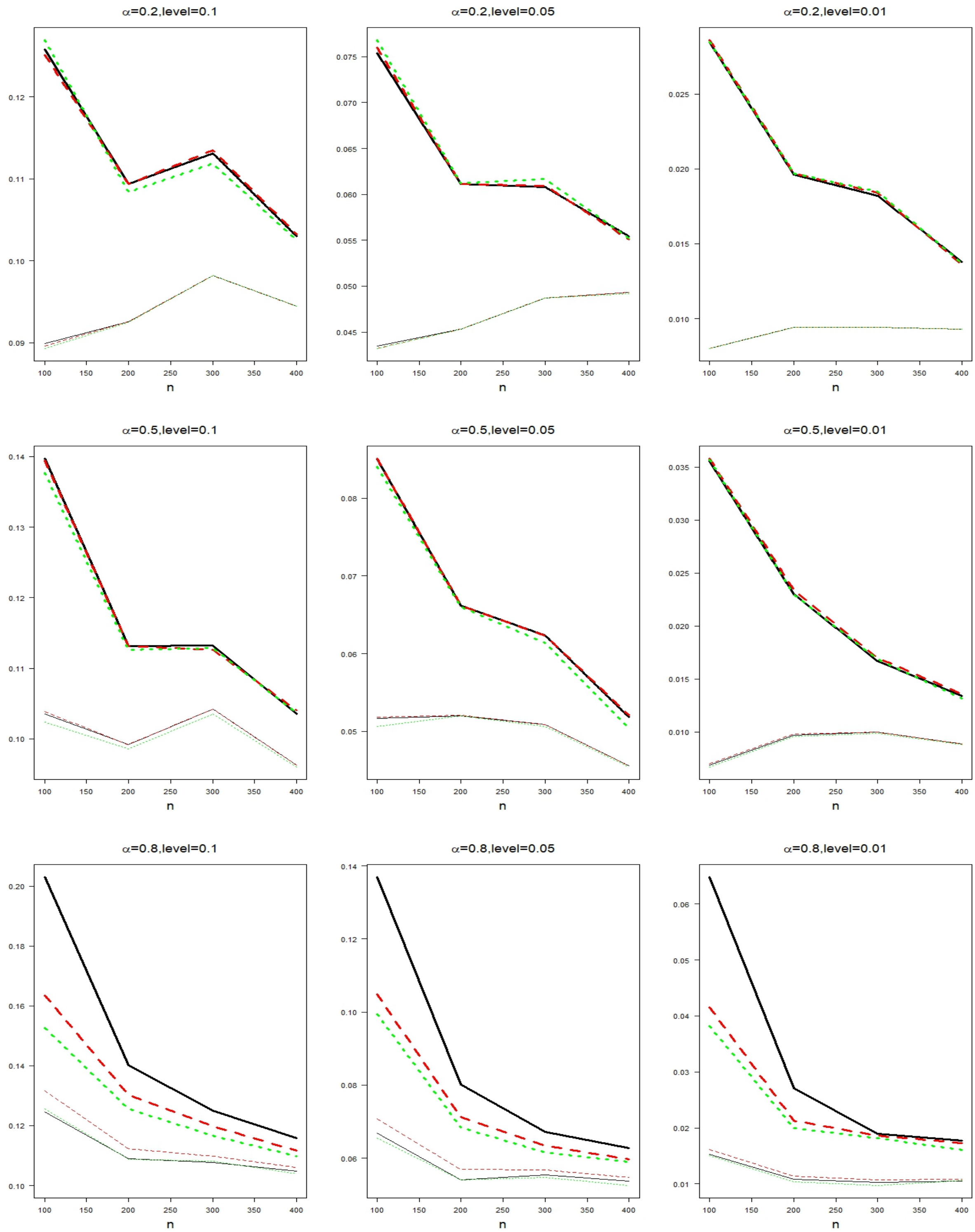

Figure 4.1: Empirical type I errors (10000 replications) of equidispersion tests having the rejection regions $|z_\#| > z(a/2)$ (thick lines) or $|z_{SW,\#}| > z(a/2)$ (thin lines), with the significant level $a = 0.1, 0.05, 0.01$, where $\# = W$(solid), YW(dashed), CLS(dotted). The upper, middle, and lower three panels correspond to $\alpha = 0.2, 0.5, 0.8$, respectively.

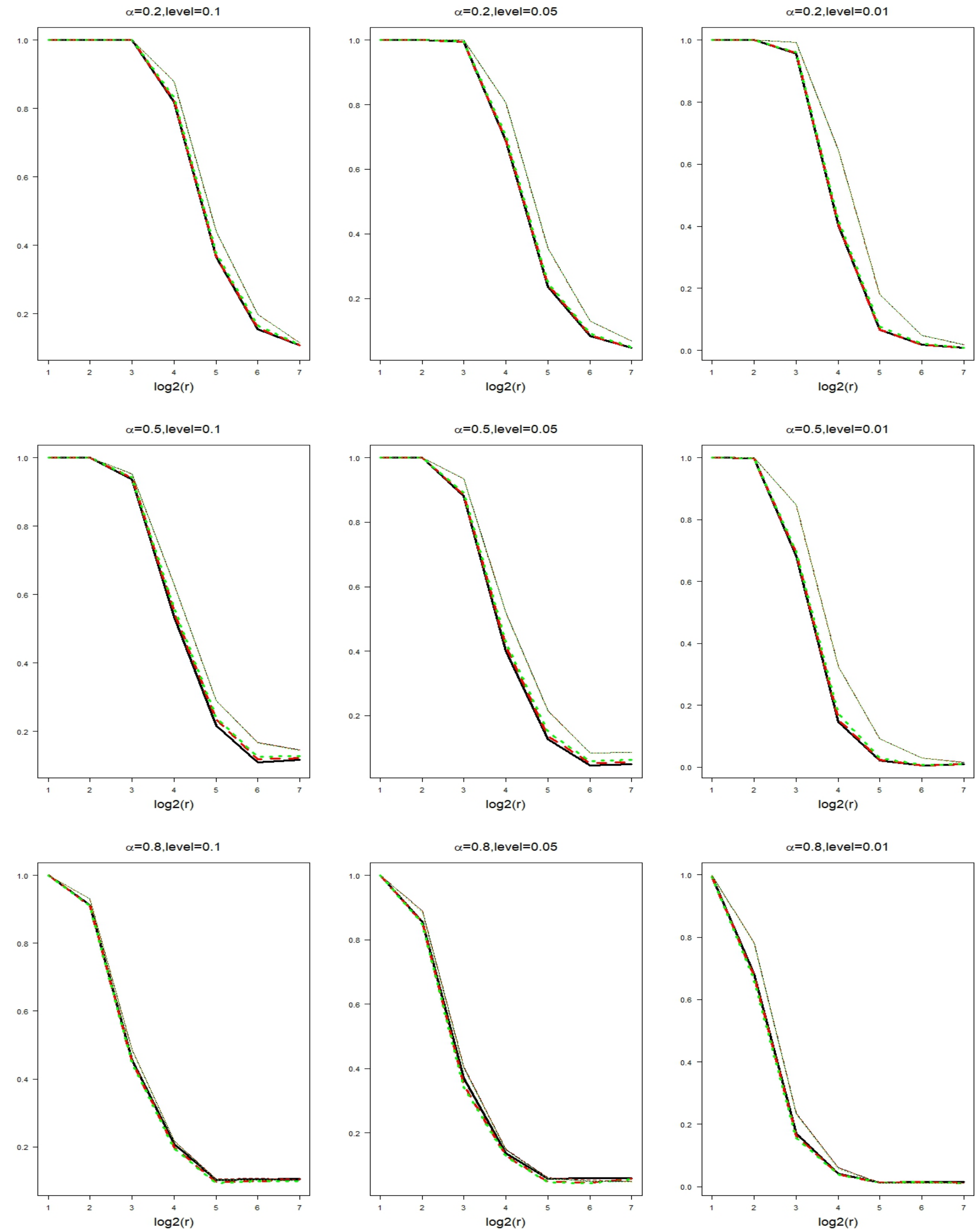

Figure 4.2: Empirical powers (10000 replications) of equidispersion tests having the rejection regions $|z_{\#}| > z(a/2)$ (thick lines) or $|z_{SW,\#}| > z(a/2)$ (thin lines), with the significant level $a = 0.1, 0.05, 0.01$, where $\# = W$(solid), YW(dashed), CLS(dotted). The upper, middle, and lower three panels correspond to $\alpha = 0.2, 0.5, 0.8$, respectively.

Note that $q(\alpha, \mu_\varepsilon) \geq q(\alpha_0, \mu_{\varepsilon,0}) = (1-\alpha_0^2)\sigma_{Y,0}^2$ *(the equality holds iff* $(\alpha, \mu_\varepsilon) = (\alpha_0, \mu_{\varepsilon,0})$*).*

Proof. For $h = 0, 1$, we have

$$\begin{aligned}
&\sup_{\theta' \in \Theta'} |\widehat{\gamma}_Y^\bullet(h) - \alpha_0^h \sigma_{Y,0}^2 - (\mu_{Y,0} - \mu_Y)^2| \\
&\leq \Big|\frac{1}{n}\sum_{t=1+h}^{n} (Y_t - \mu_{Y,0})(Y_{t-h} - \mu_{Y,0}) - \alpha_0^h \sigma_{Y,0}^2\Big| \\
&\quad + \Big(\sup_{\theta' \in \Theta'} |\mu_{Y,0} - \mu_Y|\Big)\Big|\frac{1}{n}\sum_{t=1+h}^{n} (Y_t - \mu_{Y,0} + Y_{t-h} - \mu_{Y,0})\Big| \\
&\quad + \Big(\sup_{\theta' \in \Theta'} (\mu_{Y,0} - \mu_Y)^2\Big)\frac{h}{n} \\
&\xrightarrow{a.s.} 0 \quad \text{(by the strictly stationarity and ergodicity of } \{Y_t\}\text{)}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\sup_{\theta' \in \Theta'} |\mathcal{J}_Y(\alpha, \mu_\varepsilon) - q(\alpha, \mu_\varepsilon)| \\
&= \sup_{\theta' \in \Theta'} \Big|(1+\alpha^2)[\widehat{\gamma}_Y^\bullet(0) - \{\sigma_{Y,0}^2 + (\mu_{Y,0} - \mu_Y)^2\}] - 2\alpha[\widehat{\gamma}_Y^\bullet(1) - \{\alpha_0\sigma_{Y,0}^2 + (\mu_{Y,0} - \mu_Y)\}^2\}]\Big| \\
&\leq \sup_{\theta' \in \Theta'} \Big|(1+\alpha^2)\{\widehat{\gamma}_Y^\bullet(0) - \sigma_{Y,0}^2 - (\mu_{Y,0} - \mu_Y)^2\}\Big| + \sup_{\theta' \in \Theta'} \Big|2\alpha\{\widehat{\gamma}_Y^\bullet(1) - \alpha_0\sigma_{Y,0}^2 - (\mu_{Y,0} - \mu_Y)^2\}\Big| \\
&\xrightarrow{a.s.} 0. \quad \square
\end{aligned}$$

Proof of Proposition 4.1. (i) By Lemma 4.1, we know (e.g., van der Vaart (1998)) that the Whittle estimator for θ' is strongly consistent, i.e., $(\widehat{\alpha}_W, \widehat{\mu}_{\varepsilon;W})^T \xrightarrow{a.s.} \theta'_0$. Then,

$$\widehat{\sigma^2_{\varepsilon;W}} \xrightarrow{a.s.} (1-\alpha_0^2)\sigma_{Y,0}^2 - \alpha_0\mu_{\varepsilon,0} = \sigma_{\varepsilon,0}^2.$$

(ii) The first-order conditions for $\widehat{\alpha}_W$ and $\widehat{\mu}_{\varepsilon;W}$ are given by

$$\begin{aligned}
0 &= \frac{\partial \mathcal{J}_Y(\widehat{\alpha}_W, \widehat{\mu}_{\varepsilon;W})}{\partial \alpha} \\
&= \frac{2\widehat{\alpha}_W}{n}\sum_{t=1}^{n}\Big(Y_t - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)^2 - \frac{2}{n}\sum_{t=1}^{n-1}\Big(Y_t - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)\Big(Y_{t+1} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big) \\
&\quad - \frac{2\widehat{\mu}_{\varepsilon;W}}{(1-\widehat{\alpha}_W)^2 n}\Big[(1+\widehat{\alpha}_W^2)\sum_{t=1}^{n}\Big(Y_t - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big) - 2\widehat{\alpha}_W\sum_{t=1}^{n}\Big(Y_t + \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big) + \widehat{\alpha}_W\Big(Y_1 + Y_n - \frac{2\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)\Big], \\
0 &= \frac{\partial \mathcal{J}_Y(\widehat{\alpha}_W, \widehat{\mu}_{\varepsilon;W})}{\partial \mu_\varepsilon} \\
&= -\frac{2}{(1-\widehat{\alpha}_W)n}\Big[(1+\widehat{\alpha}_W^2)\sum_{t=1}^{n}\Big(Y_t - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big) - 2\widehat{\alpha}_W\sum_{t=1}^{n}\Big(Y_t - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big) + \widehat{\alpha}_W\Big(Y_1 + Y_n - \frac{2\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)\Big],
\end{aligned}$$

respectively. Then, we have

$$\begin{aligned}
0 &= \frac{\widehat{\alpha}_W}{n}\sum_{t=1}^{n}\Big(Y_t - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)^2 - \frac{1}{n}\sum_{t=1}^{n-1}\Big(Y_t - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)\Big(Y_{t+1} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)\\
&= \frac{\widehat{\alpha}_W}{n}\sum_{t=1}^{n}(Y_t-\overline{Y})^2 - \frac{1}{n}\sum_{t=1}^{n-1}(Y_t-\overline{Y})(Y_{t+1}-\overline{Y}) + \Big(\widehat{\alpha}_W - 1 + \frac{1}{n}\Big)\Big(\overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big)^2\\
&\quad + \frac{1}{n}(Y_1+Y_n-2\overline{Y})\Big(\overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big),\\
0 &= \overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W} + \frac{\widehat{\alpha}_W}{(1-\widehat{\alpha}_W)^2 n}\Big(Y_1+Y_n - \frac{2\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W}\Big).
\end{aligned}$$

Using

$$\frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W} = \frac{\overline{Y} + \dfrac{\widehat{\alpha}_W(Y_1+Y_n)}{(1-\widehat{\alpha}_W)^2 n}}{1 + \dfrac{2\widehat{\alpha}_W}{(1-\widehat{\alpha}_W)^2 n}}, \quad \text{i.e.,} \quad \overline{Y} - \frac{\widehat{\mu}_{\varepsilon;W}}{1-\widehat{\alpha}_W} = \frac{\dfrac{\widehat{\alpha}_W(2\overline{Y}-Y_1-Y_n)}{(1-\widehat{\alpha}_W)^2 n}}{1 + \dfrac{2\widehat{\alpha}_W}{(1-\widehat{\alpha}_W)^2 n}} = O_p(n^{-1}),$$

we get

$$0 = \frac{\widehat{\alpha}_W}{n}\sum_{t=1}^{n}(Y_t-\overline{Y})^2 - \frac{1}{n}\sum_{t=1}^{n-1}(Y_t-\overline{Y})(Y_{t+1}-\overline{Y}) + O_p(n^{-2}),$$

which yields $\widehat{\alpha}_W = \widehat{\alpha}_{YW} + O_p(n^{-2})$.

Now, by Proposition 2.2(ii), we have

$$\sqrt{n}[\overline{Y} - \mu_{Y,0}, \widehat{\gamma}(0) - \sigma^2_{Y,0}, \widehat{\gamma}(1) - \alpha_0\sigma^2_{Y,0}]^T \xrightarrow{d} \mathrm{N}(0, \Xi_0), \tag{4.7}$$

where Ξ_0 is the true 3×3 matrix corresponding to Ξ (here, we set $m=1$). Noting that

$$\begin{aligned}
\widehat{\alpha}_W &= \frac{\widehat{\gamma}(1)}{\widehat{\gamma}(0)} + O_p(n^{-2})\\
&= \alpha_0 - \frac{\alpha_0}{\sigma^2_{Y,0}}\{\widehat{\gamma}(0)-\sigma^2_{Y,0}\} + \frac{1}{\sigma^2_{Y,0}}\{\widehat{\gamma}(1) - \alpha_0\sigma^2_{Y,0}\} + O_p(n^{-1}),\\
\widehat{\mu}_{\varepsilon;W} &= (1-\widehat{\alpha}_W)\overline{Y} + O_p(n^{-1})\\
&= \mu_{\varepsilon,0} + (1-\alpha_0)(\overline{Y}-\mu_{Y,0}) - \mu_{Y,0}(\widehat{\alpha}_W - \alpha_0) + O_p(n^{-1})\\
&= \mu_{\varepsilon,0} + (1-\alpha_0)(\overline{Y}-\mu_{Y,0}) + \frac{\alpha_0\mu_{Y,0}}{\sigma^2_{Y,0}}\{\widehat{\gamma}(0)-\sigma^2_{Y,0}\} - \frac{\mu_{Y,0}}{\sigma^2_{Y,0}}\{\widehat{\gamma}(1)-\alpha_0\sigma^2_{Y,0}\} + O_p(n^{-1}),\\
\widehat{\sigma^2_{\varepsilon;W}} &= (1-\widehat{\alpha}^2_W)\widehat{\gamma}(0) - \widehat{\alpha}_W\widehat{\mu}_{\varepsilon;W} + O_p(n^{-2})\\
&= \sigma^2_{\varepsilon,0} - (2\alpha_0\sigma^2_{Y,0} + \mu_{\varepsilon,0})(\widehat{\alpha}_W - \alpha_0) + (1-\alpha_0^2)\{\widehat{\gamma}(0)-\sigma^2_{Y,0}\} - \alpha_0(\widehat{\mu}_{\varepsilon;W} - \mu_{\varepsilon,0}) + O_p(n^{-1})\\
&= \sigma^2_{\varepsilon,0} - \alpha_0(1-\alpha_0)(\overline{Y}-\mu_{Y,0}) + \Big\{1 + \alpha_0^2 + \frac{\alpha_0(1-2\alpha_0)}{\sigma^2_{Y,0}}\mu_{Y,0}\Big\}\{\widehat{\gamma}(0)-\sigma^2_{Y,0}\}\\
&\quad - \Big(2\alpha_0 + \frac{1-2\alpha_0}{\sigma^2_{Y,0}}\mu_{Y,0}\Big)\{\widehat{\gamma}(1)-\alpha_0\sigma^2_{Y,0}\} + O_p(n^{-1}),
\end{aligned}$$

we have

$$\sqrt{n}(\widehat{\theta}_W - \theta_0) = \frac{\sqrt{n}}{\sigma^2_{Y,0}} A_0[\overline{Y} - \mu_{Y,0}, \widehat{\gamma}(0) - \sigma^2_{Y,0}, \widehat{\gamma}(1) - \alpha_0\sigma^2_{Y,0}]^T + O_p(n^{-1/2}),$$

where

$$A_0 = \begin{pmatrix} 0 & -\alpha_0 & 1 \\ (1-\alpha_0)\sigma^2_{Y,0} & \alpha_0\mu_{Y,0} & -\mu_{Y,0} \\ -\alpha_0(1-\alpha_0)\sigma^2_{Y,0} & (1+\alpha_0^2)\sigma^2_{Y,0} + \alpha_0(1-2\alpha_0)\mu_{Y,0} & -\{2\alpha_0\sigma^2_{Y,0} + (1-2\alpha_0)\mu_{Y,0}\} \end{pmatrix}.$$

This, together with (4.7), enables us to establish the asymptotic normality of $\widehat{\theta}_W$, with the asymptotic covariance matrix $A_0\Xi_0 A_0{}^T/\sigma^4_{Y,0}$. □

Proof of Proposition 4.2. By $\widehat{\alpha}_{(c_1,c_2)} \xrightarrow{a.s.} \alpha_0$ (see Proposition 3.1(i)), we have

$$\widehat{\mu}_{\varepsilon;(c_1,c_2)} \xrightarrow{a.s.} (1-\alpha_0)\mu_{Y,0} = \mu_{\varepsilon,0}, \widehat{\sigma^2}_{\varepsilon;(c_1,c_2)} \xrightarrow{a.s.} \text{ and } (1-\alpha_0^2)\sigma^2_{Y,0} - \alpha_0\mu_{\varepsilon,0} = \sigma^2_{\varepsilon,0}.$$

Since

$$\begin{aligned} \widehat{\alpha}_{(c_1,c_2)} &= \frac{n^{-1}\sum_{t=1}^{n-1}(Y_t - \overline{Y})(Y_{t+1} - \overline{Y})}{n^{-1}\sum_{t=1}^{n}(Y_t - \overline{Y})^2 + n^{-1}(c_1-1)(Y_1 - \overline{Y})^2 + n^{-1}(c_2-1)(Y_n - \overline{Y})^2} \\ &= \widehat{\alpha}_{YW} + O_p(n^{-1}), \\ \widehat{\mu}_{\varepsilon;(c_1,c_2)} &= (1-\widehat{\alpha}_{YW})\overline{Y} + O_p(n^{-1}), \\ \widehat{\sigma^2}_{\varepsilon;(c_1,c_2)} &= (1-\widehat{\alpha}^2_{YW})\widehat{\gamma}(0) - \widehat{\alpha}_{YW}(1-\widehat{\alpha}_{YW})\overline{Y} + O_p(n^{-1}), \end{aligned}$$

the asymptotic normality of $\widehat{\theta}_{(c_1,c_2)}$ is the same as that of $\widehat{\theta}_W$; see the proof of Proposition 4.1(ii).

Similarly, by the strictly stationarity and ergodicity of $\{Y_t\}$, we can show that

$$\widehat{\alpha}_{CLS} \xrightarrow{a.s.} \frac{E(Y_1Y_2) - \mu^2_{Y,0}}{E(Y_1^2) - \mu^2_{Y,0}} = \alpha_0 \text{ and } \widehat{\mu}_{\varepsilon;CLS} \xrightarrow{a.s.} \mu_{Y,0} - \alpha_0\mu_{Y,0} = \mu_{\varepsilon,0},$$

hence,

$$\widehat{\sigma^2}_{\varepsilon;CLS} \xrightarrow{a.s.} (1-\alpha_0^2)\sigma^2_{Y,0} - \alpha_0\mu_{\varepsilon,0} = \sigma^2_{\varepsilon,0}.$$

Also, we know that

$$\begin{aligned} \widehat{\alpha}_{CLS} &= \widehat{\alpha}_{(1,0)} + O_p(n^{-2}), \\ \widehat{\mu}_{\varepsilon;CLS} &= \frac{n\overline{Y} - Y_1}{n-1} - \{\widehat{\alpha}_{(1,0)} + O_p(n^{-2})\}\frac{n\overline{Y} - Y_n}{n-1} \\ &= (1-\widehat{\alpha}_{(1,0)})\overline{Y} + O_p(n^{-1}), \end{aligned}$$

$$\widehat{\sigma^2_{\varepsilon;CLS}} = (1 - \widehat{\alpha}^2_{(1,0)})\widehat{\gamma}(0) - \widehat{\alpha}_{(1,0)}(1 - \widehat{\alpha}_{(1,0)})\overline{Y} + O_p(n^{-1}).$$

The asymptotic normality of $\widehat{\theta}_{CLS}$ is the same as that of $\widehat{\theta}_{(1,0)}$. □

Chapter 5

Probabilistic and statistical properties of ADCINAR(1) process

5.1 Introduction

A nonnegative integer-valued autoregressive process of the first-order (INAR(1)) was proposed by Al-Osh and Alzaid (1987), based on the binomial thinning operator due to Steutel and van Harn (1979). Since then, there has been huge basic and interesting works for statistical inference of time series count data (see Weiß (2008,2018)).

On the other hand, considering that such a binomial thinning operator was defined through the independent counting series, Nastić et al. (2017) (see also Ristić et al. (2013)) recently attempted to change it into a dependent counting series and then introduced an alternative generalized binomial thinning operator. In this way, they defined an alternative dependent counting nonnegative INAR process of the first-order (ADCINAR(1)). However, to the best of our knowledge, higher autocumulant functions, except for the autocovariance function, have not been discussed so far, for such an ADCINAR(1) process. As an extension of Chapter 2, we derive, explicitly, the third and fourth autocumulant functions of the stationary ADCINAR(1) process under a general innovation, together with the structure about arbitrary higher autocumulant functions.

The rest of this chapter is organized as follows. Section 5.2 gives the definition and some useful properties of the generalized binomial thinning operator. After the introduction of the ADCINAR(1) process and its basic formulas of the moment, variance, and autocorrelation function at lag $u(\geq 0)$, Section 5.3 discusses higher autocumulant functions of the stationary ADCINAR(1) process. The technical proofs of Propositions 5.1 and 5.2 are postponed to Section 5.4.

5.2 Generalized binomial thinning operator

In Chapter 2, we introduced the binomial thinning operator $\alpha\circ$, which consists of independent counting series. Although such a counting series is fundamental, an alternative dependent formulation will be important for modelling the time series count data.

Given $0 \le \alpha \le \vartheta \le 1 (\vartheta \neq 0)$, we construct a sequence of random variables $\{S_j(\alpha,\vartheta)\}$ such that

$$S_j(\alpha,\vartheta) = B_j(\vartheta)B(\alpha/\vartheta),$$

where $\{B_j(\vartheta)\}$ is a sequence of independent and identically distributed (IID) Bernoulli random variables, which is independent of $B(\alpha/\vartheta)$. An alternative generalized binomial thinning operator was recently introduced by Nastić et al. (2017), as follows. Given a nonnegative integer-valued random variable Y, let

$$\alpha \diamond_\vartheta Y = \begin{cases} 0, & Y = 0, \\ \sum_{j=1}^{Y} S_j(\alpha,\vartheta), & Y = 1,2,\ldots, \end{cases}$$

where $\{B_j(\vartheta)\}$ and $B(\alpha/\vartheta)$ are independent of Y. Using

$$B_j(\vartheta)B(\alpha/\vartheta) = \begin{cases} 0, & \text{with probability } 1 - \dfrac{\alpha}{\vartheta}, \\ B_j(\vartheta), & \text{with probability } \dfrac{\alpha}{\vartheta}, \end{cases}$$

the probability generating function (pgf) of $S_j(\alpha,\vartheta)$ is given by

$$E[u^{B_j(\vartheta)B(\alpha/\vartheta)}] = 1 - \frac{\alpha}{\vartheta} + \frac{\alpha}{\vartheta}E[u^{B_j(\vartheta)}] = 1 - \frac{\alpha}{\vartheta} + \frac{\alpha}{\vartheta}(1 - \vartheta + \vartheta u) = 1 - \alpha + \alpha u,$$

so that $S_j(\alpha,\vartheta)$ is distributed as the Bernoulli distribution $\mathrm{Bin}(1,\alpha)$, whereas $\{S_j(\alpha,\vartheta)\}$ is, in general, a dependent sequence, i.e., for $i \neq j$,

$$E[S_i(\alpha,\vartheta)S_j(\alpha,\vartheta)] = E[B_i(\vartheta)]E[B_j(\vartheta)]E[B^2(\alpha/\vartheta)] = \alpha\vartheta,$$

hence,

$$Cov[S_i(\alpha,\vartheta), S_j(\alpha,\vartheta)] = \alpha(\vartheta - \alpha).$$

The following basic results (Lemmas 5.1 and 5.2) of the generalized binomial thinning operation are repeatedly used (Lemma 5.1 was shown by Nastić et al. (2017)).

Lemma 5.1. *(i) For* $0 < \gamma \le 1$, $0 \diamond_\gamma Y = 0$ *and* $1 \diamond_1 Y = Y$.

(ii) The pgf of $\alpha \diamond_\vartheta Y$ is given by

$$E[u^{\alpha \diamond_\vartheta Y}] = 1 - \frac{\alpha}{\vartheta} + \frac{\alpha}{\vartheta} E[(1 - \vartheta + \vartheta u)^Y].$$

(iii) For $0 \le \beta \le \gamma \le 1 (\gamma \ne 0)$ and $0 \le \delta \le \eta \le 1 (\eta \ne 0)$,

$$\delta \diamond_\eta (\beta \diamond_\gamma Y) \stackrel{d}{=} (\beta\delta) \diamond_{(\gamma\eta)} Y \stackrel{d}{=} \beta \diamond_\gamma (\delta \diamond_\eta Y),$$

where $\stackrel{d}{=}$ stands for equal in distribution (it is implicitly assumed that the operations $\beta\diamond_\gamma$ and $\delta\diamond_\eta$ are performed independently, which are independent of Y).

Proof By definition, (i) is trivial. On the other hand, noting that

$$\alpha \diamond_\vartheta Y = \begin{cases} 0, & \text{with probability } 1 - \dfrac{\alpha}{\vartheta}, \\ \vartheta \circ Y, & \text{with probability } \dfrac{\alpha}{\vartheta}, \end{cases}$$

and that, given Y, $\vartheta \circ Y (= \sum_{j=1}^{Y} B_j(\vartheta))$ is distributed as the binomial distribution $\mathrm{Bin}(Y, \vartheta)$, (ii) follows from

$$E[u^{\alpha \diamond_\vartheta Y}] = 1 - \frac{\alpha}{\vartheta} + \frac{\alpha}{\vartheta} E[u^{\vartheta \circ Y}] = 1 - \frac{\alpha}{\vartheta} + \frac{\alpha}{\vartheta} E[(1 - \vartheta + \vartheta u)^Y]$$

(use the law of total expectation). Then, (iii) is verified by means of the pgf

$$\begin{aligned} E[u^{\delta \diamond_\eta (\beta \diamond_\gamma Y)}] &= 1 - \frac{\delta}{\eta} + \frac{\delta}{\eta} E[(1 - \eta + \eta u)^{\beta \diamond_\gamma Y}] \\ &= 1 - \frac{\delta}{\eta} + \frac{\delta}{\eta}\left[1 - \frac{\beta}{\gamma} + \frac{\beta}{\gamma} E[\{1 - \gamma + \gamma(1 - \eta + \eta u)\}^Y]\right] \\ &= 1 - \frac{\delta\beta}{\eta\gamma} + \frac{\delta\beta}{\eta\gamma} E[(1 - \eta\gamma + \eta\gamma u)^Y] \\ &= E[u^{(\delta\beta) \diamond_{(\eta\gamma)} Y}]. \quad \square \end{aligned}$$

Let

$$M_\ell(x, p) = \sum_{i=1}^{\ell} p^{i-1} S\binom{\ell}{i} \sum_{j=1}^{i} s\binom{i}{j} x^j = \sum_{j=1}^{\ell} x^j \sum_{i=j}^{\ell} p^{i-1} S\binom{\ell}{i} s\binom{i}{j}$$

be a polynomial of degree ℓ in x (without constant term) for any $p \in [0, 1]$ and positive integer ℓ, where $s\binom{i}{j}$ and $S\binom{\ell}{i}$ are the Stirling numbers of the first and second kind, respectively; for these definitions, see Olver et al. (2010; Section 26.8). Given Y, $\sum_{j=1}^{Y} B_j(\vartheta)$ is distributed as the binomial distribution

$\text{Bin}(Y, \vartheta)$, so that, for any positive integer ℓ,

$$E\Big[\Big\{\sum_{j=1}^{Y} B_j(\vartheta)\Big\}^{\ell} | Y\Big] = \vartheta M_{\ell}(Y, \vartheta) \quad \text{(e.g., Johnson et al. (2005)).} \tag{5.1}$$

Lemma 5.2. *(i) For a function G (e.g., we set $G(Y) \equiv 1$ or $G(Y) = Y - E(Y)$),*

$$\begin{aligned}
E[G(Y)(\alpha \diamond_{\vartheta} Y)] &= \alpha E[G(Y)Y], \\
E[G(Y)(\alpha \diamond_{\vartheta} Y)^2] &= \alpha\Big\{\vartheta E[G(Y)Y^2] + (1-\vartheta)E[G(Y)Y]\Big\}, \\
E[G(Y)(\alpha \diamond_{\vartheta} Y)^3] &= \alpha\Big\{\vartheta^2 E[G(Y)Y^3] + 3\vartheta(1-\vartheta)E[G(Y)Y^2] + (1-\vartheta)(1-2\vartheta)E[G(Y)Y]\Big\}, \\
E[G(Y)(\alpha \diamond_{\vartheta} Y)^4] &= \alpha\Big\{\vartheta^3 E[G(Y)Y^4] + 6\vartheta^2(1-\vartheta)E[G(Y)Y^3] \\
&\quad + \vartheta(1-\vartheta)(7-11\vartheta)E[G(Y)Y^2] + (1-\vartheta)(1-6\vartheta+6\vartheta^2)E[G(Y)Y]\Big\}
\end{aligned}$$

(it is implicitly assumed that the expectations in the right-hand side exist).

(ii) Also,

$$E[G(Y)(\alpha \diamond_{\vartheta} Y - \alpha E(Y))] = \alpha E[G(Y)(Y - E(Y))],$$

hence,

$$Cov[G(Y), \alpha \diamond_{\vartheta} Y] = \alpha E[G(Y)(Y - E(Y))]$$

(it is implicitly assumed that $E[G(Y)(Y - E(Y))]$ exists).

(iii) Furthermore,

$$\begin{aligned}
&E[G(Y)(\alpha \diamond_{\vartheta} Y - \alpha E(Y))^2] \\
&= \alpha\vartheta E[G(Y)(Y - E(Y))^2] + \alpha(1-\vartheta)\Big\{E[G(Y)(Y - E(Y))] + E[G(Y)]E(Y)\Big\} \\
&\quad + \alpha(\vartheta - \alpha)\Big\{2E[G(Y)(Y - E(Y))]E(Y) + E[G(Y)][E(Y)]^2\Big\}, \\
&E[G(Y)(\alpha \diamond_{\vartheta} Y - \alpha E(Y))^3] \\
&= \alpha\vartheta^2 E[G(Y)(Y - E(Y))^3] + 3\alpha\vartheta(1-\vartheta)\Big\{E[G(Y)(Y - E(Y))^2] + E[G(Y)(Y - E(Y))]E(Y)\Big\} \\
&\quad + \alpha(1-\vartheta)(1-2\vartheta)\Big\{E[G(Y)(Y - E(Y))] + E[G(Y)]E(Y)\Big\} \\
&\quad + \alpha(\vartheta - \alpha)\Big[3\vartheta E[G(Y)(Y - E(Y))^2]E(Y) + 3(\vartheta - \alpha)E[G(Y)(Y - E(Y))][E(Y)]^2 \\
&\qquad + (\vartheta - 2\alpha)E[G(Y)][E(Y)]^3 + 3(1-\vartheta)\Big\{E[G(Y)(Y - E(Y))] + E[G(Y)][E(Y)]^2\Big\}\Big], \\
&E[G(Y)(\alpha \diamond_{\vartheta} Y - \alpha E(Y))^4] \\
&= \alpha\vartheta^3 E[G(Y)(Y - E(Y))^4] + 6\alpha\vartheta^2(1-\vartheta)\Big\{E[G(Y)(Y - E(Y))^3] + E[G(Y)(Y - E(Y))^2]E(Y)\Big\}
\end{aligned}$$

$$+\alpha\vartheta(1-\vartheta)(7-11\vartheta)E[G(Y)(Y-E(Y))^2] + 2\alpha\vartheta(1-\vartheta)(5-7\vartheta)E[G(Y)(Y-E(Y))]E(Y)$$
$$+\alpha(1-\vartheta)(1-6\vartheta+6\vartheta^2)\Big\{E[G(Y)(Y-E(Y))] + E[G(Y)]E(Y)\Big\} + 3\alpha\vartheta(1-\vartheta)^2E[G(Y)][E(Y)]^2$$
$$+\alpha(\vartheta-\alpha)\Big\{4\vartheta^2E[G(Y)(Y-E(Y))^3]E(Y) + 12\vartheta(1-\vartheta)E[G(Y)(Y-E(Y))^2]E(Y)$$
$$+6\vartheta(\vartheta-\alpha)E[G(Y)(Y-E(Y))^2][E(Y)]^2 + 4\vartheta(1-\vartheta)(1-2\vartheta)E[G(Y)(Y-E(Y))]E(Y)$$
$$+6(1-\vartheta)(3\vartheta-\alpha)E[G(Y)(Y-E(Y))][E(Y)]^2 + 4(\vartheta-\alpha)^2E[G(Y)(Y-E(Y))][E(Y)]^3$$
$$+4(1-\vartheta)(1-2\vartheta)E[G(Y)][E(Y)]^2 + 6(1-\vartheta)(\vartheta-\alpha)E[G(Y)][E(Y)]^3$$
$$+(\vartheta^2-3\alpha\vartheta+3\alpha^2)E[G(Y)][E(Y)]^4\Big\}$$

(it is implicitly assumed that the expectations in the right-hand side exist).

Proof (i) For any positive integer ℓ,

$$E\Big[\Big\{\sum_{j=1}^{Y} S_j(\alpha,\vartheta)\Big\}^{\ell}|Y\Big] = E[B^{\ell}(\alpha/\vartheta)]E\Big[\Big\{\sum_{j=1}^{Y} B_j(\vartheta)\Big\}^{\ell}|Y\Big] = \alpha M_{\ell}(Y,\vartheta) \ \ \text{(by (5.1))}.$$

The result follows from $E[G(Y)(\alpha\diamond_{\vartheta} Y)^{\ell}] = E[G(Y)E[(\alpha\diamond_{\vartheta} Y)^{\ell}|Y]]$.

(ii)&(iii) Use (i) and the binomial theorem; $(A+B)^m = \sum_{i=0}^{m} {}_mC_i A^{m-i}B^i$, $m = 1, 2, \ldots$. □

In addition to the moment of $\alpha\diamond_{\vartheta} Y$, the second, third, and fourth central moments (or cumulants) of $\alpha\diamond_{\vartheta} Y$ are derived as direct consequences of Lemma 5.2(i,iii) with $G(Y)\equiv 1$.

Corollary 5.1. *(i) When $E(Y)$ exists, then,*

$$E(\alpha\diamond_{\vartheta} Y) = \alpha E(Y).$$

(ii) The jth central moments of $\alpha\diamond_{\vartheta} Y$, $j = 2, 3, 4$ (when $E(Y^j)$ exists), are given by

$$E[(\alpha\diamond_{\vartheta} Y-\alpha E(Y))^2] = \alpha\vartheta E[(Y-E(Y))^2] + \alpha(1-\vartheta)E(Y) + \alpha(\vartheta-\alpha)[E(Y)]^2,$$
$$E[(\alpha\diamond_{\vartheta} Y-\alpha E(Y))^3] = \alpha\vartheta^2E[(Y-E(Y))^3] + 3\alpha\vartheta(1-\vartheta)E[(Y-E(Y))^2] + \alpha(1-\vartheta)(1-2\vartheta)E(Y)$$
$$+\alpha(\vartheta-\alpha)\Big\{3\vartheta E[(Y-E(Y))^2]E(Y) + 3(1-\vartheta)[E(Y)]^2 + (\vartheta-2\alpha)[E(Y)]^3\Big\},$$
$$E[(\alpha\diamond_{\vartheta} Y-\alpha E(Y))^4] = \alpha\vartheta^3E[(Y-E(Y))^4] + 6\alpha\vartheta^2(1-\vartheta)\Big\{E[(Y-E(Y))^3] + E[(Y-E(Y))^2]E(Y)\Big\}$$
$$+\alpha\vartheta(1-\vartheta)(7-11\vartheta)E[(Y-E(Y))^2] + 3\alpha\vartheta(1-\vartheta)^2[E(Y)]^2$$
$$+\alpha(1-\vartheta)(1-6\vartheta+6\vartheta^2)E(Y)$$
$$+\alpha(\vartheta-\alpha)\Big\{4\vartheta^2E[(Y-E(Y))^3]E(Y) + 12\vartheta(1-\vartheta)E[(Y-E(Y))^2]E(Y)$$
$$+6\vartheta(\vartheta-\alpha)E[(Y-E(Y))^2][E(Y)]^2 + 4(1-\vartheta)(1-2\vartheta)[E(Y)]^2$$
$$+6(1-\vartheta)(\vartheta-\alpha)[E(Y)]^3 + (\vartheta^2-3\alpha\vartheta+3\alpha^2)[E(Y)]^4\Big\}.$$

Recall that, for a random variable X with finite fourth moment,

$$
\begin{aligned}
V(X) &= E[(X-E(X))^2] \quad \text{(note that } Cum_2(X)=V(X)\text{)},\\
Cum_3(X) &= E[(X-E(X))^3],\\
Cum_4(X) &= E[(X-E(X))^4]-3[V(X)]^2.
\end{aligned}
$$

Here, instead of writing $Cum(\underbrace{X,\ldots,X}_{j\text{ times}})$, we use the notation $Cum_j(X)$ for the jth cumulant, where $j=2,3,\ldots$.

Corollary 5.2. *The jth cumulants of $\alpha\diamond_\vartheta Y$, $j=2,3,4$ (when $E(Y^j)$ exists), are given by*

$$
\begin{aligned}
V(\alpha\diamond_\vartheta Y) &= \alpha\vartheta V(Y)+\alpha(1-\vartheta)E(Y)+\alpha(\vartheta-\alpha)[E(Y)]^2,\\
Cum_3(\alpha\diamond_\vartheta Y) &= \alpha\vartheta^2 Cum_3(Y)+3\alpha\vartheta(1-\vartheta)V(Y)+\alpha(1-\vartheta)(1-2\vartheta)E(Y)\\
&\quad+\alpha(\vartheta-\alpha)\Big\{3\vartheta V(Y)E(Y)+3(1-\vartheta)[E(Y)]^2+(\vartheta-2\alpha)[E(Y)]^3\Big\},\\
Cum_4(\alpha\diamond_\vartheta Y) &= \alpha\vartheta^3 Cum_4(Y)+6\alpha\vartheta^2(1-\vartheta)Cum_3(Y)+\alpha\vartheta(1-\vartheta)(7-11\vartheta)V(Y)\\
&\quad+\alpha(1-\vartheta)(1-6\vartheta+6\vartheta^2)E(Y)\\
&\quad+\alpha(\vartheta-\alpha)\Big\{4\vartheta^2 Cum_3(Y)E(Y)+3\vartheta^2[V(Y)]^2+18\vartheta(1-\vartheta)V(Y)E(Y)\\
&\qquad+6\vartheta(\vartheta-2\alpha)V(Y)[E(Y)]^2+(1-\vartheta)(7-11\vartheta)[E(Y)]^2\\
&\qquad+6(1-\vartheta)(\vartheta-2\alpha)[E(Y)]^3+(\vartheta^2-6\alpha\vartheta+6\alpha^2)[E(Y)]^4\Big\}.
\end{aligned}
$$

We further prepare the following lemmas.

Lemma 5.3. *For $0\le\beta\le\gamma\le 1$ ($\gamma\ne 0$), $0\le\delta\le\eta\le 1$ ($\eta\ne 0$), and $0\le\iota\le\kappa\le 1$ ($\kappa\ne 0$),*

$$
\begin{aligned}
&E[G(Y)(\beta\diamond_\gamma Y-\beta E(Y))(\delta\diamond_\eta(\beta\diamond_\gamma Y)-\delta\beta E(Y))]\\
&=\delta\beta\Big[\gamma E[G(Y)(Y-E(Y))^2]+(1-\gamma)\Big\{E[G(Y)(Y-E(Y))]+E[G(Y)]E(Y)\Big\}\\
&\qquad+(\gamma-\beta)\Big\{2E[G(Y)(Y-E(Y))]E(Y)+E[G(Y)][E(Y)]^2\Big\}\Big], \qquad (5.2)\\
&E[G(Y)(\beta\diamond_\gamma Y-\beta E(Y))(\delta\diamond_\eta(\beta\diamond_\gamma Y)-\delta\beta E(Y))(\iota\diamond_\kappa(\delta\diamond_\eta(\beta\diamond_\gamma Y))-\iota\delta\beta E(Y))]\\
&=\iota\delta\beta\Big\{\eta\gamma^2E[G(Y)(Y-E(Y))^3]+\gamma(1+2\eta-3\eta\gamma)E[G(Y)(Y-E(Y))^2]\\
&\qquad+\gamma(3\eta\gamma-\eta\beta-2\delta\beta)E[G(Y)(Y-E(Y))^2]E(Y)\\
&\qquad+(1+\eta\gamma-3\eta\beta-\gamma-\eta\gamma^2+3\eta\beta\gamma)E[G(Y)(Y-E(Y))]\\
&\qquad+(2\gamma-\beta+3\eta\beta-2\delta\beta+\eta\gamma+2\delta\gamma\beta-2\eta\gamma\beta-3\eta\gamma^2)E[G(Y)(Y-E(Y))]E(Y)\\
&\qquad+(3\eta\gamma^2-4\delta\gamma\beta-2\eta\beta\gamma+3\delta\beta^2)E[G(Y)(Y-E(Y))][E(Y)]^2\\
&\qquad+(1-\gamma-2\eta\gamma+2\eta\gamma^2)E[G(Y)]E(Y)
\end{aligned}
$$

$$
\begin{aligned}
&+ (\gamma - \beta + 2\eta\gamma - 2\delta\beta + \eta\gamma\beta + 2\delta\gamma\beta - 3\eta\gamma^2)E[G(Y)][E(Y)]^2 \\
&+ (\eta\gamma^2 - \eta\gamma\beta + 2\delta\beta^2 - 2\delta\beta\gamma)E[G(Y)][E(Y)]^3\Big\} \qquad (5.3)
\end{aligned}
$$

(it is implicitly assumed that the operations $\beta\diamond_\gamma$, $\delta\diamond_\eta$, and $\iota\diamond_\kappa$ are performed mutually independently, which are independent of Y, and that the expectations in the right-hand side exist).

Proof In this proof, let $\mu = E(Y)$. We can easily prove (5.2) via the conditional expectation (given Y and all operations from $\beta\diamond_\gamma$), i.e.,

$$
\begin{aligned}
&E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)(\delta\diamond_\eta(\beta\diamond_\gamma Y) - \delta\beta\mu)] \\
&= \delta E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)^2] \\
&= \delta\Big[\beta\gamma E[G(Y)(Y-\mu)^2] + \beta(1-\gamma)\Big\{E[G(Y)(Y-\mu)] + E[G(Y)]\mu\Big\} \\
&\qquad + \beta(\gamma-\beta)\Big\{2E[G(Y)(Y-\mu)]\mu + E[G(Y)]\mu^2\Big\}\Big] \quad \text{(by Lemma 5.2(iii))}.
\end{aligned}
$$

Similarly, to prove (5.3), we use the law of total expectation after taking the conditional expectation (given Y and all operations from $\beta\diamond_\gamma$; note that $E(\beta\diamond_\gamma Y) = \beta\mu$), in relation to (5.2) with $(\beta,\gamma,\delta,\eta)$ replaced by $(\delta,\eta,\iota,\kappa)$. In this way, we obtain, after some tedious algebra,

$$
\begin{aligned}
&E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)(\delta\diamond_\eta(\beta\diamond_\gamma Y) - \delta\beta\mu)(\iota\diamond_\kappa(\delta\diamond_\eta(\beta\diamond_\gamma Y)) - \iota\delta\beta\mu)] \\
&= \iota\delta\Big[\eta E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)^3] + (1-\eta)\Big\{E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)^2] + E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)](\beta\mu)\Big\} \\
&\qquad + (\eta-\delta)\Big\{2E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)^2](\beta\mu) + E[G(Y)(\beta\diamond_\gamma Y - \beta\mu)](\beta\mu)^2\Big\}\Big] \\
&= \iota\delta\eta\Big[\beta\gamma^2 E[G(Y)(Y-\mu)^3] + 3\beta\gamma(1-\gamma)\Big\{E[G(Y)(Y-\mu)^2] + E[G(Y)(Y-\mu)]\mu\Big\} \\
&\qquad + \beta(1-\gamma)(1-2\gamma)\Big\{E[G(Y)(Y-\mu)] + E[G(Y)]\mu\Big\} \\
&\qquad + 3\beta\gamma(\gamma-\beta)E[G(Y)(Y-\mu)^2]\mu + 3\beta(\gamma-\beta)^2 E[G(Y)(Y-\mu)]\mu^2 \\
&\qquad + \beta(\gamma-\beta)(\gamma-2\beta)E[G(Y)]\mu^3 + 3\beta(1-\gamma)(\gamma-\beta)\Big\{E[G(Y)(Y-\mu)] + E[G(Y)]\mu^2\Big\}\Big] \\
&\quad + \iota\delta(1-\eta)\Big[\beta\gamma E[G(Y)(Y-\mu)^2] + \beta(1-\gamma)\Big\{E[G(Y)(Y-\mu)] + E[G(Y)]\mu\Big\} \\
&\qquad\qquad + \beta(\gamma-\beta)\Big\{2E[G(Y)(Y-\mu)]\mu + E[G(Y)]\mu^2\Big\}\Big] \\
&\quad + \iota\delta\beta^2(1-\eta)E[G(Y)(Y-\mu)]\mu \\
&\quad + 2\iota\delta\beta(\eta-\delta)\Big[\beta\gamma E[G(Y)(Y-\mu)^2] + \beta(1-\gamma)\Big\{E[G(Y)(Y-\mu)] + E[G(Y)]\mu\Big\} \\
&\qquad\qquad + \beta(\gamma-\beta)\Big\{2E[G(Y)(Y-\mu)]\mu + E[G(Y)]\mu^2\Big\}\Big]\mu \\
&\quad + \iota\delta\beta^3(\eta-\delta)E[G(Y)(Y-\mu)]\mu^2 \quad \text{(by Lemma 5.2(ii,iii))}. \qquad \square
\end{aligned}
$$

We use the notation $'^{(s)}$ in such way that $'^{(0)}$ stands for "without prime", $'^{(1)}$ stands for prime, $'^{(2)}$ stands for double prime, and so on. By definition, given $0 \le \beta \le \gamma \le 1 (\gamma \ne 0)$, for any positive integer

m and nonnegative integer-valued random variables $X, X'^{(1)}, \dots, X'^{(m)}$, it should be indicated explicitly that

$$\beta \diamond_\gamma \Big(X + \sum_{i=1}^m X'^{(i)}\Big) = \sum_{j=1}^X B_j(\gamma)B(\beta/\gamma) + \sum_{i=1}^m \sum_{j=1}^{X'^{(i)}} B'^{(i)}_j(\gamma)B(\beta/\gamma)$$
$$= \beta \diamond_\gamma X + \sum_{i=1}^m \beta \diamond'^{(i)}_\gamma X'^{(i)} \quad \text{(say)},$$

where $B'^{(i)}_j(\gamma)$'s are independent copies of $B_1(\gamma)$. Similarly to (5.1), it is easy to see that, for any positive integers $\ell'^{(0)}, \dots, \ell'^{(m)}$ (here, $m = 0$ is allowed),

$$E\Big[\prod_{j=0}^m (\beta \diamond'^{(j)}_\gamma X'^{(j)})^{\ell'^{(j)}} | X'^{(0)}, \dots, X'^{(m)}\Big] = \beta\gamma^m \prod_{j=0}^m M_{\ell'^{(j)}}(X'^{(j)}, \gamma). \tag{5.4}$$

Lemma 5.4. *Suppose that nonnegative integer-valued random variables X, X', and X'' are not necessarily independent. For $0 \le \beta \le \gamma \le 1 (\gamma \ne 0)$, assume that the operations $\beta\diamond_\gamma$, $\beta\diamond'_\gamma$, and $\beta\diamond''_\gamma$ are independent of (X, X', X''), in which the same Bernoulli random variable $B(\beta/\gamma)$ is used for $\beta\diamond_\gamma$, $\beta\diamond'_\gamma$, and $\beta\diamond''_\gamma$, whereas an IID Bernoulli sequence $\{B_i(\gamma)\}$ for $\beta\diamond_\gamma$, an IID Bernoulli sequence $\{B'_i(\gamma)\}$ for $\beta\diamond'_\gamma$, and an IID Bernoulli sequence $\{B''_i(\gamma)\}$ for $\beta\diamond''_\gamma$ are mutually independent. Then, in addition to*

$$E(\beta \diamond_\gamma X | X) = \beta X, \tag{5.5}$$
$$E[(\beta \diamond_\gamma X)(\beta \diamond'_\gamma Y)|X, Y] = \beta\gamma XY, \tag{5.6}$$
$$E[(\beta \diamond_\gamma X)(\beta \diamond'_\gamma Y)(\beta \diamond''_\gamma Z)|X, Y, Z] = \beta\gamma^2 XYZ, \tag{5.7}$$

the following hold.

(*i*) $Cov(\beta \diamond_\gamma X, \beta \diamond'_\gamma Y | X, Y) = \beta(\gamma - \beta)XY$.

(*ii*) $V(\beta \diamond_\gamma X | X) = \beta\{(\gamma - \beta)X^2 + (1 - \gamma)X\}$.

(*iii*) $Cum(\beta \diamond_\gamma X, \beta \diamond_\gamma X, \beta \diamond'_\gamma Y | X, Y) = \beta(\gamma - \beta)\{(\gamma - 2\beta)X^2Y + (1 - \gamma)XY\}$.

(*iv*) $Cum(\beta \diamond_\gamma X, \beta \diamond'_\gamma Y, \beta \diamond''_\gamma Z | X, Y, Z) = \beta(\gamma - \beta)(\gamma - 2\beta)XYZ$.

(*v*) $Cov[(\beta \diamond_\gamma X)^2, \beta \diamond'_\gamma Y | X, Y] = \beta(\gamma - \beta)\{\gamma X^2Y + (1 - \gamma)XY\}$.

(*vi*) $Cov[(\beta \diamond_\gamma X)(\beta \diamond'_\gamma Y), \beta \diamond_\gamma X | X, Y] = \beta\gamma\{(\gamma - \beta)X^2Y + (1 - \gamma)XY\}$.

(*vii*) $Cov[(\beta \diamond_\gamma X)(\beta \diamond'_\gamma Y), \beta \diamond''_\gamma Z | X, Y, Z] = \beta\gamma(\gamma - \beta)XYZ$.

Proof Using (5.4), (i)–(iv) follow from

$$Cov(H_1, H_2|*) = E(H_1H_2|*) - E(H_1|*)E(H_2|*),$$

$$Cum(H_1, H_2, H_3|*) = E\Big(\prod_{i=1}^{3} H_i|\,*\Big) - E(H_1H_2|*)E(H_3|*)\langle 3\rangle + 2\prod_{i=1}^{3} E(H_i|*)$$

(hereafter, $\langle N\rangle$ means that there are similar N terms obtained under index permutation).

On the other hand, (v)–(vii) are direct consequences of (i)–(vi), since

$$\begin{aligned}
Cov[(\beta \diamond_\gamma X)^2, \beta \diamond'_\gamma Y|X,Y] &= Cum(\beta \diamond_\gamma X, \beta \diamond_\gamma X, \beta \diamond'_\gamma Y|X,Y) \\
&\quad + 2Cov(\beta \diamond_\gamma X, \beta \diamond'_\gamma Y|X,Y)E(\beta \diamond_\gamma X|X,Y), \\
Cov[(\beta \diamond_\gamma X)(\beta \diamond'_\gamma Y), \beta \diamond_\gamma X|X,Y] &= Cum(\beta \diamond_\gamma X, \beta \diamond_\gamma X, \beta \diamond'_\gamma Y|X,Y) \\
&\quad + V(\beta \diamond_\gamma X|X,Y)E(\beta \diamond'_\gamma Y|X,Y) \\
&\quad + Cov(\beta \diamond_\gamma X, \beta \diamond'_\gamma Y|X,Y)E(\beta \diamond_\gamma X|X,Y), \\
Cov[(\beta \diamond_\gamma X)(\beta \diamond'_\gamma Y), \beta \diamond''_\gamma Z|X,Y,Z] &= Cum(\beta \diamond_\gamma X, \beta \diamond'_\gamma Y, \beta \diamond''_\gamma Z|X,Y,Z) \\
&\quad + Cov(\beta \diamond_\gamma X, \beta \diamond''_\gamma Z|X,Y,Z)E(\beta \diamond'_\gamma Y|X,Y,Z) \\
&\quad + Cov(\beta \diamond'_\gamma Y, \beta \diamond''_\gamma Z|X,Y,Z)E(\beta \diamond_\gamma X|X,Y,Z). \quad \square
\end{aligned}$$

5.3 ADCINAR(1) process and higher autocumulant functions

Nastić et al. (2017) defined the ADCINAR(1) process, as follows:

$$Y_t = \alpha \diamond_\vartheta Y_{t-1} + \varepsilon_t, \quad t = 0, \pm 1, \ldots, \tag{5.8}$$

where $\alpha \diamond_\vartheta$ is an alternative generalized binomial thinning operator with $0 \le \alpha \le \vartheta < 1 (\vartheta \ne 1)$, and $\{\varepsilon_t\}$, referred to as an innovation, is a sequence of IID nonnegative integer-valued random variables, such that ε_t and Y_{t-i} are independent for all integer t and positive integer i (it is implicitly assumed that the operations $\alpha \diamond_\vartheta$ at different times are performed mutually independently, which are also independent of $\{Y_t\}$ and $\{\varepsilon_t\}$). Given a recursive definition (5.8), Lemma 5.1(iii) enables us to see that, for any positive integer h,

$$Y_{t+h} \stackrel{d}{=} \alpha^h \diamond_{\vartheta^h} Y_t + \sum_{i=0}^{h-1} \alpha^i \diamond_{\vartheta^i} \varepsilon_{t+h-i}, \tag{5.9}$$

where Y_t is independent of $\sum_{i=0}^{h-1} \alpha^i \diamond_{\vartheta^i} \varepsilon_{t+h-i}$. In what follows, the mean, variance, jth cumulant, and jth raw moment of ε_t are denoted by $\mu_\varepsilon = E(\varepsilon_t)$, $\sigma^2_\varepsilon = V(\varepsilon_t)$, $\kappa_{j,\varepsilon} = Cum_j(\varepsilon_t)$, and $\mu'_{j,\varepsilon} = E(\varepsilon_t^j)$, respectively. Also, I_S stands for the indicator of the set S.

We emphasize that, throughout this thesis, except for some simulation experiments, we do not assume the distributional form about the innovation $\{\varepsilon_t\}$. In this sense, we treat (5.8) to be semi-parametric.

As mentioned in Nastić et al. (2017), the ADCINAR(1) process $\{Y_t\}$ is strictly stationary and ergodic when $0 \le \alpha \le \vartheta < 1 (\vartheta \neq 0)$, whose mean μ_Y and variance σ_Y^2 are given by

$$\mu_Y = \frac{\mu_\varepsilon}{1-\alpha} \quad \text{and} \quad \sigma_Y^2 = \frac{\sigma_\varepsilon^2 + \alpha\mu_\varepsilon + \alpha(\vartheta - \alpha)\Delta(\alpha, \mu_\varepsilon)}{1-\alpha\vartheta}, \tag{5.10}$$

respectively, where

$$\Delta(\alpha, \mu_\varepsilon) = -\frac{\mu_\varepsilon}{1-\alpha} + \left(\frac{\mu_\varepsilon}{1-\alpha}\right)^2,$$

since

$$E(Y_t) = E(\alpha \diamond_\vartheta Y_{t-1}) + E(\varepsilon_t) = \alpha E(Y_{t-1}) + \mu_\varepsilon$$

(use Corollary 5.1(i)) and, by independence between $\alpha \diamond_\vartheta Y_{t-1}$ and ε_t,

$$\begin{aligned} V(Y_t) &= V(\alpha \diamond_\vartheta Y_{t-1}) + V(\varepsilon_t) \\ &= \alpha\vartheta V(Y_{t-1}) + \alpha(1-\vartheta)E(Y_{t-1}) + \alpha(\vartheta - \alpha)[E(Y_{t-1})]^2 + \sigma_\varepsilon^2 \end{aligned}$$

(use Corollary 5.2). Here, when μ_ε exists, μ_Y exists; when $\mu'_{2,\varepsilon}$ exists (in this case, μ_ε and σ_ε^2 exist), σ_Y^2 exists. Further, using (5.9) and Lemma 5.2(ii), the autocovariance function at lag $u \ge 0$ of the stationary ADCINAR(1) process $\{Y_t\}$ is given by

$$\begin{aligned} Cov(Y_t, Y_{t+u}) &= Cov\Big(Y_t, \alpha^u \diamond_{\vartheta^u} Y_t + I_{\{u>0\}} \sum_{i=0}^{u-1} \alpha^i \diamond_{\vartheta^i} \varepsilon_{t+u-i}\Big) \\ &= Cov(Y_t, \alpha^u \diamond_{\vartheta^u} Y_t) = \alpha^u \sigma_Y^2, \end{aligned} \tag{5.11}$$

hence, the autocorrelation function at lag $u (\ge 0)$ of the stationary ADCINAR(1) process $\{Y_t\}$ is given by α^u, as in the usual stationary autoregressive process of the first-order (e.g., Brockwell and Davis (1987)).

Similarly, it is possible to compute the third and fourth cumulants of Y_t (denoted by $\kappa_{j,Y}$), as follows. Noting that

$$Cum_j(Y_t) = Cum_j(\alpha \diamond_\vartheta Y_{t-1}) + Cum_j(\varepsilon_t)$$

by independence between $\alpha \diamond_\vartheta Y_{t-1}$ and ε_t, and using Corollary 5.2, together with strictly stationarity of $\{Y_t\}$, we obtain

$$\begin{aligned} \kappa_{3,Y} &= \frac{1}{1-\alpha\vartheta^2}\Big[\kappa_{3,\varepsilon} + 3\alpha\vartheta(1-\vartheta)\sigma_Y^2 + \alpha(1-\vartheta)(1-2\vartheta)\mu_Y \\ &\qquad + \alpha(\vartheta-\alpha)\Big\{3\vartheta\mu_Y\sigma_Y^2 + 3(1-\vartheta)\mu_Y^2 + (\vartheta - 2\alpha)\mu_Y^3\Big\}\Big], \end{aligned} \tag{5.12}$$

$$\kappa_{4,Y} = \frac{1}{1-\alpha\vartheta^3}\Big[\kappa_{4,\varepsilon} + 6\alpha\vartheta^2(1-\vartheta)\kappa_{3,Y} + \alpha\vartheta(1-\vartheta)(7-11\vartheta)\sigma_Y^2$$

$$
\begin{aligned}
&+\alpha(1-\vartheta)(1-6\vartheta+6\vartheta^2)\mu_Y \\
&+\alpha(\vartheta-\alpha)\Big\{4\vartheta^2\mu_Y\kappa_{3,Y}+3\vartheta^2\sigma_Y^4+18\vartheta(1-\vartheta)\mu_Y\sigma_Y^2 \\
&\qquad+6\vartheta(\vartheta-2\alpha)\mu_Y^2\sigma_Y^2+(1-\vartheta)(7-11\vartheta)\mu_Y^2 \\
&\qquad+6(1-\vartheta)(\vartheta-2\alpha)\mu_Y^3+(\vartheta^2-6\alpha\vartheta+6\alpha^2)\mu_Y^4\Big\}\Big]. \qquad (5.13)
\end{aligned}
$$

It is worth noting that, if $\mu'_{J,\varepsilon}$ exists for $J \geq 3$ (in this case, μ_ε, σ_ε^2, and $\kappa_{j,\varepsilon}$, $j = 3, \ldots, J$, exist), then, $\kappa_{j,Y}$ exists for $j = 3, \ldots, J$.

Remark 5.1. (i) By similar computations of the (unconditional) expectations given in Section 5.2, we see that

$$E(\alpha \diamond_\vartheta Y|Y) = \alpha Y \text{ and } E[(\alpha \diamond_\vartheta Y)^2|Y] = \alpha\{\vartheta Y^2 + (1-\vartheta)Y\},$$

hence,

$$E[(\alpha \diamond_\vartheta Y - \alpha Y)^2|Y] = \alpha\{(\vartheta-\alpha)Y^2 + (1-\vartheta)Y\}.$$

For the stationary ADCINAR(1) process $\{Y_t\}$ (see (5.8)), we have

$$
\begin{aligned}
E(Y_t|Y_{t-1}) &= \alpha Y_{t-1} + \mu_\varepsilon, \\
E[(Y_t - \alpha Y_{t-1} - \mu_\varepsilon)^2|Y_{t-1}] &= E[(\alpha \diamond_\vartheta Y_{t-1} - \alpha Y_{t-1})^2|Y_{t-1}] + E[(\varepsilon_t - \mu_\varepsilon)^2|Y_{t-1}] \\
&= \alpha(\vartheta-\alpha)Y_{t-1}^2 + \alpha(1-\vartheta)Y_{t-1} + \sigma_\varepsilon^2.
\end{aligned}
$$

(ii) If $\vartheta = \alpha (\in (0,1])$, the operator $\alpha\diamond_\vartheta$ under consideration reduces to the binomial thinning operator $\alpha \circ Y$, which means that the INAR(1) process given by $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$ is a special case of the ADCINAR(1) process. The expressions of μ_Y, σ_Y^2, and $\kappa_{j,Y}$, $j = 3, 4$ (see (5.10), (5.12), and (5.13)) are extensions of those in Chapter 2.

For asymptotic theory of the stationary time series, the autocumulant functions (equivalently, the central automoment functions) are fundamental. Recall that, given a stationary process $\{X_t\}$ with mean μ_X, the $(r+1)$th central automoment function (if it exists), where r is a positive integer, is defined by

$$\mu_X(u_1, \ldots, u_r) = E\Big[(X_t - \mu_X)\prod_{i=1}^{r}(X_{t+u_i} - \mu_X)\Big]$$

for nonnegative integers $u_r \geq \cdots \geq u_1 \geq 0$. The second central automoment function is nothing but the autocovariance function $\gamma_X(\cdot)$ and the third central automoment function is equal to the third autocumulant function $\gamma_X(\cdot,\cdot)$, i.e.,

$$\gamma_X(u_1) = \mu_X(u_1) \text{ and } \gamma_X(u_1,u_2) = \mu_X(u_1,u_2).$$

Further, the fourth autocumulant function $\gamma_X(\cdot,\cdot,\cdot)$ is given by

$$\gamma_X(u_1,u_2,u_3) = \mu_X(u_1,u_2,u_3) - \gamma_X(u_1)\gamma_X(u_3-u_2) - \gamma_X(u_2)\gamma_X(u_3-u_1) - \gamma_X(u_3)\gamma_X(u_2-u_1).$$

The following result is an extension of Proposition 2.1 to the ADCINAR(1) case.

Proposition 5.1. *In addition to the autocovariance function $\gamma_Y(u) = \alpha^u\sigma_Y^2$ (see (5.11)), the third and fourth autocumulant functions of the stationary ADCINAR(1) process $\{Y_t\}$ are, respectively, given by, for nonnegative integers $w \ge v \ge u \ge 0$,*

$$\gamma_Y(u,v) = \alpha^v\vartheta^u Q_{2:3,Y} + \alpha^v\sigma_Y^2 + 2\alpha^v\Big[\vartheta^u - \alpha^u + (1-\alpha)\Big(\frac{\vartheta-\vartheta^u}{1-\vartheta} - \frac{\alpha-\alpha^u}{1-\alpha}\Big)\Big]\sigma_Y^2\mu_Y, \tag{5.14}$$

$$\begin{aligned}\gamma_Y(u,v,w) = {}& \alpha^w\vartheta^{u+v}Q_{2:4,Y} + \alpha^w(\vartheta^u + 2\vartheta^v)Q_{2:3,Y} + \alpha^w\sigma_Y^2 \\ & + \alpha^w\Big[\Big\{(\vartheta^u-\alpha^u)\vartheta^v + 2\vartheta^u(\vartheta^v-\alpha^v)\Big\}(Q_{2:3,Y}\mu_Y + \sigma_Y^4) \\ & \quad + \Big\{(\vartheta^u-\alpha^u)(2+\vartheta^{v-u}) + 2\alpha^u(2-\vartheta^u)(\vartheta^{v-u}-\alpha^{v-u})\Big\}\sigma_Y^2\mu_Y \\ & \quad + 3(\vartheta^u-\alpha^u)\Big\{(\vartheta^v - 2\alpha^v)\sigma_Y^2\mu_Y^2 + (\vartheta^v - \vartheta^{v-u})\sigma_Y^2\Big\} \\ & \quad + \delta_Y(u,v)\Big],\end{aligned} \tag{5.15}$$

where $Q_{2:3,Y} = \kappa_{3,Y} - \sigma_Y^2$ and $Q_{2:4,Y} = \kappa_{4,Y} - 3\kappa_{3,Y} + 2\sigma_Y^2$ (when $\vartheta = \alpha (\in (0,1))$, the terms of the square brackets in $\gamma_Y(u,v)$ and $\gamma_Y(u,v,w)$ vanish).

Complicated formula $\delta_Y(u,v)$, depending on $\alpha, \vartheta, \mu_Y, \sigma_Y^2, \kappa_{3,Y}$, and $\kappa_{4,Y}$, is given, as follows:

$$\delta_Y(u,v) = \delta_1(u,v) + 2\delta_2(u,v) + 2\delta_3(u,v) + \delta_4(u,v), \tag{5.16}$$

where

$$\begin{aligned}\delta_1(u,v) = {}& 2(1-\alpha)^2\Big(\vartheta\frac{1-\vartheta^{v-u}}{1-\vartheta} - \alpha\frac{1-\alpha^{v-u}}{1-\alpha}\Big)\Big(\frac{\vartheta-\vartheta^u}{1-\vartheta} - \frac{\alpha-\alpha^u}{1-\alpha}\Big)\sigma_Y^2\mu_Y^2 \\ & + (1-\alpha)^2\Big[\Big(\vartheta\frac{\vartheta^{v-u}}{1-\vartheta} - \alpha\frac{\alpha^{v-u}}{1-\alpha}\Big)\Big(\frac{\vartheta^2-\vartheta^u}{1-\vartheta} - \frac{\alpha^2-\alpha^u}{1-\alpha}\Big) \\ & \quad + \frac{\vartheta^{v-u}}{1-\vartheta}\Big(\frac{\alpha^2\vartheta^2-\alpha^u\vartheta^u}{1-\alpha\vartheta} - \frac{\vartheta^4-\vartheta^{2u}}{1-\vartheta^2}\Big) + \frac{\alpha^{v-u}}{1-\alpha}\Big(\frac{\alpha^2\vartheta^2-\alpha^u\vartheta^u}{1-\alpha\vartheta} - \frac{\alpha^4-\alpha^{2u}}{1-\alpha^2}\Big) \\ & \quad + \alpha^{v-u}\Big\{\frac{\alpha^2-\alpha^u}{1-\alpha}\Big(\frac{\alpha}{1-\alpha} - \frac{\vartheta}{1-\vartheta}\Big) + \frac{1}{1-\vartheta}\frac{\alpha^2\vartheta^2-\alpha^u\vartheta^u}{1-\alpha\vartheta} - \frac{1}{1-\alpha}\frac{\alpha^4-\alpha^{2u}}{1-\alpha^2}\Big\}\Big]\sigma_Y^2\mu_Y^2 \\ & + \vartheta^{v-u}\Big(\frac{\vartheta^2-\vartheta^{2u}}{1-\vartheta^2} - \frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta}\Big)\{(1-\alpha\vartheta)\sigma_Y^4 - \alpha(1-\vartheta)\sigma_Y^2\mu_Y - \alpha(\vartheta-\alpha)\sigma_Y^2\mu_Y^2\} \\ & + (1-\alpha)\Big\{\frac{\vartheta-\vartheta^u}{1-\vartheta} - \frac{\alpha-\alpha^u}{1-\alpha} + \vartheta^{v-u}\Big(\frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta} - \frac{\vartheta^2-\vartheta^{2u}}{1-\vartheta^2}\Big)\Big\}\sigma_Y^2\mu_Y\end{aligned}$$

$$
\begin{aligned}
&+(1-\alpha)^2\Big\{\vartheta^{v-u}\Big(\frac{\vartheta^2-\vartheta^{2u}}{1-\vartheta^2}-\frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta}\Big)+2\alpha^{v-u}\Big(\frac{\alpha^2-\alpha^{2u}}{1-\alpha^2}-\frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta}\Big)\Big\}\sigma_Y^2\mu_Y^2\\
&+(1-\alpha)^2\Big\{\Big(\vartheta^{v-u}\frac{\vartheta-\vartheta^{u-1}}{1-\vartheta}-\alpha^{v-u}\frac{\alpha-\alpha^{u-1}}{1-\alpha}\Big)\Big(\frac{\alpha^u}{1-\alpha}-\frac{\vartheta^u}{1-\vartheta}\Big)+\frac{\alpha^v}{1-\alpha}\Big(\frac{\vartheta-\vartheta^{u-1}}{1-\vartheta}-\frac{\alpha-\alpha^{u-1}}{1-\alpha}\Big)\\
&\qquad+\frac{\vartheta}{1-\vartheta}\Big(\vartheta^{v-u}\frac{\vartheta^2-\vartheta^{2(u-1)}}{1-\vartheta^2}-\alpha^{v-u}\frac{\alpha\vartheta-\alpha^{u-1}\vartheta^{u-1}}{1-\alpha\vartheta}\Big)\\
&\qquad+\frac{\alpha}{1-\alpha}\Big(\alpha^{v-u}\frac{\alpha^2-\alpha^{2(u-1)}}{1-\alpha^2}-\vartheta^{v-u}\frac{\alpha\vartheta-\alpha^{u-1}\vartheta^{u-1}}{1-\alpha\vartheta}\Big)\\
&\qquad+\frac{\alpha^{v-u+1}}{1-\alpha}\Big(\frac{\alpha^2-\alpha^{2(u-1)}}{1-\alpha^2}-\frac{\alpha\vartheta-\alpha^{u-1}\vartheta^{u-1}}{1-\alpha\vartheta}\Big)\Big\}\sigma_Y^2\mu_Y^2,
\end{aligned}
$$

$$
\begin{aligned}
\delta_2(u,v)&=(\vartheta^{v-u}-\alpha^{v-u})\{(1-\alpha\vartheta)\sigma_Y^4-\alpha(1-\vartheta)\sigma_Y^2\mu_Y-\alpha(\vartheta-\alpha)\sigma_Y^2\mu_Y^2\}\\
&+(1-\alpha)^2\Big(\vartheta\frac{1-\vartheta^{v-u}}{1-\vartheta}-\alpha\frac{1-\alpha^{v-u}}{1-\alpha}\Big)\Big(\frac{\vartheta-\vartheta^u}{1-\vartheta}-\frac{\alpha-\alpha^u}{1-\alpha}\Big)\sigma_Y^2\mu_Y^2\\
&+(1-\alpha)^2\Big[\Big(\vartheta\frac{\vartheta^{v-u}}{1-\vartheta}-\alpha\frac{\alpha^{v-u}}{1-\alpha}\Big)\Big(\frac{\vartheta^2-\vartheta^u}{1-\vartheta}-\frac{\alpha^2-\alpha^u}{1-\alpha}\Big)\\
&\qquad+\frac{\vartheta^{v-u}}{1-\vartheta}\Big(\frac{\alpha^2\vartheta^2-\alpha^u\vartheta^u}{1-\alpha\vartheta}-\frac{\vartheta^4-\vartheta^{2u}}{1-\vartheta^2}\Big)+\frac{\alpha^{v-u}}{1-\alpha}\Big(\frac{\alpha^2\vartheta^2-\alpha^u\vartheta^u}{1-\alpha\vartheta}-\frac{\alpha^4-\alpha^{2u}}{1-\alpha^2}\Big)\\
&\qquad+\alpha^{v-u}\Big\{\frac{\alpha^2-\alpha^u}{1-\alpha}\Big(\frac{\alpha}{1-\alpha}-\frac{\vartheta}{1-\vartheta}\Big)+\frac{1}{1-\vartheta}\frac{\alpha^2\vartheta^2-\alpha^u\vartheta^u}{1-\alpha\vartheta}-\frac{1}{1-\alpha}\frac{\alpha^4-\alpha^{2u}}{1-\alpha^2}\Big\}\Big]\sigma_Y^2\mu_Y^2\\
&+\Big(\vartheta^{v-u}\frac{\vartheta^2-\vartheta^{2u}}{1-\vartheta^2}-\alpha^{v-u}\frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta}\Big)\{(1-\alpha\vartheta)\sigma_Y^4-\alpha(1-\vartheta)\sigma_Y^2\mu_Y-\alpha(\vartheta-\alpha)\sigma_Y^2\mu_Y^2\}\\
&+(1-\alpha)\Big\{\vartheta^{v-u}\Big(\frac{\vartheta-\vartheta^u}{1-\vartheta}-\frac{\vartheta^2-\vartheta^{2u}}{1-\vartheta^2}\Big)+\alpha^{v-u}\Big(\frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta}-\frac{\alpha-\alpha^u}{1-\alpha}\Big)\Big\}\sigma_Y^2\mu_Y\\
&+(1-\alpha)^2\Big\{\vartheta^{v-u}\Big(\frac{\vartheta^2-\vartheta^{2u}}{1-\vartheta^2}-\frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta}\Big)+2\alpha^{v-u}\Big(\frac{\alpha^2-\alpha^{2u}}{1-\alpha^2}-\frac{\alpha\vartheta-\alpha^u\vartheta^u}{1-\alpha\vartheta}\Big)\Big\}\sigma_Y^2\mu_Y^2\\
&+(1-\alpha)^2\Big\{\frac{\vartheta^{v-u+1}}{1-\vartheta}\Big(\frac{\vartheta^2-\vartheta^{2(u-1)}}{1-\vartheta^2}-\frac{\alpha\vartheta-\alpha^{u-1}\vartheta^{u-1}}{1-\alpha\vartheta}\Big)\\
&\qquad+\frac{2\alpha^{v-u+1}}{1-\alpha}\Big(\frac{\alpha^2-\alpha^{2(u-1)}}{1-\alpha^2}-\frac{\alpha\vartheta-\alpha^{u-1}\vartheta^{u-1}}{1-\alpha\vartheta}\Big)\\
&\qquad+\Big(\frac{2\alpha^v}{1-\alpha}-\frac{\vartheta^v}{1-\vartheta}\Big)\Big(\frac{\vartheta-\vartheta^{u-1}}{1-\vartheta}-\frac{\alpha-\alpha^{u-1}}{1-\alpha}\Big)\Big\}\sigma_Y^2\mu_Y^2,
\end{aligned}
$$

$$
\begin{aligned}
\delta_3(u,v)&=(1-\alpha)\Big[\Big(\frac{\vartheta-\vartheta^v}{1-\vartheta}-\frac{\alpha-\alpha^v}{1-\alpha}\Big)\{\vartheta^uQ_{2:3,Y}+\sigma_Y^2+2(\vartheta^u-\alpha^u)\sigma_Y^2\mu_Y\}\mu_Y\\
&\qquad+2\alpha^v\Big(\frac{\alpha-\alpha^u}{1-\alpha}-\frac{\vartheta-\vartheta^u}{1-\vartheta}\Big)\sigma_Y^2\mu_Y^2\Big],\\
\delta_4(u,v)&=(1-\alpha)\Big(\frac{\vartheta-\vartheta^u}{1-\vartheta}-\frac{\alpha-\alpha^u}{1-\alpha}\Big)\{\vartheta^vQ_{2:3,Y}+\sigma_Y^2+2(\vartheta^v-2\alpha^v)\sigma_Y^2\mu_Y\}\mu_Y.
\end{aligned}
$$

Remark 5.2. Note that $\delta_j(u,v) = c_{1j}(u,v)\sigma_Y^4 + c_{2j}(u,v)\sigma_Y^2\mu_Y + c_{3j}(u,v)\sigma_Y^2\mu_Y^2$, $j = 1,2$, where $c_{ij}(u,v)$'s heavily depend on α and ϑ. However, the re-arrangement of (5.16) is not made here, since it would not be practically important.

Remark 5.3. (i) Especially, we have, for $w \ge v \ge 1$,

$$
\begin{aligned}
\gamma_Y(0,v,w)=\alpha^w\Big[&\vartheta^vQ_{2:4,Y}+(1+2\vartheta^v)Q_{2:3,Y}+\sigma_Y^2\\
&+2(\vartheta^v-\alpha^v)(Q_{2:3,Y}\mu_Y+\sigma_Y^2\mu_Y+\sigma_Y^4)\\
&+2(1-\alpha)\Big(\frac{\vartheta-\vartheta^v}{1-\vartheta}-\frac{\alpha-\alpha^v}{1-\alpha}\Big)(Q_{2:3,Y}\mu_Y+\sigma_Y^2\mu_Y)\Big],
\end{aligned}
$$

$$
\begin{aligned}
\gamma_Y(1,v,w) = \alpha^w \Big[& \vartheta^{v+1} Q_{2:4,Y} + (\vartheta + 2\vartheta^v) Q_{2:3,Y} + \sigma_Y^2 \\
& + \{(\vartheta-\alpha)\vartheta^v + 2\vartheta(\vartheta^v - \alpha^v)\}(Q_{2:3,Y}\mu_Y + \sigma_Y^4) \\
& + \{(\vartheta-\alpha)(2+\vartheta^{v-1}) + 2\alpha(2-\vartheta)(\vartheta^{v-1} - \alpha^{v-1})\}\sigma_Y^2\mu_Y \\
& + 3(\vartheta-\alpha)\{(\vartheta^v - 2\alpha^v)\sigma_Y^2\mu_Y^2 - \vartheta^{v-1}(1-\vartheta)\sigma_Y^2\} \\
& + 2(\vartheta^{v-1} - \alpha^{v-1})\{(1-\alpha\vartheta)\sigma_Y^4 - \alpha(1-\vartheta)\sigma_Y^2\mu_Y - \alpha(\vartheta-\alpha)\sigma_Y^2\mu_Y^2\} \\
& + 2(1-\alpha)\Big(\frac{\vartheta-\vartheta^v}{1-\vartheta} - \frac{\alpha-\alpha^v}{1-\alpha}\Big)\{\vartheta Q_{2:3,Y}\mu_Y + \sigma_Y^2\mu_Y + 2(\vartheta-\alpha)\sigma_Y^2\mu_Y^2\}\Big].
\end{aligned}
$$

(ii) Subsection 6.3.2 below needs

$$
\begin{aligned}
\mu_Y(1,k,k) - \alpha\mu_Y(0,k,k) = \alpha^k \Big[& 2(1-\alpha)\alpha^k Q_{2:3,Y} + (1-\alpha)\sigma_Y^2 + 2(1-\alpha^2)\alpha^{k-1}\sigma_Y^4 \\
& + (\vartheta-\alpha)\vartheta^k Q_{2:4,Y} + \{\vartheta - \alpha + 2(1-\alpha)(\vartheta^k - \alpha^k)\}Q_{2:3,Y} \\
& + \frac{\vartheta-\alpha}{1-\vartheta}\{(1+2\alpha-3\vartheta)\vartheta^k + 2(\vartheta-\alpha)\}Q_{2:3,Y}\mu_Y \\
& + \frac{\vartheta-\alpha}{1-\vartheta}\{(2\alpha\vartheta - 1 - \vartheta)\vartheta^{k-1} + 2(2-\alpha-\vartheta)\}\sigma_Y^2\mu_Y \\
& + \frac{\vartheta-\alpha}{1-\vartheta}\{(6\alpha\vartheta - 2\alpha - \vartheta - 3\vartheta^2)\vartheta^{k-1} + 4(\vartheta-\alpha)\}\sigma_Y^2\mu_Y^2 \\
& + 3(\vartheta-\alpha)(\vartheta-1)\vartheta^{k-1}\sigma_Y^2 \\
& + \{(\vartheta-\alpha)(3\vartheta-2\alpha)\vartheta^{k-1} + 2(1-\alpha^2)(\vartheta^{k-1} - \alpha^{k-1})\}\sigma_Y^4\Big],
\end{aligned}
$$

where $k = 1, 2, \ldots$.

By Proposition 5.1, it is easy to understand that, for $r = 1, 2, 3$, the $(r+1)$th autocumulant function of the stationary ADCINAR(1) process is proportional to α^{u_r} when $u_r \geq \cdots \geq u_1 \geq 0$. Note that this fact is correct up to $r = 5$, for the stationary INAR(1) process under a general innovation (see Chapter 2). The following result proves this structure about higher autocumulant functions of the stationary ADCINAR(1) process (the proof is postponed to Section 5.4).

Proposition 5.2. *Given a positive integer r, suppose that ε_t has the $(r+1)$th moment. Consider the stationary ADCINAR(1) process $\{Y_t\}$, given by (5.8). Then, for any integer t and nonnegative integers $u_r \geq \cdots \geq u_1 \geq 0$, there exists a constant $C_r > 0$ (independent of t and $u_1, \ldots, u_r$) such that*

$$
|Cum(Y_t, Y_{t+u_1}, \ldots, Y_{t+u_r})| \leq C_r \alpha^{u_r}.
$$

Remark 5.4. For $r = 4, 5, \ldots$, the closed-form expression of the $(r+1)$th autocumulant function would be unrealistically not available for the stationary ADCINAR(1) process.

Suppose that the observation $\{Y_1, \ldots, Y_n\}$ of length n is generated by (5.8). Define the sample mean, variance, and cumulants by $\overline{Y} = (1/n)\sum_{t=1}^n Y_t$, $\widehat{\sigma}_Y^2 = (1/n)\sum_{t=1}^n (Y_t - \overline{Y})^2$, $\widehat{\kappa}_{3,Y} = (1/n)\sum_{t=1}^n (Y_t - \overline{Y})^3$, and $\widehat{\kappa}_{4,Y} = (1/n)\sum_{t=1}^n (Y_t - \overline{Y})^4 - 3\widehat{\sigma}_Y^4$, respectively

We prepare two lemmas, which will be used in the next chapter.

Lemma 5.5. *Suppose that* $0 \le \alpha \le \vartheta < 1 (\vartheta \neq 0)$. *The following hold.*

(i) If $E(\varepsilon_t^4)$ *exists, then,* $\overline{Y} \xrightarrow{a.s.} \mu_Y$, $\widehat{\sigma}_Y^2 \xrightarrow{a.s.} \sigma_Y^2$, *and* $\widehat{\kappa}_{j,Y} \xrightarrow{a.s.} \kappa_{j,Y}$, $j = 3, 4$.

(ii) Let J *be a positive integer. If* $E(\varepsilon_t^{J+1})$ *exists, then,*

$$\frac{1}{n}\sum_{t=1}^{n-u_\ell}(Y_t - \overline{Y})\prod_{j=1}^{\ell}(Y_{t+u_j} - \overline{Y}) \xrightarrow{a.s.} \mu_Y(u_1, \ldots, u_\ell)$$

for fixed nonnegative integers $0 \le u_1 \le \cdots \le u_\ell (\ell = 1, 2, \ldots, J)$.

Proof For $i = 1, 2, 3, 4$, $(1/n)\sum_{t=1}^{n} Y_t^i \xrightarrow{a.s.} E(Y_1^i)$ by strictly stationarity and ergodicity of $\{Y_t\}$. Similarly, for $\ell = 1, 2, \ldots, J$, $(1/n)\sum_{t=1}^{n-u_\ell} Y_t \prod_{j=1}^{\ell} Y_{t+u_j} \xrightarrow{a.s.} E(Y_1 \prod_{j=1}^{\ell} Y_{1+u_j})$. We thus have (i) and (ii). □

Lemma 5.6. *Let* $Z_t = Y_t - \mu_Y$. *Suppose that* $0 \le \alpha \le \vartheta < 1 (\vartheta \neq 0)$ *and that* $E(\varepsilon_t^8)$ *exists. Then,*

$$\frac{1}{\sqrt{n}}\sum_{t=2}^{n}[Z_{t-1}^i Z_t^j - E(Z_{t-1}^i Z_t^j)] = O_p(1), \quad 0 \le i + j \le 4, j = 0, 1((i, j) \neq (0, 0)).$$

Proof According to Proposition 5.2, we see that $|Cov(Z_{s-1}^i Z_s^j, Z_{t-1}^i Z_t^j)| \le C_{ij}\alpha^{|t-s|}$ for some constant $C_{ij} > 0$, independent of s and t, if $0 \le i + j \le 4, j = 0, 1((i, j) \neq (0, 0))$, using McCullagh's (2018) table of complementary set partitions ($\tau = 1|2$, $\tau = 12|34$, $\tau = 123|456$, and $\tau = 1234|5678$). The result follows from

$$V\Big(\frac{1}{\sqrt{n}}\sum_{t=2}^{n} Z_{t-1}^i Z_t^j\Big) \le \frac{1}{n}\Big\{\sum_{t=2}^{n} V(Z_{t-1}^i Z_t^j) + 2\sum_{s,t=2}^{n} I_{\{s<t\}}|Cov(Z_{s-1}^i Z_s^j, Z_{t-1}^i Z_t^j)|\Big\} = O(1). \quad \square$$

5.4 Proofs of Propositions 5.1 and 5.2

We emphasize that, for the proofs of Propositions 5.1 and 5.2, the time at which the operation is performed should be indicated, in such a way that $Y_t = \alpha \diamond_\vartheta^{(t)} Y_{t-1} + \varepsilon_t$, $t = 0, \pm 1, \ldots$, where

$$\alpha \diamond_\vartheta^{(t)} Y_{t-1} = \begin{cases} 0, & Y_{t-1} = 0, \\ \sum_{j=1}^{Y_{t-1}} S_j^{(t)}(\alpha, \vartheta), & Y_{t-1} = 1, 2, \ldots, \end{cases}$$

with $S_j^{(t)}(\alpha, \vartheta) = B_j^{(t)}(\vartheta) B^{(t)}(\alpha/\vartheta)$. Here, $\{B_j^{(t)}(\vartheta)\}$ is a sequence of IID Bernoulli random variables at time t, which is independent of Bernoulli random variable $B^{(t)}(\alpha/\vartheta)$ at time t (these generating random variables at different times are mutually independent); besides, $\{B_j^{(t)}(\vartheta)\}$ and $\{B^{(t)}(\alpha/\vartheta)\}$ are mutually independent of $\{Y_t\}$ and $\{\varepsilon_t\}$.

Additionally, we must give a more precise expression for Y_{t+h} (rather than (5.9)) involving the operations $\alpha \diamond_\vartheta^{(t)}$ at different times. For this purpose, it is natural to introduce the ℓ-times repeated operation, as follows. For any integers i and j, let $\binom{j}{i} = \{i, \ldots, j\}$ if $i \le j$, and $\binom{j}{i} = \emptyset$ if $i > j$. If we consider all operations at times from $t+1$ to $t+\ell$, given by $\binom{t+\ell}{t+1}$ (ℓ is a positive integer), where t is an arbitrary integer, we define

$$\alpha\binom{t+\ell}{t+1} \diamond_{\vartheta} Y_t = (\underbrace{\alpha \diamond_{\vartheta}^{(t+\ell)} \cdots \alpha \diamond_{\vartheta}^{(t+2)} \alpha \diamond_{\vartheta}^{(t+1)}}_{\ell \text{ times}}) Y_t = \begin{cases} 0, & Y_t = 0, \\ \sum_{j=1}^{Y_t} \prod_{\nu=t+1}^{t+\ell} S_j^{(\nu)}(\alpha, \vartheta), & Y_t = 1, 2, \ldots, \end{cases} \tag{5.17}$$

for which we can use the fact $\alpha\binom{t+\ell}{t+1} \diamond_{\vartheta} Y_t \overset{d}{=} \alpha^{\ell} \diamond_{\vartheta^{\ell}} Y_t$; See Lemma 5.1(iii), where $|\binom{t+\ell}{t+1}| = \ell$ corresponds to the length of times from $t+1$ to $t+\ell$. It is convenient for us to use the conventional notation $\alpha(\emptyset)\diamond_{\vartheta} = 1\diamond_1$ (in this case, let $|\emptyset| = 0$).

Remark 5.5. The meaning of the repeated operation (5.17) is clear from the birth-and-death interpretation. That is, suppose that there are Y_t objects (parents) at time t, and that every descendant of the jth object produces $S_j^{(\nu)}(\alpha, \vartheta)$ objects randomly for any time $\nu(> t)$. Then, the number of the descendant at time $t+\ell$, starting at time t, is given by (5.17).

We can prove inductively that, for any positive integer h,

$$Y_{t+h} = \alpha\binom{t+h}{t+1} \diamond_{\vartheta} Y_t + \sum_{i=0}^{h-1} \alpha_{h-i}\binom{t+h}{t+h-i+1} \diamond_{\vartheta} \varepsilon_{t+h-i}, \quad t = 0, \pm 1, \ldots \tag{5.18}$$

(with $\alpha_h\binom{t+h}{t+h+1}\diamond_{\vartheta} = 1\diamond_1$), once we define, for $i = 1, \ldots, h-1$ (when $h > 1$),

$$\alpha_{h-i}\binom{t+h}{t+h-i+1} \diamond_{\vartheta} \varepsilon_{t+h-i} = \begin{cases} 0, & \varepsilon_{t+h-i} = 0, \\ \sum_{j=1}^{\varepsilon_{t+h-i}} \prod_{\nu=t+h-i+1}^{t+h} S_{j,h-i}^{(\nu)}(\alpha, \vartheta), & \varepsilon_{t+h-i} = 1, 2, \ldots, \end{cases} \tag{5.19}$$

where

$$S_{j,h-i}^{(\nu)}(\alpha, \vartheta) = B_{j,h-i}^{(\nu)}(\vartheta) B^{(\nu)}(\alpha/\vartheta),$$

with $B_{j,h-i}^{(\nu)}(\vartheta)$'s being independent copies of $B_1(\vartheta)$. Such a notation is rather cumbersome, but there will be no confusion, since the time at which the operation is performed and the independence of the Bernoulli random variables $B_{j,h-i}^{(\nu)}(\vartheta)$'s (over Y_t and ε_{t+h-i}'s) are indicated by superscript of parenthesized index and subscript "$h-i$", respectively.

Remark 5.6. In order to make the expression (5.18) clear, we regard the ADCINAR(1) process as a birth-and-death process with immigration, in which the innovation $\{\varepsilon_s\}$ is a sequence of the numbers of the immigration (at time s). In this setting, if the counting is started at time t where there are Y_t objects (parents), and furthermore, if, at each time $\nu(> t)$, there are possibly immigrations which can produce the objects randomly, then, (5.18) denotes the number of the descendant at time $t+h$. Note that (5.19) denotes the number of the descendant at time $t+h$, produced from the immigration at time $t+h-i$.

Remark 5.7. We mention some useful properties of the repeated operation $\alpha\binom{s+h}{s+1}\diamond_\vartheta$, as follows (similar properties hold for $\alpha_{h-i}\binom{t+h}{t+h-i+1}\diamond_\vartheta$). By definition,

$$\alpha\binom{s+h}{s+1}\diamond_\vartheta Y_s = \underbrace{(\alpha\binom{s+h}{s+h_{m-1}+1}\diamond_\vartheta \cdots \alpha\binom{s+h_2}{s+h_1+1}\diamond_\vartheta \alpha\binom{s+h_1}{s+1}\diamond_\vartheta)}_{m \text{ times}}Y_s$$

for a given decomposition (m-disjoint sets of times); $\binom{s+h}{s+1} = \binom{s+h}{s+h_{m-1}+1}\cup\cdots\cup\binom{s+h_2}{s+h_1+1}\cup\binom{s+h_1}{s+1}$, where the repeated operations at disjoint sets of times are performed independently, which are mutually independent of $\{Y_t\}$ and $\{\varepsilon_t\}$. Also, commutative and associative properties hold. For instance,

$$\alpha\binom{s+3}{s+1}\diamond_\vartheta Y_s = (\alpha\binom{s+3}{s+3}\diamond_\vartheta \alpha\binom{s+2}{s+1}\diamond_\vartheta)Y_s = (\alpha\binom{s+2}{s+1}\diamond_\vartheta \alpha\binom{s+3}{s+3}\diamond_\vartheta)Y_s$$

and

$$\alpha\binom{s+3}{s+1}\diamond_\vartheta Y_s = ((\alpha\binom{s+3}{s+3}\diamond_\vartheta \alpha\binom{s+2}{s+2}\diamond_\vartheta)\alpha\binom{s+1}{s+1}\diamond_\vartheta)Y_s = (\alpha\binom{s+3}{s+3}\diamond_\vartheta (\alpha\binom{s+2}{s+2}\diamond_\vartheta \alpha\binom{s+1}{s+1}\diamond_\vartheta))Y_s.$$

Outline[1] of the proof of Proposition 5.1. By Lemma 5.3, we can derive

$$Cum(Y_t, \alpha^{u_1}\diamond_{\vartheta^{u_1}} Y_t, \ldots, \alpha^{u_\ell}\diamond_{\vartheta^{u_\ell}} Y_t) = C_{Y,1}(u_1,\ldots,u_\ell) \quad \text{(say), independent of } t,$$

where $\ell = 2, 3$. It suffices to compute

$$Cum(Y_t, Y_{t+u}, Y_{t+v}) - C_{Y,1}(u,v) = \Delta_Y(u,v;t) \quad \text{(say)},$$
$$Cum(Y_t, Y_{t+u}, Y_{t+v}, Y_{t+w}) - C_{Y,1}(u,v,w) = \Delta_Y(u,v,w;t) \quad \text{(say)}.$$

Note that they obviously vanish for the INAR(1) process. Below, $w \geq v \geq u \geq 1$ is assumed.

Using the expression (5.18), we start with

$$\begin{aligned}\Delta_Y(u,v;t) &= Cum\Big(Y_t, \sum_{i=0}^{u-1}\alpha^i\diamond_{\vartheta^i}\varepsilon_{t+u-i}, \sum_{j=0}^{v-1}\alpha^j\diamond_{\vartheta^j}\varepsilon_{t+v-j}\Big)\\ &\quad + Cum\Big\{Y_t, \alpha\binom{t+u}{t+1}\diamond_\vartheta Y_t, I_{\{v>1\}}\sum_{j=1}^{v-1}\alpha_{v-j}\binom{t+v}{t+v-j+1}\diamond_\vartheta \varepsilon_{t+v-j} + \varepsilon_{t+v}\Big\}\\ &\quad + Cum\Big\{Y_t, I_{\{u>1\}}\sum_{i=1}^{u-1}\alpha_{u-i}\binom{t+u}{t+u-i+1}\diamond_\vartheta \varepsilon_{t+u-i} + \varepsilon_{t+u}, \alpha\binom{t+v}{t+1}\diamond_\vartheta Y_t\Big\}\\ &= C_{Y,2}(u,v;t) + C_{Y,3}(u,v;t) + C_{Y,4}(u,v;t) \quad \text{(say)}.\end{aligned}$$

[1]The complete proof, being lengthy and technical, is available: Zeng, X. and Kakizawa, Y. (2022). Higher autocumulant functions for ADCINAR(1) process and bias-correction of some estimators. *Discussion Paper (Series A):* No.367.

By independence between Y_t and ε_{t+i} $(i > 0)$, it is obvious that $C_{Y,2}(u, v; t) \equiv 0$ and that some terms of $C_{Y,3}(u, v; t)$ and $C_{Y,4}(u, v; t)$ are zero, i.e.,

$$C_{Y,3}(u, v; t) = I_{\{v>1\}} I_{\{u>1\}} \sum_{j=v-u+1}^{v-1} Cum\{Y_t, \alpha\binom{t+u}{t+1} \diamond_\vartheta Y_t, \alpha_{v-j}\binom{t+v}{t+v-j+1} \diamond_\vartheta \varepsilon_{t+v-j}\}, \tag{5.20}$$

$$C_{Y,4}(u, v; t) = I_{\{u>1\}} \sum_{i=1}^{u-1} Cum\{Y_t, \alpha_{u-i}\binom{t+u}{t+u-i+1} \diamond_\vartheta \varepsilon_{t+u-i}, \alpha\binom{t+v}{t+1} \diamond_\vartheta Y_t\}. \tag{5.21}$$

Now, the repeated operations can be treated, in the following way:

- with regard to (5.20) when $u, v > 1$, for each $j = v - u + 1, \ldots, v - 1$,

$$\alpha\binom{t+u}{t+1} \diamond_\vartheta Y_t = \alpha\left[\binom{t+v-j}{t+1} \cup \binom{t+u}{t+v-j+1}\right] \diamond_\vartheta Y_t,$$
$$\alpha_{v-j}\binom{t+v}{t+v-j+1} \diamond_\vartheta \varepsilon_{t+v-j} = \alpha_{v-j}\left[\binom{t+u}{t+v-j+1} \cup \binom{t+v}{t+u+1}\right] \diamond_\vartheta \varepsilon_{t+v-j};$$

- with regard to (5.21) when $u > 1$, for each $i = 1, \ldots, u - 1$,

$$\alpha\binom{t+v}{t+1} \diamond_\vartheta Y_t = \alpha\left[\binom{t+u-i}{t+1} \cup \binom{t+u}{t+u-i+1} \cup \binom{t+v}{t+u+1}\right] \diamond_\vartheta Y_t,$$
$$\alpha_{u-i}\binom{t+u}{t+u-i+1} \diamond_\vartheta \varepsilon_{t+u-i} = \alpha_{u-i}\binom{t+u}{t+u-i+1} \diamond_\vartheta \varepsilon_{t+u-i}.$$

Applying the law of total cumulant (e.g., McCullagh (2018)) and Lemma 5.4(i), it is shown that

$$C_{Y,3}(u, v; t) = C_{Y,4}(u, v; t) = \alpha^v \Big(\frac{\vartheta - \vartheta^u}{1 - \vartheta} - \frac{\alpha - \alpha^u}{1 - \alpha}\Big)\sigma_Y^2 \mu_\varepsilon,$$

hence,

$$\Delta_Y(u, v; t) = 2\alpha^v (1 - \alpha)\Big(\frac{\vartheta - \vartheta^u}{1 - \vartheta} - \frac{\alpha - \alpha^u}{1 - \alpha}\Big)\sigma_Y^2 \mu_Y, \quad \text{independent of } t.$$

Similarly, applying the law of total cumulant and Lemma 5.4, a heavy algebra shows that

$$\Delta_Y(u, v, w; t) = \alpha^w \delta_Y(u, v), \quad \text{independent of } t. \quad \square$$

Proof of Proposition 5.2 In what follows, given a nonnegative integer h, we introduce $Y_{t,h} = \alpha\binom{t+h}{t+1} \diamond_\vartheta Y_t$ and $\varepsilon_{t,h} = I_{\{h>0\}} \sum_{i=0}^{h-1} \varepsilon_{t,h,i}$, where $\varepsilon_{t,h,i} = \alpha_{h-i}\binom{t+h}{t+h-i+1} \diamond_\vartheta \varepsilon_{t+h-i}$; see (5.18). Also, let $\mathcal{T}_{t,h} = \binom{t+h}{t+1}$ (or $\mathcal{T}_{t,h,i} = \binom{t+h}{t+h-i+1}$) be a set of times at which the repeated operation $\alpha\binom{t+h}{t+1}\diamond_\vartheta$ (or $\alpha_{h-i}\binom{t+h}{t+h-i+1}\diamond_\vartheta$) is performed. Throughout this proof, $C_{\#}^{[\cdot]}$'s are positive constants, independent of t and $u_1, \ldots, u_r$.

To make the proof clearer, we first deal with $r = 2, 3$. Below, we assume that $u_r \geq \cdots \geq u_1 \geq 1$.

Now, for the case $r = 2$,

$$Cum(Y_t, Y_{t+u_1}, Y_{t+u_2}) = Cum(Y_t, \varepsilon_{t,u_1}, \varepsilon_{t,u_2}) + Cum(Y_t, Y_{t,u_1}, \varepsilon_{t,u_2}) + Cum(Y_t, Y_{t,u_1} + \varepsilon_{t,u_1}, Y_{t,u_2})$$
$$= D_0^{[2]} + D_1^{[2]} + D_2^{[2]} \quad \text{(say), with } D_0^{[2]} \equiv 0.$$

To deal with the terms $D_2^{[2]}$ and $D_1^{[2]}$ in turn, we apply

$$Cum(X_0, X_1, X_2) = E(X_0X_1X_2) - E(X_0X_1)E(X_2)\langle 3\rangle + 2E(X_0)E(X_1)E(X_2).$$

The repeated use of the conditioning argument, in relation to (5.4), shows that, given I, I_1, I_2, whose values are 0 or 1, with $I + I_1 + I_2 \geq 1$, $|E[Y_t^I(Y_{t,u_1} + \varepsilon_{t,u_1})^{I_1}Y_{t,u_2}^{I_2}]| \leq C_2^{[2]}\alpha^{I_2u_2}$, which yields

$$|D_2^{[2]}| \leq [C_2^{[2]} + 3(C_2^{[2]})^2 + 2(C_2^{[2]})^3]\alpha^{u_2}.$$

On the other hand, we notice that

$$D_1^{[2]} = \sum_{i_2=u_2-u_1+1}^{u_2-1} Cum(Y_t, Y_{t,u_1}, \varepsilon_{t,u_2,i_2}).$$

As long as $u_2 - u_1 + 1 \leq i_2$ $(\leq u_2 - 1)$, $|E(Y_t^IY_{t,u_1}^{I_1}\varepsilon_{t,u_2,i_2}^{I_2})| \leq C_1^{[2]}\alpha^{I_1u_1+I_2(u_2-u_1)}\vartheta^{I_2(i_2+u_1-u_2)}$, which enables us to see that

$$|D_1^{[2]}| \leq [C_1^{[2]} + 3(C_1^{[2]})^2 + 2(C_1^{[2]})^3] \sum_{i_2=u_2-u_1+1}^{u_2-1} \alpha^{u_2}\vartheta^{i_2+u_1-u_2} \leq [C_1^{[2]} + 3(C_1^{[2]})^2 + 2(C_1^{[2]})^3]\frac{\alpha^{u_2}}{1-\vartheta}.$$

This completes the proof for the case $r = 2$.

When $r = 3$, it is easy to see that

$$Cum(Y_t, Y_{t+u_1}, Y_{t+u_2}, Y_{t+u_3})$$
$$= Cum(Y_t, \varepsilon_{t,u_1}, \varepsilon_{t,u_2}, \varepsilon_{t,u_3}) + Cum(Y_t, Y_{t,u_1}, \varepsilon_{t,u_2}, \varepsilon_{t,u_3})$$
$$+ Cum(Y_t, Y_{t,u_1} + \varepsilon_{t,u_1}, Y_{t,u_2}, \varepsilon_{t,u_3}) + Cum(Y_t, Y_{t,u_1} + \varepsilon_{t,u_1}, Y_{t,u_2} + \varepsilon_{t,u_2}, Y_{t,u_3})$$
$$= D_0^{[3]} + D_1^{[3]} + D_2^{[3]} + D_3^{[3]} \quad \text{(say), with } D_0^{[3]} \equiv 0.$$

To deal with the terms $D_3^{[3]}$, $D_2^{[3]}$, and $D_1^{[3]}$ in turn, we apply

$$Cum(X_0, X_1, X_2, X_3) = E(X_0X_1X_2X_3) - E(X_0X_1X_2)E(X_3)\langle 4\rangle - E(X_0X_1)E(X_2X_3)\langle 3\rangle$$
$$+ 2E(X_0X_1)E(X_2)E(X_3)\langle 6\rangle - 6E(X_0)E(X_1)E(X_2)E(X_3).$$

Step 1. The repeated use of the conditioning argument, in relation to (5.4), shows that, given I, I_1, I_2, I_3, whose values are 0 or 1, with $I + \sum_{j=1}^{3} I_j \geq 1$,

$$\left| E\left[Y_t^I \left\{ \prod_{j=1}^{2} (Y_{t,u_j} + \varepsilon_{t,u_j})^{I_j} \right\} Y_{t,u_3}^{I_3} \right] \right| \leq C_3^{[3]} \alpha^{I_3 u_3},$$

which yields

$$|D_3^{[3]}| \leq [C_3^{[3]} + 7(C_3^{[3]})^2 + 12(C_3^{[3]})^3 + 6(C_3^{[3]})^4]\alpha^{u_3} \leq 6B_4[1 + (C_3^{[3]})^4]\alpha^{u_3}$$

(B_r stands for the Bell number; note $B_4 = 1 + 7 + 6 + 1 = 15$).

Step 2. We notice that

$$\begin{aligned} D_2^{[3]} &= \sum_{i_3=u_3-u_2+1}^{u_3-1} Cum(Y_t, Y_{t,u_1} + \varepsilon_{t,u_1}, Y_{t,u_2}, \varepsilon_{t,u_3,i_3}), \\ D_1^{[3]} &= \sum_{i_3=u_3-u_1+1}^{u_3-1} Cum(Y_t, Y_{t,u_1}, \varepsilon_{t,u_2}, \varepsilon_{t,u_3,i_3}) \\ &\quad + \sum_{i_2=u_2-u_1+1}^{u_2-1} \sum_{i_3=u_3-u_2+1}^{u_3-u_1} Cum(Y_t, Y_{t,u_1}, \varepsilon_{t,u_2,i_2}, \varepsilon_{t,u_3,i_3}) \\ &= D_{1,1}^{[3]} + D_{1,2}^{[3]} \quad \text{(say)}. \end{aligned}$$

Similarly to Step 1, we deal with three terms $D_2^{[3]}$, $D_{1,1}^{[3]}$, and $D_{1,2}^{[3]}$ in the following way:

$D_2^{[3]}$. As long as $u_3 - u_2 + 1 \leq i_3 (\leq u_3 - 1)$,

$$|E[Y_t^I (Y_{t,u_1} + \varepsilon_{t,u_1})^{I_1} Y_{t,u_2}^{I_2} \varepsilon_{t,u_3,i_3}^{I_3}]| \leq C_{2,1}^{[3]} \alpha^{I_2 u_2 + I_3(u_3-u_2)} \vartheta^{I_3(i_3+u_2-u_3)},$$

which enables us to see that

$$|D_2^{[3]}| \leq [C_{2,1}^{[3]} + 7(C_{2,1}^{[3]})^2 + 12(C_{2,1}^{[3]})^3 + 6(C_{2,1}^{[3]})^4]\alpha^{u_3} \sum_{i_3=u_3-u_2+1}^{u_3-1} \vartheta^{i_3+u_2-u_3} \leq 6B_4[1 + (C_{2,1}^{[3]})^4]\frac{\alpha^{u_3}}{1-\vartheta}.$$

$D_{1,1}^{[3]}$. As long as $u_3 - u_1 + 1 \leq i_3 (\leq u_3 - 1)$,

$$|E(Y_t^I Y_{t,u_1}^{I_1} \varepsilon_{t,u_2}^{I_2} \varepsilon_{t,u_3,i_3}^{I_3})| \leq C_{1,1}^{[3]} \alpha^{I_1 u_1 + I_3(u_3-u_1)} \vartheta^{I_3(i_3+u_1-u_3)},$$

which enables us to see that

$$|D_{1,1}^{[3]}| \leq [C_{1,1}^{[3]} + 7(C_{1,1}^{[3]})^2 + 12(C_{1,1}^{[3]})^3 + 6(C_{1,1}^{[3]})^4]\alpha^{u_3} \sum_{i_3=u_3-u_1+1}^{u_3-1} \vartheta^{i_3+u_1-u_3} \leq 6B_4[1 + (C_{1,1}^{[3]})^4]\frac{\alpha^{u_3}}{1-\vartheta}.$$

$D_{1,2}^{[3]}$. As long as $u_3 - u_2 + 1 \leq i_3 \leq u_3 - u_1$ and $u_2 - u_1 + 1 \leq i_2 (\leq u_2 - 1)$ (in this case, $\mathcal{T}_{t,u_3,i_3}$

intersects with $\mathcal{T}_{t,u_2,i_2}$),

$$|E(Y_t^I Y_{t,u_1}^{I_1} \varepsilon_{t,u_2,i_2}^{I_2} \varepsilon_{t,u_3,i_3}^{I_3})| \leq C_{1,2}^{[3]} \alpha^{I_1 u_1 + \sum_{j=2}^{3} I_j (u_j - u_{j-1})} \vartheta^{\sum_{j=2}^{3} I_j (i_j + u_{j-1} - u_j)},$$

which enables us to see that

$$\begin{aligned}|D_{1,2}^{[3]}| &\leq [C_{1,2}^{[3]} + 7(C_{1,2}^{[3]})^2 + 12(C_{1,2}^{[3]})^3 + 6(C_{1,2}^{[3]})^4] \alpha^{u_3} \sum_{i_2=u_2-u_1+1}^{u_2-1} \sum_{i_3=u_3-u_2+1}^{u_3-u_1} \vartheta^{\sum_{j=2}^{3}(i_j+u_j-u_{j-1})} \\ &\leq 6B_4[1 + (C_{1,2}^{[3]})^4] \frac{\alpha^{u_3}}{(1-\vartheta)^2}.\end{aligned}$$

This completes the proof of $r = 3$.

Even for the general case $r = 4, 5, \ldots$, we obviously have $Cum(Y_t, \varepsilon_{t,u_1}, \ldots, \varepsilon_{t,u_r}) \equiv 0$, hence,

$$Cum(Y_t, Y_{t+u_1}, \ldots, Y_{t+u_r}) = \sum_{J=1}^{r} D_J^{[r]},$$

where

$$D_J^{[r]} = \begin{cases} Cum(Y_t, Y_{t,u_1} + \varepsilon_{t,u_1}, \ldots, Y_{t,u_{r-1}} + \varepsilon_{t,u_{r-1}}, Y_{t,u_r}), & J = r, \\ Cum(Y_t, Y_{t,u_1} + \varepsilon_{t,u_1}, \ldots, Y_{t,u_{J-1}} + \varepsilon_{t,u_{J-1}}, Y_{t,u_J}, \varepsilon_{t,u_{J+1}}, \ldots, \varepsilon_{t,u_r}), & J = 1, \ldots, r-1. \end{cases}$$

The basic tool for dealing with the terms $D_J^{[r]}$'s is the cumulant-moment relation, given by

$$Cum(X_0, \ldots, X_r) = E\Big(\prod_{i=0}^{r} X_i\Big) - E\Big(\prod_{i=0}^{r-1} X_i\Big) E(X_r) \langle r+1 \rangle + \cdots + (-1)^r r! \prod_{i=0}^{r} E(X_i)$$

(e.g., McCullagh (2018)). We adopt the conventional notation of the empty sum/product, i.e., $\sum_{j\in\emptyset} a_j = 0$ and $\prod_{j\in\emptyset} a_j = 1$.

In the same manner as the term $D_3^{[3]}$, we can deal with the last term $D_r^{[r]}$. That is, the repeated use of the conditioning argument, in relation to (5.4), shows that, given $I, I_1, \ldots, I_r$, whose values are 0 or 1, with $I + \sum_{j=1}^{r} I_j \geq 1$,

$$\Big|E\Big[Y_t^I \Big\{\prod_{j=1}^{r-1} (Y_{t,u_j} + \varepsilon_{t,u_j})^{I_j}\Big\} Y_{t,u_r}^{I_r}\Big]\Big| \leq C_r^{[r]} \alpha^{I_r u_r},$$

which enables us to see that

$$|D_r^{[r]}| \leq \Big\{\sum_{i=1}^{r+1} (i-1)! S\big(\begin{smallmatrix} r+1 \\ i \end{smallmatrix}\big) (C_r^{[r]})^i\Big\} \alpha^{u_r} \leq r! B_{r+1} [1 + (C_r^{[r]})^{r+1}] \alpha^{u_r}.$$

To complete the proof, we need to consider other terms $D_J^{[r]}$, $J = 1, \ldots, r-1$.

For every case of $J \in \{1, \ldots, r-1\}$, we denote by K_J the rightmost position of $k \in \{J+1, \ldots, r\}$

such that $\mathcal{T}_{t,u_J} \cap \mathcal{T}_{t,u_k,i_k} \neq \emptyset$, i.e., by definition,

(i) $u_{K_J} - u_J + 1 \le i_{K_J} (\le u_{K_J} - 1)$ and (ii) when $K_J \neq r$, $0 \le i_\ell \le u_\ell - u_J$ for $\ell = K_J + 1, \ldots, r$.

On the basis of this proof strategy, let

$$C^{[r]}_{i_{K_J},\ldots,i_r} = Cum(Y_t, Y_{t,u_1} + \varepsilon_{t,u_1}, \ldots, Y_{t,u_{J-1}} + \varepsilon_{t,u_{J-1}}, Y_{t,u_J}, \varepsilon_{t,u_{J+1}}, \ldots, \varepsilon_{t,u_{K_J-1}}, \varepsilon_{t,u_{K_J},i_{K_J}}, \ldots, \varepsilon_{t,u_r,i_r}).$$

Besides, when $K_J \neq r$, $(i_{K_J+1}, \ldots, i_r)$ must be restricted to the condition that $u_\ell - u_{\ell-1} + 1 \le i_\ell (\le u_\ell - u_J)$ for $\ell = K_J + 1, \ldots, r$. That is, for every $J \in \{1, \ldots, r-1\}$,

$$D^{[r]}_J = \sum_{K_J=J+1}^{r} D^{[r]}_{J,r-K_J+1},$$

where

$$D^{[r]}_{J,r-K_J+1} = \sum_{i_{K_J}=u_{K_J}-u_J+1}^{u_{K_J}-1} \sum_{i_{K_J+1}=u_{K_J+1}-u_{K_J}+1}^{u_{K_J+1}-u_J} \cdots \sum_{i_r=u_r-u_{r-1}+1}^{u_r-u_J} C^{[r]}_{i_{K_J},\ldots,i_r} = \sum_{(i_{K_J},\ldots,i_r)\in \mathcal{I}^{[r]}_J} C^{[r]}_{i_{K_J},\ldots,i_r} \quad \text{(say)},$$

noting that $C^{[r]}_{i_{K_J},\ldots,i_r} \equiv 0$ when $(i_{K_J}, \ldots, i_r) \in [0, u_{K_J} - 1] \times \cdots \times [0, u_r - 1] \backslash \mathcal{I}^{[r]}_J$. For every $J \in \{1, \ldots, r-1\}$, given $I, I_1, \ldots, I_r$, whose values are 0 or 1, with $I + \sum_{j=1}^r I_j \ge 1$, as long as $(i_{K_J}, \ldots, i_r) \in \mathcal{I}^{[r]}_J$,

$$\begin{aligned}
&\left| E\left[Y_t^I \left\{ \prod_{j=1}^{J-1} (Y_{t,u_j} + \varepsilon_{t,u_j})^{I_j} \right\} Y_{t,u_J}^{I_J} \left\{ \prod_{j=J+1}^{K_J-1} \varepsilon_{t,u_j}^{I_j} \right\} \left\{ \prod_{j=K_J}^{r} \varepsilon_{t,u_j,i_j}^{I_j} \right\} \right] \right| \\
&\le C^{[r]}_{J,r-K_J+1} \alpha^{I_J u_J + I_{K_J}(u_{K_J}-u_J) + \sum_{j=K_J+1}^r I_j (u_j - u_{j-1})} \vartheta^{I_{K_J}(i_{K_J}+u_J-u_{K_J}) + \sum_{j=K_J+1}^r I_j (i_j + u_{j-1} - u_j)},
\end{aligned}$$

which enables us to see that

$$\begin{aligned}
|D^{[r]}_{J,r-K_J+1}| &\le \left\{ \sum_{i=1}^{r+1} (i-1)! S\binom{r+1}{i} \right\} \alpha^{u_r} \sum_{(i_{J_K},\ldots,i_r)\in \mathcal{I}^{[r]}_J} \vartheta^{i_{K_J}+u_J-u_{K_J}+\sum_{j=K_J+1}^r (i_j+u_{j-1}-u_j)} \\
&\le r! B_{r+1} \{1 + (C^{[r]}_{J,r-K_J+1})^{r+1}\} \frac{\alpha^{u_r}}{(1-\vartheta)^{r-K_J+1}} \\
&\le r! B_{r+1} \frac{1 + (C^{[r]}_{J,r-K_J+1})^{r+1}}{(1-\vartheta)^{r-J}} \alpha^{u_r}.
\end{aligned}$$

Summing over $K_J = J + 1, \ldots, r$, it follows that

$$|D^{[r]}_J| \le r! B_{r+1} \frac{\sum_{i=1}^{r-J} \{1 + (C^{[r]}_{J,j})^{r+1}\}}{(1-\vartheta)^{r-J}} \alpha^{u_r}, \quad J = 1, \ldots, r-1.$$

In this way, we can prove that

$$Cum(Y_t, Y_{t+u_1}, \ldots, Y_{t+u_r}) \leq r! B_{r+1} \Big[1 + (C_r^{[r]})^{r+1} + \sum_{J=1}^{r-1} \frac{\sum_{i=1}^{r-J} \{1 + (C_{J,j}^{[r]})^{r+1}\}}{(1-\vartheta)^{r-J}} \Big] \alpha^{u_r}. \quad \square$$

Chapter 6

Some estimators in ADCINAR(1) process

6.1 Introduction

Nastić et al. (2017) considered the stationary alternative dependent counting nonnegative integer-valued autoregressive process of the first-order (ADCINAR(1)), defined by $Y_t = \alpha \diamond_\vartheta Y_{t-1} + \varepsilon_t$ under the specific innovation $\{\varepsilon_t\}$ (see Theorem 3 in Nastić et al. (2017); the marginal distribution of $\{Y_t\}$ is then geometric distribution Geo$(\mu/(1+\mu))$ for $\mu > 0$), where $\alpha \diamond_\vartheta$ is an alternative generalized binomial thinning operator with $0 \leq \alpha \leq \vartheta < 1 (\vartheta \neq 0)$. They gave asymptotic normality of the estimator for the new parameter ϑ when other parameter (α, μ) is unrealistically known.

However, in practice, the distributional assumption on the innovation can not be specified in advance. Also, the conditional expectation is given by $E(Y_t|Y_{t-1}) = \alpha Y_{t-1} + E(\varepsilon_t)$, hence, α is an important parameter to be inferenced. This motivates us to revisit asymptotic properties of an estimator for ϑ without assuming that α is known.

One of the contributions (here, we are interested in a general innovation, not the one considered by Nastić et al. (2017)) is Proposition 6.3 in Section 6.2, which elucidates that plugging a certain $n^{1/2}$-consistent estimator for α affects the resulting asymptotic variance, whereas plugging the sample mean and variance for $E(Y_1)$ and $V(Y_1)$ has no effect. The phenomenon can not be grasped from two examples discussed in Karlsen and Tjøstheim (1988). Since two-step procedure is one of the commonly used techniques in time series analysis, this finding gives a warning of the use of Karlsen and Tjøstheim's (1988) theory in the recent literature of the stationary nonnegative integer-valued process.

In the statistical analysis of time series count data, the conditional least squares (CLS) method due to Klimko and Nelson (1978) has been widely used. See, e.g., Al-Osh and Alzaid (1987) and Park and Oh (1997) for the stationary nonnegative integer-valued autoregressive process of the first-order (INAR(1)) defined by $Y_t = \alpha \circ Y_{t-1} + \varepsilon_t$, and Nastić et al. (2017) for the stationary ADCINAR(1) process defined by $Y_t = \alpha \diamond_\vartheta Y_{t-1} + \varepsilon_t$. Here, we do not assume the specific distributional form about the innovation $\{\varepsilon_t\}$. The Yule–Walker (YW) estimator for the parameter α can be applied easily, since the autocorrelation structures of various nonnegative integer-valued time series models are the same as the usual stationary autoregressive process of the first-order. A general estimator for the parameter α (see Chapter 3) can be also applied even for the ADCINAR(1) process.

It is well known that these estimators are strong consistent and asymptotic normal under moment conditions on a general innovation, but are biased in a finite-sample. Some authors derived asymptotic expansions of the biases in order to perform an analytical bias-correction for the stationary INAR(1) process (see Bourguignon and Vasconcellos (2015), Weiß and Schweer (2016), and Zeng and Kakizawa (2022)). As an extension of Chapter 3, we also provide the asymptotic expansions of the biases of the YW and CLS estimators for the parameter α even in the stationary ADCINAR(1) process. Not surprisingly, the resulting analytical bias-corrected estimators are complicated, involving the estimator for the new parameter ϑ. Therefore, the major contribution of this chapter is to develop, in a nonparametric way, other bias-corrected estimators, without computing the closed-form expression for the asymptotic expansions of the biases.

The rest of this chapter is organized as follows. Section 6.2 shows consistency and asymptotic normality of two-step CLS (2CLS) estimator for the new parameter ϑ in the stationary ADCINAR(1) process. Section 6.3 develops the bias-corrected YW and CLS estimators for the parameter α, together with their theoretical justifications. Simulation experiments are conducted in Section 6.4, to assess the finite-sample performances of the estimators (without/with bias-correction). Section 6.5 concludes this chapter. The proof of Proposition 6.3(i) is postponed to Section 6.6.

6.2 Estimation for the parameters α and ϑ

Suppose that the observation $\{Y_1, \dots, Y_n\}$ of length n is generated by $Y_t = \alpha \diamond_\vartheta Y_{t-1} + \varepsilon_t$. Here, we assume that $0 \le \alpha \le \vartheta < 1 (\vartheta \ne 0)$, i.e., the ADCINAR(1) process is strictly stationary and ergodic; see Nastić et al. (2017). As usual, we define the sample mean, variance, and cumulants by $\overline{Y} = (1/n)\sum_{t=1}^n Y_t$, $\widehat{\sigma}_Y^2 = (1/n)\sum_{t=1}^n (Y_t - \overline{Y})^2$, $\widehat{\kappa}_{3,Y} = (1/n)\sum_{t=1}^n (Y_t - \overline{Y})^3$, and $\widehat{\kappa}_{4,Y} = (1/n)\sum_{t=1}^n (Y_t - \overline{Y})^4 - 3\widehat{\sigma}_Y^4$, respectively. Note that $\overline{Y} \xrightarrow{a.s.} \mu_Y$, $\widehat{\sigma}_Y^2 \xrightarrow{a.s.} \sigma_Y^2$, and $\widehat{\kappa}_{j,Y} \xrightarrow{a.s.} \kappa_{j,Y}$, $j = 3, 4$. See Lemma 5.5(i). Also, I_S stands for the indicator of the set S.

Estimation for the parameter α

The CLS method is widely used for estimating the parameter α in the INAR-type processes. That is, minimizing the criterion

$$J(\alpha, \mu_\varepsilon) = \sum_{t=2}^n (Y_t - E[Y_t|Y_{t-1}])^2 = \sum_{t=2}^n (Y_t - \alpha Y_{t-1} - \mu_\varepsilon)^2, \tag{6.1}$$

the CLS estimator for α and μ_ε is defined by

$$\widehat{\alpha}_{CLS} = \frac{\sum_{t=2}^{n} Y_t Y_{t-1} - \frac{1}{n-1}\sum_{t=2}^{n} Y_t \sum_{t=2}^{n} Y_{t-1}}{\sum_{t=2}^{n} Y_{t-1}^2 - \frac{1}{n-1}\Big(\sum_{t=2}^{n} Y_{t-1}\Big)^2} \quad \text{and} \quad \widehat{\mu}_{\varepsilon;CLS} = \frac{1}{n-1}\Big(\sum_{t=2}^{n} Y_t - \widehat{\alpha}_{CLS}\sum_{t=2}^{n} Y_{t-1}\Big)$$

(see also Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)). Also, since α is the autocorrelation at lag 1 for the ADCINAR(1) process, a general estimator for the parameter α is defined by

$$\widehat{\alpha}_{c_1,c_2} = \frac{\sum_{t=2}^{n}(Y_t - \overline{Y})(Y_{t-1} - \overline{Y})}{c_1(Y_1 - \overline{Y})^2 + \sum_{t=2}^{n-1}(Y_t - \overline{Y})^2 + c_2(Y_n - \overline{Y})^2}, \quad c_1, c_2 \geq 0$$

(note that $\widehat{\alpha}_{c_1,c_2}$ includes the YW ($c_1 = c_2 = 1$) and Burg ($c_1 = c_2 = 1/2$) estimators as special cases; see Chapter 3).

We state (without proof) that the estimators $\widehat{\alpha}_{c_1,c_2}$ and $\widehat{\alpha}_{CLS}$ have the desirable asymptotic properties (asymptotic normality of $\widehat{\alpha}_{c_1,c_2}$ and $\widehat{\alpha}_{CLS}$ will be given latter).

Proposition 6.1. *Suppose that* $0 \leq \alpha \leq \vartheta < 1 (\vartheta \neq 0)$. *The following hold.*

(i) If $E(\varepsilon_t^2)$ *exists,* $\widehat{\alpha}_{c_1,c_2} \xrightarrow{a.s.} \alpha$ *and* $\widehat{\alpha}_{CLS} \xrightarrow{a.s.} \alpha$.

(ii) If $E(\varepsilon_t^4)$ *exists, the general estimator* $\widehat{\alpha}_{c_1,c_2}$ *and CLS estimator* $\widehat{\alpha}_{CLS}$ *admit the stochastic expansion in the form*

$$\sqrt{n}(\widehat{\alpha} - \alpha) = \frac{1}{\sigma_Y^2\sqrt{n}}\sum_{t=2}^{n}\{Y_t - \mu_Y - \alpha(Y_{t-1} - \mu_Y)\}(Y_{t-1} - \mu_Y) + o_p(1).$$

Also, $\widehat{\alpha}_{CLS} = \widehat{\alpha}_{1,0} + O_p(n^{-2})$.

Estimation for the new parameter ϑ

On the other hand, the criterion (6.1) is free of the parameter ϑ. Another tool is needed to estimate ϑ. According to Karlsen and Tjøstheim (1988) (see also Nastić et al. (2017)), we apply the 2CLS method. It is the CLS criterion as if the observations Y_t's were the squared residuals $\{Y_t - E(Y_t|Y_{t-1})\}^2$'s.

Note that, after some algebra, by Remark 5.1(i) and (5.10), in the stationary ADCINAR(1) process $\{Y_t\}$, we have:

Claim 6.1. *For* $t = 2, \ldots, n$ *and arbitrary function* G,

$$E[Y_t G(Y_{t-1})|Y_{t-1}] = \{\alpha Y_{t-1} + (1-\alpha)\mu_Y\}G(Y_{t-1}),$$
$$E[g_t(\alpha, \mu_Y)G(Y_{t-1})|Y_{t-1}] = \{f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) + \vartheta\alpha f_{2,t-1}(\mu_Y, \sigma_Y^2)\}G(Y_{t-1}),$$

where

$$g_t(\alpha, \mu_Y) = \{Y_t - \mu_Y - \alpha(Y_{t-1} - \mu_Y)\}^2,$$
$$f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) = -\alpha^2(Y_{t-1} - \mu_Y)^2 + \alpha(1 - 2\alpha\mu_Y)(Y_{t-1} - \mu_Y) + \sigma_Y^2,$$
$$f_{2,t-1}(\mu_Y, \sigma_Y^2) = (Y_{t-1} - \mu_Y)^2 - (1 - 2\mu_Y)(Y_{t-1} - \mu_Y) - \sigma_Y^2.$$

Now, we define

$$\begin{aligned} J(\alpha, \mu_Y, \sigma_Y^2, \vartheta) &= \sum_{t=2}^{n} [\{Y_t - E(Y_t|Y_{t-1})\}^2 - E\{(Y_t - E(Y_t|Y_{t-1}))^2|Y_{t-1}\}]^2 \\ &= \sum_{t=2}^{n} \{g_t(\alpha, \mu_Y) - f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) - \vartheta\alpha f_{2,t-1}(\mu_Y, \sigma_Y^2)\}^2. \end{aligned}$$

Suppose that an estimator $\widehat{\alpha}$ for $\alpha(>0)$ is available. Minimizing $J(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2, \vartheta)$ with respect to ϑ (here, the restriction $\alpha \leq \vartheta < 1$ is ignored), the 2CLS estimator for ϑ can be constructed in closed-form

$$\widehat{\vartheta}_{2CLS} = \frac{\sum_{t=2}^{n} \{g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2)\} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)}{\widehat{\alpha} \sum_{t=2}^{n} f_{2,t-1}^2(\overline{Y}, \widehat{\sigma}_Y^2)} = \widehat{\vartheta}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) \quad \text{(say)}.$$

Proposition 6.2. *Suppose that* $0 < \alpha \leq \vartheta < 1$ *and that* $E(\varepsilon_t^4)$ *exists. The following hold.*

(i) $(1/n) \sum_{t=2}^{n} f_{2,t-1}^2(\overline{Y}, \widehat{\sigma}_Y^2) = \widehat{B}$ *(say) converges to* $E[f_{2,1}^2(\mu_Y, \sigma_Y^2)] = B$ *(say) a.s.*

(ii) Assume that $B > 0$. *If* $\widehat{\alpha} \xrightarrow{a.s.} \alpha$ $(\widehat{\alpha} \xrightarrow{p} \alpha)$, *then, the 2CLS estimator* $\widehat{\vartheta}_{2CLS}$ *is strongly (weakly) consistent, i.e.,* $\widehat{\vartheta}_{2CLS} \xrightarrow{a.s.} \vartheta$ $(\widehat{\vartheta}_{2CLS} \xrightarrow{p} \vartheta)$.

Proof (i) The result can be shown easily by Lemma 5.5.

(ii) We start with

$$\widehat{\vartheta}_{2CLS} - \vartheta = \frac{\frac{1}{n} \sum_{t=2}^{n} \{g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) - \vartheta\widehat{\alpha} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)\} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)}{\frac{\widehat{\alpha}}{n} \sum_{t=2}^{n} f_{2,t-1}^2(\overline{Y}, \widehat{\sigma}_Y^2)}, \tag{6.2}$$

where $(1/n) \sum_{t=2}^{n} f_{2,t-1}^2(\overline{Y}, \widehat{\sigma}_Y^2) \xrightarrow{a.s.} B$. By Lemma 5.5(ii) and Claim 6.1, we see that, if $\widehat{\alpha} \xrightarrow{a.s.} \alpha$,

$$\begin{aligned} &\frac{1}{n} \sum_{t=2}^{n} \{g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) - \vartheta\widehat{\alpha} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)\} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2) \\ &\xrightarrow{a.s.} E[\{g_2(\alpha, \mu_Y) - f_{1,1}(\alpha, \mu_Y, \sigma_Y^2) - \vartheta\alpha f_{2,1}(\mu_Y, \sigma_Y^2)\} f_{2,1}(\mu_Y, \sigma_Y^2)] = 0 \end{aligned}$$

(it converges in probability to zero, if $\widehat{\alpha} \xrightarrow{p} \alpha$), which completes the proof. □

Before presenting the next result (Proposition 6.3), we define

$$A = (2\alpha - \vartheta)\frac{\kappa_{4,Y} + 2\sigma_Y^4}{\sigma_Y^2} + \{2\vartheta(1 - 2\mu_Y) - 1 - 2\alpha + 8\alpha\mu_Y\}\frac{\kappa_{3,Y}}{\sigma_Y^2} + (1 - 2\mu_Y)\{1 - 4\alpha\mu_Y - \vartheta(1 - 2\mu_Y)\},$$

and, for $t = 2, \ldots, n$,

$$\begin{aligned} M_{\alpha,t} &= \{Y_t - \mu_Y - \alpha(Y_{t-1} - \mu_Y)\}(Y_{t-1} - \mu_Y), \\ M_{\vartheta,t} &= \{g_t(\alpha, \mu_Y) - f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) - \vartheta\alpha f_{2,t-1}(\mu_Y, \sigma_Y^2)\} f_{2,t-1}(\mu_Y, \sigma_Y^2). \end{aligned}$$

By Claim 6.1, the expectations $\omega_{11} = E(M_{\alpha,2}^2)$, $\omega_{12} = E(M_{\alpha,2}M_{\vartheta,2})$, and $\omega_{22} = E(M_{\vartheta,2}^2)$ are, respectively, given by

$$\begin{aligned} \omega_{11} &= E[(Y_1 - \mu_Y)(Y_2 - \mu_Y)M_{\alpha,2}] \\ &= E[\{Y_2 - \mu_Y - \alpha(Y_1 - \mu_Y)\}(Y_2 - \mu_Y)(Y_1 - \mu_Y)^2], \qquad (6.3) \\ \omega_{12} &= E[g_2(\alpha, \mu_Y) f_{2,1}(\mu_Y, \sigma_Y^2) M_{\alpha,2}] \\ &= E[\{Y_2 - \mu_Y - \alpha(Y_1 - \mu_Y)\}^3 (Y_1 - \mu_Y) f_{2,1}(\mu_Y, \sigma_Y^2)], \qquad (6.4) \\ \omega_{22} &= E[g_2(\alpha, \mu_Y) f_{2,1}(\mu_Y, \sigma_Y^2) M_{\vartheta,2}] \\ &= E[\{Y_2 - \mu_Y - \alpha(Y_1 - \mu_Y)\}^2 f_{2,1}(\mu_Y, \sigma_Y^2) M_{\vartheta,2}]. \qquad (6.5) \end{aligned}$$

Proposition 6.3. *Suppose that* $0 < \alpha \leq \vartheta < 1$ *and that* $E(\varepsilon_t^8)$ *exists. Assume that* $B > 0$*. If*

$$\sqrt{n}(\widehat{\alpha} - \alpha) = \frac{1}{\sigma_Y^2\sqrt{n}}\sum_{t=2}^{n} M_{\alpha,t} + o_p(1) \qquad (6.6)$$

(the estimators $\widehat{\alpha}_{YW}$ *and* $\widehat{\alpha}_{CLS}$ *satisfy this condition; see Proposition 6.1(ii)), then, the following hold.*

(i) $\sqrt{n}(\widehat{\vartheta}_{2CLS} - \vartheta) = \frac{1}{\alpha B\sqrt{n}}\sum_{t=2}^{n}(AM_{\alpha,t} + M_{\vartheta,t}) + o_p(1)$.

(ii) $\sqrt{n}(\widehat{\alpha} - \alpha, \widehat{\vartheta}_{2CLS} - \vartheta)^T \xrightarrow{d} \mathrm{N}(0, \Psi)$,

with

$$\Psi = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{12} & \psi_{22} \end{pmatrix},$$

where $\psi_{11} = \omega_{11}/\sigma_Y^4$, $\psi_{12} = (A\omega_{11} + \omega_{12})/(\alpha B\sigma_Y^2)$, *and* $\psi_{22} = (A^2\omega_{11} + 2A\omega_{12} + \omega_{22})/(\alpha B)^2$.

Proof The proof of the stochastic expansion (i) is postponed to Section 6.6. The asymptotic normality

(ii) is readily established by Slutsky's theorem, once the following is shown:

$$\frac{1}{\sqrt{n}}\sum_{t=2}^{n}(M_{\alpha,t}, M_{\vartheta,t})^T \xrightarrow{d} \mathrm{N}(0,\Omega), \tag{6.7}$$

with

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{pmatrix}.$$

Thus, we have only to prove (6.7), as follows.

Let $\mathscr{F}_t = \sigma\{Y_1, \ldots, Y_t\}$ be a sigma-field, and let $n \geq 2$. Note that $\{n^{-1/2}M_{\#,t}, t = 2, \ldots, n\}$, $\# = \alpha, \vartheta$, are martingale difference arrays, i.e., $E(n^{-1/2}M_{\#,t}|\mathscr{F}_{t-1}) = 0$. See Claim 6.1.

Now, for $\# = \alpha, \vartheta$, we have $E(M^2_{\#,t}) = E(M^2_{\#,2})(< \infty)$, by strictly stationarity of $\{Y_t\}$. Then, for any $\varepsilon > 0$,

$$\begin{aligned} P\Big(\max_{t=2,\ldots,n} |n^{-1/2}M_{\#,t}| > \varepsilon\Big) &\leq \sum_{t=2}^{n} P(|M_{\#,t}| > \sqrt{n}\varepsilon) \\ &= (n-1)P(|M_{\#,2}| > \sqrt{n}\varepsilon) \quad \text{(by strictly stationarity of } \{Y_t\}) \\ &\leq \frac{1}{\varepsilon^2}E[I_{\{|M_{\#,2}|>\sqrt{n}\varepsilon\}}M^2_{\#,2}] \\ &\to 0 \quad \text{(by Lebesgue's dominated convergence theorem)}, \end{aligned}$$

which yields

$$\max_{t=2,\ldots,n} |n^{-1/2}M_{\#,t}| = o_p(1), \quad \# = \alpha, \vartheta. \tag{6.8}$$

Also, for $\# = \alpha, \vartheta$, we see that

$$\frac{1}{n}E\Big(\max_{t=2,\ldots,n} M^2_{\#,t}\Big) \leq \frac{1}{n}\sum_{t=2}^{n} E(M^2_{\#,t}) \leq E(M^2_{\#,2}), \tag{6.9}$$

and that, by strictly stationarity and ergodicity of $\{Y_t\}$,

$$\frac{1}{n}\sum_{t=2}^{n} M^2_{\#,t} \xrightarrow{a.s.} E(M^2_{\#,2}) \ \text{ and } \ \frac{1}{n}\sum_{t=2}^{n} M_{\alpha,t}M_{\vartheta,t} \xrightarrow{a.s.} E(M_{\alpha,2}M_{\vartheta,2}). \tag{6.10}$$

By means of the Cramér–Wold device, the asymptotic normality (6.7) is shown by martingale central limit theorem (see McLeish (1974)) and (6.8)–(6.10). □

Remark 6.1. Although it is unrealistic, suppose that $(\alpha, \mu_Y, \sigma^2_Y)$ is known in advance. As in the above-mentioned proof of Proposition 6.3, we can prove that

$$\sqrt{n}[\widehat{\vartheta}(\alpha, \mu_Y, \sigma_Y^2) - \vartheta] \xrightarrow{d} \mathrm{N}(0, \omega_{22}/(\alpha B)^2).$$

If $(A^2\omega_{11} + 2A\omega_{12})/(\alpha B)^2 = \mathcal{D}$ (say) is negative, the infeasible estimator $\widehat{\vartheta}(\alpha, \mu_Y, \sigma_Y^2)$ underperforms the feasible estimator $\widehat{\vartheta}_{2CLS}$ asymptotically.

By Lemma 5.5(i) and Proposition 6.2(ii), it is easy to see that, if $\widehat{\alpha} \xrightarrow{a.s.} \alpha(>0)$,

$$\begin{aligned}\widehat{A} &= (2\widehat{\alpha} - \widehat{\vartheta}_{2CLS})\frac{\widehat{\kappa}_{4,Y} + 2(\widehat{\sigma}_Y^2)^2}{\widehat{\sigma}_Y^2} + \{2\widehat{\vartheta}_{2CLS}(1 - 2\overline{Y}) - 1 - 2\widehat{\alpha} + 8\widehat{\alpha}\overline{Y}\}\frac{\widehat{\kappa}_{3,Y}}{\widehat{\sigma}_Y^2} \\ &\quad + (1 - 2\overline{Y})\{1 - 4\widehat{\alpha}\overline{Y} - \widehat{\vartheta}_{2CLS}(1 - 2\overline{Y})\} \\ &\xrightarrow{a.s.} A \ \ (\text{if } \widehat{\alpha} \xrightarrow{p} \alpha(>0), \text{ then, } \widehat{A} \xrightarrow{p} A).\end{aligned}$$

This, together with Lemma 5.5 and Proposition 6.2, implies that estimators for ψ_{ij}'s (see also (6.3)–(6.5)) can be constructed as

$$\widehat{\psi}_{11} = \frac{\widehat{\omega}_{11}}{\widehat{\sigma}_Y^4}, \ \ \widehat{\psi}_{12} = \frac{\widehat{A}\widehat{\omega}_{11} + \widehat{\omega}_{12}}{\widehat{\alpha}\widehat{B}\widehat{\sigma}_Y^2}, \ \ \text{and} \ \ \widehat{\psi}_{22} = \frac{\widehat{A}^2\widehat{\omega}_{11} + 2\widehat{A}\widehat{\omega}_{12} + \widehat{\omega}_{22}}{(\widehat{\alpha}\widehat{B})^2},$$

respectively, where

$$\begin{aligned}\widehat{\omega}_{11} &= \frac{1}{n}\sum_{t=2}^{n}\{Y_t - \overline{Y} - \widehat{\alpha}(Y_{t-1} - \overline{Y})\}(Y_t - \overline{Y})(Y_{t-1} - \overline{Y})^2, \\ \widehat{\omega}_{12} &= \frac{1}{n}\sum_{t=2}^{n}\{Y_t - \overline{Y} - \widehat{\alpha}(Y_{t-1} - \overline{Y})\}^3(Y_{t-1} - \overline{Y})f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2), \\ \widehat{\omega}_{22} &= \frac{1}{n}\sum_{t=2}^{n}\{Y_t - \overline{Y} - \widehat{\alpha}(Y_{t-1} - \overline{Y})\}^2\{g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) - \widehat{\vartheta}_{2CLS}\widehat{\alpha}f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)\}f_{2,t-1}^2(\overline{Y}, \widehat{\sigma}_Y^2).\end{aligned}$$

The standard errors (SEs) of the estimators $\widehat{\alpha}$ and $\widehat{\vartheta}_{2CLS}$ are then given by $\sqrt{\widehat{\psi}_{11}/n}$ and $\sqrt{\widehat{\psi}_{22}/n}$, respectively.

6.3 Bias-corrections of YW and CLS estimators for the parameter α

It is well known that the estimators for the parameter α are biased in a finite-sample. We focus on the bias-correction of the two commonly used estimators; one is the YW estimator $\widehat{\alpha}_{YW} = \widehat{\alpha}_{1,1}$, and the other is the CLS estimator $\widehat{\alpha}_{CLS}$. From now on, let # $= YW, CLS$, unless otherwise stated.

Let $X_t = (Y_t - \mu_Y)/\sigma_Y$, $t = 1, \ldots, n$. In the same way as in Chapter 3, it is shown that, even for the stationary ADCINAR(1) process,

$$E(\widehat{\alpha}_{c_1,c_2} - \alpha) = -\frac{1 + (c_1 + c_2)\alpha}{n} - Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2\Big] + o(n^{-1}), \qquad (6.11)$$

where

$$\begin{aligned}
&Cov\Big[\frac{1}{n}\sum_{t=2}^{n}(X_t - \alpha X_{t-1})X_{t-1}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2\Big] \\
&= \frac{1}{n^2}\sum_{s,t=2}^{n} I_{\{s<t\}} Cov[(X_s - \alpha X_{s-1})X_{s-1}, X_{t-1}^2] \\
&= \frac{1}{n^2\sigma_Y^4}\sum_{s,t=2}^{n} I_{\{s<t\}}[\mu_Y(1, t-s, t-s) - \alpha\mu_Y(0, t-s, t-s)] \\
&= \frac{1}{n\sigma_Y^4}\sum_{\ell=1}^{n-2}\Big(1 - \frac{\ell+1}{n}\Big)[\mu_Y(1,\ell,\ell) - \alpha\mu_Y(0,\ell,\ell)] \\
&= \frac{1}{n\sigma_Y^4}\{M(1) - \alpha M(0)\} + O(n^{-2}), \qquad (6.12)
\end{aligned}$$

with

$$M(u) = \sum_{\ell=1}^{\infty}\mu_Y(u,\ell,\ell), \quad u = 0, 1.$$

Below, we will estimate $M(1) - \alpha M(0)$, in two different ways.

6.3.1 Lag window-type bias-correction

Define

$$\widehat{\mu}_Y(u,\ell,\ell) = \frac{1}{n}\sum_{t=1}^{n-\ell}(Y_t - \overline{Y})(Y_{t+u} - \overline{Y})(Y_{t+\ell} - \overline{Y})^2, \quad u = 0, 1 \text{ and } \ell = 1, \dots, n-1.$$

The estimation of $M(1) - \alpha M(0)$ (see (6.12)) is proceeded in a nonparametric way, referring to the issue of the spectral density estimation (e.g., Anderson (1971)) and the robust standard error estimation in the regression model with heteroscedastic and autocorrelated disturbances. Theoretical justification will be given later.

Now, $\widehat{\alpha}_{YW} = \widehat{\alpha}_{1,1}$ and $\widehat{\alpha}_{CLS} = \widehat{\alpha}_{1,0} + O_p(n^{-2})$, as mentioned in Proposition 6.1(ii). We can construct the bias-corrected YW and CLS estimators, in the form of

$$\widetilde{\alpha}_{YW} = \widehat{\alpha}_{YW} + \frac{1}{n}\Big[1 + 2\widehat{\alpha}_{YW} + \frac{1}{\widehat{\sigma}_Y^4}\{\widehat{M}(1) - \widehat{\alpha}_{YW}\widehat{M}(0)\}\Big], \qquad (6.13)$$

$$\widetilde{\alpha}_{CLS} = \widehat{\alpha}_{CLS} + \frac{1}{n}\Big[1 + \widehat{\alpha}_{CLS} + \frac{1}{\widehat{\sigma}_Y^4}\{\widehat{M}(1) - \widehat{\alpha}_{CLS}\widehat{M}(0)\}\Big], \qquad (6.14)$$

where

$$\widehat{M}(u) = \sum_{\ell=1}^{L_n} k(\ell/L_n)\widehat{\mu}_Y(u,\ell,\ell), \quad u = 0, 1,$$

for given weight function k and positive integer $L_n(< n)$, referred to as lag window function and truncation parameter, respectively (assume that $L_n \to \infty$ as $n \to \infty$, with a suitable rate; see Proposition 6.4 below).

Justifications of (6.13) and (6.14)

It remains to justify $\widehat{M}(1) - \widehat{\alpha}_{\#}\widehat{M}(0)$, where $\# = YW, CLS$.

For notational simplicity, with $Z_t = Y_t - \mu_Y$, $t = 1, \ldots, n$, let

$$\widetilde{\mu}_Y(u,\ell,\ell) = (1/n)\sum_{t=1}^{n-\ell} Z_t Z_{t+u} Z_{t+\ell}^2 \quad \text{and} \quad \widetilde{M}(u) = \sum_{\ell=1}^{L_n} k(\ell/L_n)\widetilde{\mu}_Y(u,\ell,\ell), \quad u = 0, 1.$$

We here restrict ourselves to the form of $k(x) = 1 - |x|^q$ for some $q > 0$, which is famous as a class of Parzen's (1957) lag window functions, with $\sup_{0\le x\le 1}|k(x)| \le 1$ (other lag window functons are available in, e.g., Anderson (1971)).

Lemma 6.1. *Suppose that* $0 \le \alpha \le \vartheta < 1 (\vartheta \ne 0)$ *and that* $E(\varepsilon_t^8)$ *exists. The following hold.*

(i) If $1/L_n + L_n^q/n = o(1)$*, then,*

$$E[\widetilde{M}(1) - \alpha\widetilde{M}(0) - \{M(1) - \alpha M(0)\}] = -L_n^{-q}\sum_{\ell=1}^{\infty} \ell^q\{\mu_Y(1,\ell,\ell) - \alpha\mu_Y(0,\ell,\ell)\} + o(L_n^{-q}).$$

(ii) If $L_n \to \infty$*, then,*

$$V[\widetilde{M}(1) - \alpha\widetilde{M}(0)] = O(L_n^2/n).$$

(iii) If $1/L_n + L_n^{\min(q,2)}/n = o(1)$*, then,*

$$\widetilde{M}(1) - \alpha\widetilde{M}(0) = M(1) - \alpha M(0) + o_p(1).$$

(iv) If $L_n = O(n^{1/(2+2q)})$*, then,*

$$MSE[\widetilde{M}(1) - \alpha\widetilde{M}(0)] = O(n^{-q/(1+q)}).$$

Proof It is easy to see that

$$E[\widetilde{M}(1) - \alpha\widetilde{M}(0) - \{M(1) - \alpha M(0)\}]$$

$$= \sum_{\ell=1}^{L_n} \{k(\ell/L_n) - 1\}\{\mu_Y(1,\ell,\ell) - \alpha\mu_Y(0,\ell,\ell)\}$$
$$- \sum_{\ell=1}^{L_n} k(\ell/L_n)\frac{\ell}{n}\{\mu_Y(1,\ell,\ell) - \alpha\mu_Y(0,\ell,\ell)\} - \sum_{\ell=L_n+1}^{\infty} \{\mu_Y(1,\ell,\ell) - \alpha\mu_Y(0,\ell,\ell)\}.$$

Then, (i) is a direct consequence of an exponential decay of $|\mu_Y(1,\ell,\ell) - \alpha\mu_Y(0,\ell,\ell)| \le C\alpha^{\ell}$ (see Remark 5.3(ii)) for some constant $C > 0$, independent of ℓ.

To prove (ii), we notice that

$$V[\widetilde{M}(1) - \alpha\widetilde{M}(0)] \le 2\{V[\widetilde{M}(1)] + V[\widetilde{M}(0)]\},$$

where

$$V[\widetilde{M}(u)] = \frac{1}{n^2}\sum_{\ell=1}^{L_n}\sum_{\ell'=1}^{L_n} k(\ell/L_n)k(\ell'/L_n)\sum_{s=1}^{n-\ell}\sum_{t=1}^{n-\ell'} Cov(Z_s Z_{s+u} Z_{s+\ell}^2, Z_t Z_{t+u} Z_{t+\ell}^2), \quad u = 0, 1.$$

By McCullagh (2018), $Cov(Z_s Z_{s+u} Z_{s+\ell}^2, Z_t Z_{t+u} Z_{t+\ell}^2)$ can be expressed in terms of cumulants (we used his table of complementary set partitions for the case $\tau = 1234|5678$). Then, Proposition 5.2 enables us to show that

$$V[\widetilde{M}(u)] = V_1 + O(L_n/n),$$

where

$$\begin{aligned} V_1 = \frac{1}{n^2}\sum_{\ell=1}^{L_n}\sum_{\ell'=1}^{L_n} k(\ell/L_n)k(\ell'/L_n)\sum_{s=1}^{n-\ell}\sum_{t=1}^{n-\ell'}\{&Cum(Z_s, Z_{s+u}, Z_t, Z_{t+u})V(Z_{s+\ell})V(Z_{t+\ell'}) \\ &+ Cov(Z_s, Z_t)Cov(Z_{s+u}, Z_{t+u})V(Z_{s+\ell})V(Z_{t+\ell'}) \\ &+ Cov(Z_s, Z_{t+u})Cov(Z_{s+u}, Z_t)V(Z_{s+\ell})V(Z_{t+\ell'})\} \\ = O(L_n^2/n).& \end{aligned}$$

By (i) and (ii),

$$\widetilde{M}(1) - \alpha\widetilde{M}(0) - \{M(1) - \alpha M(0)\} = O(L_n^{-q}) + O_p(L_n/\sqrt{n}),$$

which proves (iii). On the other hand, by (i) and (ii) again,

$$MSE[\widetilde{M}(1) - \alpha\widetilde{M}(0)] = O(L_n^{-2q} + L_n^2/n),$$

so that (iv) corresponds to the optimal order when $L_n = O(n^{1/(2+2q)})$. □

We are ready to justify the estimator $\widehat{M}(1) - \widehat{\alpha}_{\#}\widehat{M}(0)$. Note that $\widehat{\alpha}_{\#} - \alpha$ is of the ratio form, i.e.,

$$\widehat{\alpha}_{\#} - \alpha = \frac{O_p(1/\sqrt{n})}{\sigma_Y^2 + O_p(1/\sqrt{n})} = O_p(1/\sqrt{n}).$$

Proposition 6.4. *Suppose that* $0 \le \alpha \le \vartheta < 1 (\vartheta \ne 0)$ *and that* $E(\varepsilon_t^8)$ *exists. If* $\widehat{\alpha} - \alpha = O_p(1/\sqrt{n})$, *then,*

$$\widehat{M}(1) - \widehat{\alpha}\widehat{M}(0) = M(1) - \alpha M(0) + o_p(1),$$

provided that $1/L_n + L_n^{\min(q,2)}/n = o(1)$.

Proof As usual, we decompose

$$\begin{aligned}\widehat{M}(1) - \widehat{\alpha}\widehat{M}(0) - \{M(1) - \alpha M(0)\} &= \{\widehat{M}(1) - \widetilde{M}(1)\} - \widehat{\alpha}\{\widehat{M}(0) - \widetilde{M}(0)\} - (\widehat{\alpha} - \alpha)\widetilde{M}(0) \\ &\quad + [\widetilde{M}(1) - \alpha\widetilde{M}(0) - \{M(1) - \alpha M(0)\}]. \end{aligned} \tag{6.15}$$

It is easy to see that

$$E[|\widetilde{M}(0)|] \le \sum_{\ell=1}^{L_n} E(Z_1{}^2 Z_{1+\ell}^2) \le L_n \mu_{4,Y}$$

by the fourth-order stationarity of $\{Y_t\}$ and Cauchy–Schwarz's inequality, which yields $\widetilde{M}(0) = O_p(L_n)$, hence,

$$(\widehat{\alpha} - \alpha)\widetilde{M}(0) = O_p(L_n/\sqrt{n}). \tag{6.16}$$

On the other hand, for $u = 0, 1$, $\widehat{M}(u) - \widetilde{M}(u)$ can be written as

$$\begin{aligned}\widehat{M}(u) - \widetilde{M}(u) &= \sum_{\ell=1}^{L_n} k(\ell/L_n)\{\widehat{\mu}_Y(u,\ell,\ell) - \widetilde{\mu}_Y(u,\ell,\ell)\} \\ &= \frac{1}{n}\sum_{\ell=1}^{L_n} k(\ell/L_n) \sum_{t=1}^{n-\ell} \{-\overline{Z}(2Z_t Z_{t+u} Z_{t+\ell} + Z_t Z_{t+\ell}^2 + Z_{t+u} Z_{t+\ell}^2) \\ &\qquad + \overline{Z}^2(Z_t Z_{t+u} + 2Z_t Z_{t+\ell} + 2Z_{t+u} Z_{t+\ell} + 2Z_{t+\ell}^2) \\ &\qquad - \overline{Z}^3(Z_t + Z_{t+u} + 2Z_{t+\ell}) + \overline{Z}^4\} \\ &= \sum_{j=1}^{3} W_{n,j}^{[u]}\overline{Z}^j + \sum_{\ell=1}^{L_n} k(\ell/L_n)\Big(1 - \frac{\ell}{n}\Big)\overline{Z}^4 \quad \text{(say)}.\end{aligned}$$

By the strictly stationarity of $\{Y_t\}$, $W_{n,j}^{[u]} = O_p(L_n)$, $j = 1, 2, 3$, are verified, since

$$E(|W_{n,1}^{[u]}|) \le \sum_{\ell=1}^{L_n} \{2E(|Z_1 Z_{1+u}||Z_{1+\ell}|) + E(|Z_1|Z_{1+\ell}^2) + E(|Z_{1+u}|Z_{1+\ell}^2)\} \le 4L_n \sigma_Y \sqrt{\mu_{4,Y}},$$
$$E(|W_{n,2}^{[u]}|) \le \sum_{\ell=1}^{L_n} \{E(|Z_1||Z_{1+u}|) + 2E(|Z_1||Z_{1+\ell}|) + 2E(|Z_{1+u}||Z_{1+\ell}|) + 2E(Z_{1+\ell}^2)\} \le 7L_n \sigma_Y^2,$$
$$E(|W_{n,3}^{[u]}|) \le \sum_{\ell=1}^{L_n} \{E(|Z_1|) + E(|Z_{1+u}|) + 2E(|Z_{1+\ell}|)\} \le 4L_n \sigma_Y$$

(we used Cauchy–Schwarz's inequality). Then,

$$\widehat{M}(u) - \widetilde{M}(u) = O_p(L_n/\sqrt{n}), \quad u = 0, 1, \quad \text{and} \quad \widehat{\alpha}\{\widehat{M}(0) - \widetilde{M}(0)\} = O_p(L_n/\sqrt{n}). \tag{6.17}$$

The result follows from (6.15)–(6.17) and Lemma 6.1(iii). □

Remark 6.2. As is well-known, the choice of the truncation parameter L_n is crucial for this approach. In Section 6.4, some simulation studies will be given by letting $L_n = \lfloor n^{1/(2+2q)} \rfloor$ for the lag window $k(x) = 1 - |x|^q$, $q = 1, 2$, on the basis of Lemma 6.1(iv). We leave a suitable data-driven choice of L_n for future.

Remark 6.3. After completing the paper (Zeng and Kakizawa (2023)), we know the recent work about a novel interesting thinning operation (and model) based on a generalized Neyman type distribution theory (see Amiri et al. (2022)). Since such a model is strictly stationary and ergodic, whose autocorrelation structures are the same as the ADCINAR(1) process, the bias-corrected YW and CLS estimators (6.13) and (6.14) are expected to be applicable even for their model.

6.3.2 Analytical bias-correction

A tedious computation of (6.12) gives the following asymptotic expansion for the bias of $\widehat{\alpha}_{c_1,c_2}$, which is the extension of Proposition 3.2 to the ADCINAR(1) case.

Proposition 6.5. *Suppose that* $0 \le \alpha \le \vartheta < 1 (\vartheta \ne 0)$ *and that* $E(\varepsilon_t^8)$ *exists. The bias of* $\widehat{\alpha}_{c_1,c_2}$ *is given by*

$$E(\widehat{\alpha}_{c_1,c_2} - \alpha) = -\frac{1}{n}\Big(1 + B_{c_1,c_2} + \frac{\vartheta - \alpha}{1 - \alpha\vartheta} D\Big) + o(n^{-1}),$$

where

$$B_{c_1,c_2} = \alpha\Big\{\frac{2\alpha Q_{2:3,Y}}{(1+\alpha)\sigma_Y^4} + \frac{1}{\sigma_Y^2} + 2 + c_1 + c_2\Big\},$$
$$D = \alpha\Big[\frac{\vartheta Q_{2:4,Y}}{\sigma_Y^4} + \frac{\{3 - \alpha - \alpha(1+\alpha)\vartheta\}Q_{2:3,Y}}{(1-\alpha^2)\sigma_Y^4} + \frac{\{(3-\alpha)\vartheta - 2\alpha\}Q_{2:3,Y}\mu_Y}{(1-\alpha)\sigma_Y^4}$$

$$+\frac{(3-\alpha-2\alpha\vartheta)\mu_Y}{(1-\alpha)\sigma_Y^2}+\frac{\{(3+\alpha)\vartheta-2\alpha(3-\alpha)\}\mu_Y^2}{(1-\alpha)\sigma_Y^2}+\frac{3(\vartheta-1)}{\sigma_Y^2}+3\vartheta\Big].$$

Proof After some straightforward but tedious algebra, we obtain

$$M(1)-\alpha M(0)=\frac{2\alpha^2Q_{2:3,Y}}{1+\alpha}+\alpha\sigma_Y^2+2\alpha\sigma_Y^4+\frac{\vartheta-\alpha}{1-\alpha\vartheta}D \quad \text{(by Remark 5.3(ii))}.$$

This, together with (6.11) and (6.12), completes the proof. □

In order to estimate $Q_{2:j,Y}, j=3,4$, we define $\widehat{Q}_{2:3,Y}=\widehat{\kappa}_{3,Y}-\widehat{\sigma}_Y^2$ and $\widehat{Q}_{2:4,Y}=\widehat{\kappa}_{4,Y}-3\widehat{\kappa}_{3,Y}+2\widehat{\sigma}_Y^2$. Then, with the replacement of α,σ_Y^2, and $Q_{2:3,Y}$ by $\widehat{\alpha}_\#,\widehat{\sigma}_Y^2$, and $\widehat{Q}_{2:3,Y}$, we can naturally estimate B_{c_1,c_2} as $\widehat{B}_{c_1,c_2}(\widehat{\alpha}_\#)$.

On the other hand, in order to estimate the additional bias term $(\vartheta-\alpha)D/(1-\alpha\vartheta)$, we further need an estimator for ϑ. Following Nastić et al. (2017), we here take an estimator of ϑ, given by

$$\widehat{\vartheta}(\widehat{\alpha}_\#)=\frac{\widehat{C}_Y-\widehat{\alpha}_\#\{1+2(1-\widehat{\alpha}_\#)\overline{Y}\}\widehat{\sigma}_Y^2}{\widehat{\alpha}_\#(\widehat{Q}_{2:3,Y}+2\widehat{\sigma}_Y^2\overline{Y})},$$

where

$$\widehat{C}_Y=\frac{1}{n-1}\sum_{t=2}^{n}Y_t^2Y_{t-1}-\frac{1}{n-1}\sum_{t=2}^{n}Y_t^2\frac{1}{n-1}\sum_{t=2}^{n}Y_{t-1}.$$

Similarly, with the replacement of $\alpha,\vartheta,\mu_Y,\sigma_Y^2$, and $Q_{2:j,Y}$ with $\widehat{\alpha}_\#,\widehat{\vartheta}(\widehat{\alpha}_\#),\overline{Y},\widehat{\sigma}_Y^2$, and $\widehat{Q}_{2:j,Y}$, $j=3,4$, we can naturally estimate D as $\widehat{D}_\#$.

Lemma 6.2. *Suppose that* $0<\alpha\le\vartheta<1$ *and that* $E(\varepsilon_t^4)$ *exists. Then,*

$$\widehat{B}_{c_1,c_2}(\widehat{\alpha}_\#)\xrightarrow{a.s.}B_{c_1,c_2},\ \ \widehat{\vartheta}(\widehat{\alpha}_\#)\xrightarrow{a.s.}\vartheta,$$

and

$$\frac{\{\widehat{\alpha}_\#-\widehat{\vartheta}(\widehat{\alpha}_\#)\}\widehat{D}_\#}{1-\widehat{\alpha}_\#\widehat{\vartheta}(\widehat{\alpha}_\#)}\xrightarrow{a.s.}\frac{(\alpha-\vartheta)D}{1-\alpha\vartheta}.$$

Proof We have only to the strong consistency $\widehat{\vartheta}_\#\xrightarrow{a.s.}\vartheta$.

By Lemma 5.5(i) and Proposition 6.1(i), we have $\overline{Y}\xrightarrow{a.s.}\mu_Y$, $\widehat{\sigma}_Y^2\xrightarrow{a.s.}\sigma_Y^2$, $\widehat{\kappa}_{j,Y}\xrightarrow{a.s.}\kappa_{j,Y}$, $j=3,4$, and $\widehat{\alpha}_\#\xrightarrow{a.s.}\alpha$. By strictly stationarity and ergodicity of $\{Y_t\}$, one can show that

$$\widehat{C}_Y\xrightarrow{a.s.}Cov(Y_2^2,Y_1)=\gamma_Y(1,1)+2\gamma_Y(1)\mu_Y=\alpha[\vartheta(Q_{2:3,Y}+2\sigma_Y^2\mu_Y)+\{1+2(1-\alpha)\mu_Y\}\sigma_Y^2].$$

Hence,

$$\widehat{\vartheta}(\widehat{\alpha}_{\#}) \xrightarrow{a.s.} \frac{Cov(Y_2^2, Y_1) - \alpha\{1 + 2(1-\alpha)\mu_Y\}\sigma_Y^2}{\alpha(Q_{2:3,Y} + 2\sigma_Y^2\mu_Y)} = \vartheta. \quad \square$$

In this way, analytical bias-corrected YW and CLS estimators for the parameter α in the stationary ADCINAR(1) process are defined by

$$\widetilde{\alpha}_{YW}^{\text{Ana}} = \widehat{\alpha}_{YW} + \frac{1}{n}\Big\{1 + \widehat{B}_{1,1}(\widehat{\alpha}_{YW}) + \frac{\widehat{\vartheta}(\widehat{\alpha}_{YW}) - \widehat{\alpha}_{YW}}{1 - \widehat{\vartheta}(\widehat{\alpha}_{YW})\widehat{\alpha}_{YW}}\widehat{D}_{YW}\Big\}, \tag{6.18}$$

$$\widetilde{\alpha}_{CLS}^{\text{Ana}} = \widehat{\alpha}_{CLS} + \frac{1}{n}\Big\{1 + \widehat{B}_{1,0}(\widehat{\alpha}_{CLS}) + \frac{\widehat{\vartheta}(\widehat{\alpha}_{CLS}) - \widehat{\alpha}_{CLS}}{1 - \widehat{\vartheta}(\widehat{\alpha}_{CLS})\widehat{\alpha}_{CLS}}\widehat{D}_{CLS}\Big\}. \tag{6.19}$$

Remark 6.4. (i) For the stationary INAR(1) process under a general innovation, analytical bias-corrected YW and CLS estimators were suggested in Chapter 3, as follows:

$$\widetilde{\alpha}_{YW}^{\prime\text{Ana}} = \widehat{\alpha}_{YW} + \frac{1}{n}\{1 + \widehat{B}_{1,1}(\widehat{\alpha}_{YW})\},$$

$$\widetilde{\alpha}_{CLS}^{\prime\text{Ana}} = \widehat{\alpha}_{CLS} + \frac{1}{n}\{1 + \widehat{B}_{1,0}(\widehat{\alpha}_{CLS})\}.$$

Under the Poisson innovation (in this case, $Q_{2:3,Y} = 0$), one can use

$$\widetilde{\alpha}_{YW}^{\prime\prime\text{Ana}} = \widehat{\alpha}_{YW} + \frac{1}{n}\Big(1 + \frac{\widehat{\alpha}_{YW}}{\widehat{\sigma}_Y^2} + 4\widehat{\alpha}_{YW}\Big),$$

$$\widetilde{\alpha}_{CLS}^{\prime\prime\text{Ana}} = \widehat{\alpha}_{CLS} + \frac{1}{n}\Big(1 + \frac{\widehat{\alpha}_{CLS}}{\widehat{\sigma}_Y^2} + 3\widehat{\alpha}_{CLS}\Big).$$

(ii) Unlike the lag window-type bias-corrections (6.13) and (6.14), the analytical bias-corrections in this subsection suffer from model misspecification problems (i.e., the under-fitting or over-fitting), depending on the situation without/with the constraint $\vartheta = \alpha$, $Q_{2:3,Y} = 0$, or $Q_{2:4,Y} = 0$.

6.4 Simulation results

Simulation studies of 2CLS estimator for the new parameter ϑ

We conduct simulation experiments ($\alpha_0 = 0.3, 0.4, 0.5, 0.6, 0.7$, $\vartheta_0 = 0.9$, and $n = 100, 200, 300, 1000$) about the infeasible/feasible estimators for ϑ, under the set-up that $\{\varepsilon_t\}$ follows the Poisson distribution $\text{Po}((1-\alpha)\mu)$ or negative binomial (NB) distribution $\text{NB}(r, r/\{r + (1-\alpha)\mu\})$ (we set $\mu = 10$ and $r = 10$). We consider $\widehat{\vartheta}_0 = \widehat{\vartheta}(\alpha_0, \mu_{Y,0}, \sigma_{Y,0}^2)$ as if the parameter $(\alpha, \mu_Y, \sigma_Y^2) = (\alpha_0, \mu_{Y,0}, \sigma_{Y,0}^2)$ were known, together with the feasible estimator $\widehat{\vartheta}_{2CLS} = \widehat{\vartheta}(\widehat{\alpha}_{CLS}, \overline{Y}, \widehat{\sigma}_Y^2)$. To keep the inequality constraint, we also compute

$$\widehat{\vartheta}_0^{\dagger} = \begin{cases} \alpha_0, & \widehat{\vartheta}(\alpha_0, \mu_{Y,0}, \sigma_{Y,0}^2) \leq \alpha_0, \\ \widehat{\vartheta}(\alpha_0, \mu_{Y,0}, \sigma_{Y,0}^2), & \alpha_0 < \widehat{\vartheta}(\alpha_0, \mu_{Y,0}, \sigma_{Y,0}^2) < 1, \\ 1, & 1 \leq \widehat{\vartheta}(\alpha_0, \mu_{Y,0}, \sigma_{Y,0}^2), \end{cases}$$

$$\widehat{\vartheta}_{2CLS}^{\dagger} = \begin{cases} \widehat{\alpha}^{\dagger}, & \widehat{\vartheta}(\widehat{\alpha}^{\dagger}, \overline{Y}, \widehat{\sigma}_Y^2) \leq \widehat{\alpha}^{\dagger}, \\ \widehat{\vartheta}(\widehat{\alpha}^{\dagger}, \overline{Y}, \widehat{\sigma}_Y^2), & \widehat{\alpha}^{\dagger} < \widehat{\vartheta}(\widehat{\alpha}^{\dagger}, \overline{Y}, \widehat{\sigma}_Y^2) < 1, \\ 1, & 1 \leq \widehat{\vartheta}(\widehat{\alpha}^{\dagger}, \overline{Y}, \widehat{\sigma}_Y^2), \end{cases}$$

where $\widehat{\alpha}^{\dagger} = \max\{0, \min(1, \widehat{\alpha}_{\text{CLS}})\}$ is the positive part of the truncated CLS estimator for α. Note that the former $\widehat{\vartheta}_0^{\dagger}$ is infeasible, whereas the latter $\widehat{\vartheta}_{2CLS}^{\dagger}$ is feasible.

We discuss the simulation results (2000 replications) for the Poisson innovation. The similar findings shown for the NB innovation, as well as the estimators with $\widehat{\alpha}_{CLS}$ replaced by $\widehat{\alpha}_{YW}$, are omitted to save space. It is obvious from Table 6.1 that the biases, variances, and mean standard errors (MSEs) of all estimators decrease as n increases. More details are as follows.

(A) In our experiance, it seems that, the more α_0 approches to zero, the less numerically stable the estimators for ϑ are. The bad performance of $\widehat{\vartheta}_{2CLS}$ when α_0 is small, compared to $\widehat{\vartheta}_0$ (the mark # means that the variance and MSE of $\widehat{\vartheta}_{2CLS}$ are too large) is caused by the fact that, for small sample size n, (i) the CLS (or YW) estimator for α can be negative value or greater than 1, and (ii) the resulting 2CLS estimator for ϑ can be smaller than the estimator for α (of course, the consistency tells us that these phenomena do not occur in the large sample). We conclude that the estimator $\widehat{\vartheta}_{2CLS}$ is dangerous to use in practice, unless the sample size n is moderately large.

(B) We thus compare $\widehat{\vartheta}_0^{\dagger}$ with $\widehat{\vartheta}_{2CLS}^{\dagger}$. We find that, unless the sample size n is large, the information $(\alpha, \mu_Y, \sigma_Y^2) = (\alpha_0, \mu_{Y,0}, \sigma_{Y,0}^2)$ does not always yield good estimator, depending on the values of α_0, i.e., when α_0 is small and n is small, the variance of $\widehat{\vartheta}_0^{\dagger}$ is smaller than that of $\widehat{\vartheta}_{2CLS}^{\dagger}$. On the other hand, the simulation results when $n = 1000$ reveal that feasible estimator $\widehat{\vartheta}_{2CLS}^{\dagger}$ is superior to infeasible estimator $\widehat{\vartheta}_0^{\dagger}$.

(C) To illusrate the finding of (B) theoretically, recall Remark 6.1, i.e., the key quantity is $\mathcal{D}$, for which the explicit but tedius formulas of A, B, and ω_{11} are available (we can apply Proposition 5.1), whereas the expectation ω_{12} (see (6.4)) contains $E[(Y_1 - \mu_Y)^i (Y_0 - \mu_Y)^j], i + j = 5, 6 (i = 0, 1, 2, 3)$; they are difficult to evaluate. After estimating $\mathcal{D}$ via the simulation experiment, it turns out that $\mathcal{D}$ is almost negative, at least for the Poisson and NB innovations. This is the reason why the plug-in case seems to be superior to the known case, for large sample size $n = 1000$ (say); in that case, there will be no worry about (A).

Table 6.1: Biases, variances, and MSEs of infeasible estimators ($\widehat{\vartheta}_0$ and $\widehat{\vartheta}_0^{\dagger}$) and feasible estimators ($\widehat{\vartheta}_{2CLS}$ and $\widehat{\vartheta}_{2CLS}^{\dagger}$) in the stationary ADCINAR(1) process ($\vartheta_0 = 0.9$) under the Poisson and NB innovations. The mark # means that the variance and MSE of $\widehat{\vartheta}_{2CLS}$ are too large for small sample size n. The bold-faced values show that $\widehat{\vartheta}_0^{\dagger}$ is superior to $\widehat{\vartheta}_{2CLS}^{\dagger}$.

Poisson case		Biases (×10)				Variances (×100)				MSEs (×100)			
α_0	n	$\widehat{\vartheta}_0$	$\widehat{\vartheta}_{2CLS}$	$\widehat{\vartheta}_0^{\dagger}$	$\widehat{\vartheta}_{2CLS}^{\dagger}$	$\widehat{\vartheta}_0$	$\widehat{\vartheta}_{2CLS}$	$\widehat{\vartheta}_0^{\dagger}$	$\widehat{\vartheta}_{2CLS}^{\dagger}$	$\widehat{\vartheta}_0$	$\widehat{\vartheta}_{2CLS}$	$\widehat{\vartheta}_0^{\dagger}$	$\widehat{\vartheta}_{2CLS}^{\dagger}$
0.3	100	-0.796	0.563	-0.880	-0.482	1.809	#	**1.367**	1.886	2.443	#	2.142	2.118
	200	-0.531	-0.090	-0.601	-0.341	1.301	#	0.991	0.919	1.582	#	1.352	1.036
	300	-0.423	-0.197	-0.484	-0.282	1.082	1.137	0.830	0.635	1.261	1.176	1.064	0.714
	1000	-0.172	-0.131	-0.198	-0.138	0.499	0.234	0.416	0.210	0.529	0.251	0.455	0.229
0.4	100	-0.279	0.118	-0.296	-0.242	0.567	#	**0.513**	0.869	0.645	#	**0.601**	0.927
	200	-0.160	-0.049	-0.166	-0.131	0.324	#	**0.308**	0.393	0.350	#	**0.335**	0.410
	300	-0.146	-0.112	-0.151	-0.124	0.258	0.321	**0.245**	0.282	0.279	0.334	**0.268**	0.298
	1000	-0.051	-0.050	-0.051	-0.050	0.097	0.087	0.096	0.087	0.100	0.089	0.099	0.089
0.5	100	0.063	0.013	0.044	-0.074	0.393	#	**0.346**	0.471	0.397	#	**0.348**	0.477
	200	0.060	-0.017	0.054	-0.024	0.230	0.244	**0.215**	0.227	0.234	0.245	**0.218**	0.227
	300	0.029	-0.030	0.027	-0.031	0.172	0.168	0.168	0.166	0.173	0.168	0.169	0.167
	1000	0.011	-0.012	0.011	-0.012	0.055	0.047	0.055	0.047	0.055	0.048	0.055	0.048
0.6	100	0.290	0.067	0.228	0.046	0.525	0.393	0.393	0.328	0.609	0.397	0.445	0.330
	200	0.152	0.030	0.130	0.026	0.358	0.185	0.310	0.176	0.381	0.185	0.327	0.177
	300	0.121	0.026	0.115	0.026	0.234	0.116	0.221	0.115	0.249	0.117	0.234	0.116
	1000	0.040	0.013	0.040	0.013	0.090	0.037	0.090	0.037	0.092	0.037	0.092	0.037
0.7	100	0.307	0.059	0.226	0.046	0.603	0.286	0.425	0.255	0.697	0.289	0.476	0.257
	200	0.181	0.052	0.151	0.049	0.392	0.143	0.327	0.138	0.425	0.146	0.350	0.141
	300	0.134	0.043	0.121	0.042	0.293	0.107	0.266	0.106	0.311	0.108	0.280	0.108
	1000	0.063	0.018	0.063	0.018	0.096	0.032	0.096	0.032	0.100	0.033	0.100	0.033
NB case													
0.3	100	-0.826	-0.377	-0.934	-0.689	2.228	#	**1.646**	2.447	2.911	#	**2.518**	2.922
	200	-0.524	-0.390	-0.623	-0.408	1.619	#	1.160	1.097	1.894	#	1.548	1.264
	300	-0.437	-0.215	-0.512	-0.335	1.225	1.610	0.914	0.802	1.416	1.656	1.176	0.914
	1000	-0.174	-0.131	-0.210	-0.140	0.583	0.321	0.460	0.298	0.613	0.339	0.504	0.317
0.4	100	-0.336	-0.040	-0.358	-0.366	0.740	#	**0.664**	1.082	0.853	#	**0.792**	1.215
	200	-0.209	-0.170	-0.218	-0.214	0.430	0.784	**0.404**	0.526	0.474	0.813	**0.451**	0.572
	300	-0.149	-0.146	-0.157	-0.161	0.345	0.415	**0.323**	0.363	0.367	0.436	**0.348**	0.389
	1000	-0.071	-0.076	-0.072	-0.077	0.126	0.112	0.125	0.112	0.131	0.118	0.130	0.118
0.5	100	0.038	-0.016	-0.0001	-0.127	0.536	#	**0.437**	0.588	0.537	#	**0.437**	0.604
	200	0.038	-0.051	0.028	-0.064	0.289	0.338	**0.265**	0.302	0.291	0.341	**0.266**	0.306
	300	0.022	-0.051	0.019	-0.054	0.204	0.210	0.197	0.203	0.205	0.213	0.197	0.206
	1000	0.008	-0.018	0.008	-0.018	0.072	0.065	0.072	0.065	0.072	0.065	0.072	0.065
0.6	100	0.231	0.003	0.162	-0.022	0.595	0.461	0.436	0.394	0.649	0.461	0.462	0.394
	200	0.178	0.030	0.149	0.024	0.389	0.226	0.323	0.211	0.420	0.227	0.345	0.212
	300	0.111	0.028	0.095	0.026	0.296	0.159	0.260	0.154	0.308	0.160	0.269	0.155
	1000	0.045	0.014	0.045	0.014	0.106	0.052	0.106	0.052	0.108	0.052	0.108	0.052
0.7	100	0.275	0.044	0.187	0.026	0.653	0.339	0.450	0.293	0.728	0.341	0.484	0.293
	200	0.207	0.057	0.174	0.055	0.399	0.170	0.328	0.164	0.442	0.174	0.358	0.167
	300	0.155	0.037	0.141	0.036	0.293	0.117	0.263	0.115	0.317	0.119	0.283	0.116
	1000	0.060	0.024	0.060	0.024	0.111	0.038	0.110	0.038	0.115	0.039	0.114	0.039

Simulation studies of bias-corrected estimators for the parameter α

We here conduct some simulations about the YW and CLS estimators (without/with bias-correction) for the parameter α in the stationary INAR(1) and ADCINAR(1) processes (we set $\vartheta = 0.9$ for the ADCINAR(1) case), under the set-up that the innovation $\{\varepsilon_t\}$ follows the Poisson distribution $\mathrm{Po}((1-\alpha)\mu)$ or NB distribution $\mathrm{NB}(r, r/\{r + (1-\alpha)\mu\})$ (we set $\mu = 10$ and $r = 10$). We consider $\alpha = 0.2, 0.5, 0.8$ and $n = 100, 200, 300$, with 2000 replications. Note that the INAR(1) process is a special case of the ADCINAR(1) process when $\alpha = \vartheta$ (see Remark 5.1(ii)). We use the lag windows $k(x) = 1 - |x|^q$ for $q = 1, 2$. The resulting bias-corrected estimator, with $L_n = \lfloor n^{1/(2+2q)} \rfloor$, is denoted by $\widetilde{\alpha}_{\#}^{(q)}$, where $\# = YW, CLS$ and $q = 1, 2$. By way of comparison, we also compute the analytical bias-corrected estimators $\widetilde{\alpha}_{\#}^{\mathrm{Ana}}$, $\widetilde{\alpha}_{\#}^{\dagger\mathrm{Ana}}$, and $\widetilde{\alpha}_{\#}^{\prime\mathrm{Ana}}$. Here,

$$\widetilde{\alpha}_{YW}^{\dagger\mathrm{Ana}} = \widehat{\alpha}_{YW}^{\dagger} + \frac{1}{n}\left\{1 + \widehat{B}_{1,1}(\widehat{\alpha}_{YW}^{\dagger}) + \frac{\widehat{\vartheta}^{\dagger}(\widehat{\alpha}_{YW}^{\dagger}) - \widehat{\alpha}_{YW}^{\dagger}}{1 - \widehat{\vartheta}^{\dagger}(\widehat{\alpha}_{YW}^{\dagger})\widehat{\alpha}_{YW}^{\dagger}} \widehat{D}_{YW}^{\dagger}\right\}, \tag{6.20}$$

$$\widetilde{\alpha}_{CLS}^{\dagger\mathrm{Ana}} = \widehat{\alpha}_{CLS}^{\dagger} + \frac{1}{n}\left\{1 + \widehat{B}_{1,0}(\widehat{\alpha}_{CLS}^{\dagger}) + \frac{\widehat{\vartheta}^{\dagger}(\widehat{\alpha}_{CLS}^{\dagger}) - \widehat{\alpha}_{CLS}^{\dagger}}{1 - \widehat{\vartheta}^{\dagger}(\widehat{\alpha}_{CLS}^{\dagger})\widehat{\alpha}_{CLS}^{\dagger}} \widehat{D}_{CLS}^{\dagger}\right\}, \tag{6.21}$$

where $\widehat{\alpha}_{\#}^{\dagger} = \max\{0, \min(1, \widehat{\alpha}_{\#})\}$ is the positive part of the truncated estimator $\widehat{\alpha}_{\#}$,

$$\widehat{\vartheta}^{\dagger}(\widehat{\alpha}_{\#}^{\dagger}) = \begin{cases} \widehat{\alpha}_{\#}^{\dagger}, & \widehat{\vartheta}(\widehat{\alpha}_{\#}^{\dagger}) \le \widehat{\alpha}_{\#}^{\dagger}, \\ \widehat{\vartheta}(\widehat{\alpha}_{\#}^{\dagger}), & \widehat{\alpha}_{\#}^{\dagger} < \widehat{\vartheta}(\widehat{\alpha}_{\#}^{\dagger}) < 1, \\ 1, & 1 \le \widehat{\vartheta}(\widehat{\alpha}_{\#}^{\dagger}), \end{cases}$$

and $\widehat{D}_{\#}^{\dagger}$ is an estimator of D, replacing $\alpha, \vartheta, \mu_Y, \sigma_Y^2$, and $Q_{2:j,Y}$ by $\widehat{\alpha}_{\#}^{\dagger}, \widehat{\vartheta}^{\dagger}(\widehat{\alpha}_{\#}^{\dagger}), \overline{Y}, \widehat{\sigma}_Y^2$, and $\widehat{Q}_{2:j,Y}$, $j = 3, 4$, respectively.

The biases, variances, and MSEs of the estimators $\widehat{\alpha}_{\#}$, $\widetilde{\alpha}_{\#}^{(q)}$, $q = 1, 2$, $\widetilde{\alpha}_{\#}^{\mathrm{Ana}}$, $\widetilde{\alpha}_{\#}^{\dagger\mathrm{Ana}}$, and $\widetilde{\alpha}_{\#}^{\prime\mathrm{Ana}}$ are shown in Tables 6.2–6.5. It is obvious that the biases, variances, and MSEs of all estimators decrease as n increases (overall, the estimators have the downward biases). More details are as follows.

- Remarkably, the bias-corrections effectively reduce the bias of the estimator $\widehat{\alpha}_{\#}$, except that, as pointed out in Remark 6.4(ii), careful attention must be paid to misspecification problems. As a serious matter, we see that $\widetilde{\alpha}_{\#}^{\prime\mathrm{Ana}}$ for the under-fitted case (Tables 6.4 and 6.5) is still biased when α is small, due to the (incorrect) analytical bias-correction. On the other hand, the bias of $\widetilde{\alpha}_{\#}^{\prime\mathrm{Ana}}$ for the over-fitted case (Tables 6.2 and 6.3) seems to be almost removed similar to other bias-corrections, whereas $\widetilde{\alpha}_{\#}^{\prime\mathrm{Ana}}$ in this case underperforms $\widetilde{\alpha}_{\#}^{(q)}$ in terms of the variance and MSE.

- The analytical bias-corrected estimator $\widetilde{\alpha}_{\#}^{\mathrm{Ana}}$ ((6.18) or (6.19)) is extremely dangerous to use when $\alpha = 0.2$ and n is small (see Zeng and Kakizawa (2023)), since, as we saw before, estimating ϑ is unstable unless n is moderately large. This is the reason why (6.20) and (6.21) are employed here (note that (6.20) and (6.21) are improved but still dangerous to use when α is small; see Tables 6.4 and 6.5).

- In terms of the variance and MSE, the analytical bias-corrected estimator $\widetilde{\alpha}_{\#}^{\dagger\mathrm{Ana}}$ outperforms the lag window-type bias-corrected estimator $\widetilde{\alpha}_{\#}^{(q)}$ when n is small for the INAR(1) process (Tables 6.2 and 6.3), whereas $\widetilde{\alpha}_{\#}^{(q)}$ is superior to $\widetilde{\alpha}_{\#}^{\dagger\mathrm{Ana}}$ for the ADCINAR(1) process (Tables 6.4 and 6.5).
- If α is large, then, the bias-corrected estimator $\widetilde{\alpha}_{\#}^{(q)}$ outperforms the (bias-uncorrected) estimator $\widehat{\alpha}_{\#}$ in terms of the MSE.
- Although the variance of $\widetilde{\alpha}_{\#}^{(2)}$ is overall smaller than that of $\widetilde{\alpha}_{\#}^{(1)}$, their performances are comparable.

6.5 Concluding remarks

We have revisited the 2CLS estimator for the new parameter ϑ in the stationary ADCINAR(1) process under a general innovation. It turns out that plugging a certain $n^{1/2}$-consistent estimator $\widehat{\alpha}$ for α (see (6.6)) yields the additional term $n^{-1/2}\sum_{t=2}^{n} AM_{\alpha,t}/(\alpha B)$ to the leading term of the stochastic expansion given in Proposition 6.3(i), whereas plugging $(\overline{Y}, \widehat{\sigma}_Y^2)$ for (μ_Y, σ_Y^2) has no effect. The finding about the 2CLS approach seemed to be new in the recent literature of the stationary INAR-type process, to the best of our knowledge.

We have developed two kinds of the bias-corrections of the YW and CLS estimators. One is the lag window-type bias-correction ((6.13) and (6.14)) and the other is the analytical bias-correction ((6.18) and (6.19)). Although the data-driven choice of the truncation parameter L_n remains to be explored (in the simulations, we have used $L_n = \lfloor n^{1/(2+2q)} \rfloor$, $q = 1, 2$, as a first step), the merit of the lag window-type bias-correction is that there is no need for the computation of the closed-form expression of the biases of $\widehat{\alpha}_{YW}$ and $\widehat{\alpha}_{CLS}$. On the other hand, the analytical bias-corrected YW and CLS estimators, involving the estimation for the parameter ϑ, are more complicated than expected; besides, the resulting estimators have, of course, the model misspecification problems.

6.6 Proof of Proposition 6.3(i)

We notice that

$$\sqrt{n}(\widehat{\vartheta}_{2CLS} - \vartheta) = \frac{\dfrac{1}{\sqrt{n}}\displaystyle\sum_{t=2}^{n}\{g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) - \vartheta\widehat{\alpha} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)\} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)}{\widehat{\alpha}\{B + o_p(1)\}} \tag{6.22}$$

(see (6.2)). Below, we will deal with the numerator of (6.22).

Table 6.2: Biases, variances, and MSEs of the estimators $\widehat{\alpha}_{YW}$ (without bias-correction), $\widetilde{\alpha}^{(q)}_{YW}$, $q = 1, 2$, $\widetilde{\alpha}^{\text{Ana}}_{YW}$, $\widetilde{\alpha}^{\dagger\text{Ana}}_{YW}$, and $\widetilde{\alpha}'^{\text{Ana}}_{YW}$ in the stationary INAR(1) process under the Poisson and NB innovations (note that $\widetilde{\alpha}'^{\text{Ana}}_{YW}$ is used in a over-fitted case).

Poisson Case		Biases (×10)			Variances (×100)			MSEs (×100)		
$\alpha(=\vartheta)$	n	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}^{(1)}_{YW}$	$\widetilde{\alpha}^{(2)}_{YW}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}^{(1)}_{YW}$	$\widetilde{\alpha}^{(2)}_{YW}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}^{(1)}_{YW}$	$\widetilde{\alpha}^{(2)}_{YW}$
0.2	100	-0.200	-0.042	-0.039	0.968	1.032	1.035	1.008	1.034	1.036
	200	-0.091	-0.009	-0.008	0.475	0.491	0.492	0.483	0.491	0.492
	300	-0.068	-0.013	-0.013	0.319	0.326	0.326	0.323	0.326	0.326
0.5	100	-0.325	-0.081	-0.080	0.781	0.826	0.822	0.887	0.832	0.829
	200	-0.151	-0.026	-0.025	0.397	0.408	0.407	0.420	0.409	0.408
	300	-0.123	-0.035	-0.038	0.268	0.274	0.273	0.283	0.275	0.274
0.8	100	-0.452	-0.152	-0.158	0.507	0.520	0.518	0.711	0.543	0.543
	200	-0.195	-0.042	-0.045	0.206	0.208	0.207	0.244	0.210	0.209
	300	-0.138	-0.030	-0.038	0.138	0.139	0.139	0.157	0.140	0.140
		$\widetilde{\alpha}^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{YW}$	$\widetilde{\alpha}'^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{YW}$	$\widetilde{\alpha}'^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{YW}$	$\widetilde{\alpha}'^{\text{Ana}}_{YW}$
0.2	100	-0.016	0.001	-0.027	1.053	0.967	1.049	1.054	0.967	1.050
	200	0.002	0.002	-0.001	0.494	0.491	0.495	0.494	0.491	0.495
	300	-0.007	-0.007	-0.009	0.327	0.327	0.327	0.327	0.327	0.327
0.5	100	-0.025	-0.025	-0.033	0.840	0.840	0.847	0.841	0.841	0.848
	200	0.0003	0.0003	-0.002	0.412	0.412	0.414	0.412	0.412	0.414
	300	-0.022	-0.022	-0.023	0.275	0.275	0.276	0.275	0.275	0.276
0.8	100	-0.032	-0.032	-0.044	0.533	0.533	0.549	0.534	0.534	0.551
	200	0.019	0.019	0.015	0.210	0.210	0.214	0.210	0.210	0.214
	300	0.005	0.005	0.002	0.140	0.140	0.142	0.140	0.140	0.142
NB Case										
$\alpha(=\vartheta)$	n	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}^{(1)}_{YW}$	$\widetilde{\alpha}^{(2)}_{YW}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}^{(1)}_{YW}$	$\widetilde{\alpha}^{(2)}_{YW}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}^{(1)}_{YW}$	$\widetilde{\alpha}^{(2)}_{YW}$
0.2	100	-0.197	-0.039	-0.036	0.960	1.028	1.032	0.999	1.030	1.033
	200	-0.078	0.003	0.004	0.474	0.492	0.493	0.481	0.492	0.493
	300	-0.065	-0.010	-0.010	0.315	0.324	0.323	0.319	0.324	0.323
0.5	100	-0.303	-0.059	-0.058	0.818	0.868	0.865	0.911	0.872	0.868
	200	-0.126	-0.0005	-0.0003	0.377	0.388	0.387	0.393	0.388	0.387
	300	-0.108	-0.021	-0.024	0.265	0.272	0.271	0.277	0.272	0.271
0.8	100	-0.435	-0.135	-0.141	0.485	0.497	0.496	0.674	0.516	0.516
	200	-0.212	-0.059	-0.063	0.212	0.214	0.214	0.257	0.218	0.218
	300	-0.136	-0.028	-0.036	0.137	0.138	0.138	0.155	0.139	0.139
		$\widetilde{\alpha}^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{YW}$	$\widetilde{\alpha}'^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{YW}$	$\widetilde{\alpha}'^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\text{Ana}}_{YW}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{YW}$	$\widetilde{\alpha}'^{\text{Ana}}_{YW}$
0.2	100	-0.006	0.001	-0.023	1.041	0.976	1.041	1.041	0.976	1.041
	200	0.015	0.014	0.011	0.491	0.493	0.494	0.491	0.493	0.494
	300	-0.004	-0.004	-0.006	0.324	0.324	0.324	0.324	0.324	0.324
0.5	100	0.0005	0.0004	-0.011	0.881	0.882	0.887	0.881	0.882	0.887
	200	0.027	0.027	0.024	0.390	0.390	0.392	0.391	0.391	0.393
	300	-0.006	-0.006	-0.008	0.272	0.272	0.273	0.272	0.272	0.273
0.8	100	-0.013	-0.013	-0.025	0.513	0.513	0.525	0.514	0.514	0.526
	200	0.005	0.005	-0.003	0.215	0.215	0.221	0.215	0.215	0.221
	300	0.008	0.008	0.005	0.138	0.138	0.140	0.138	0.138	0.140

Table 6.3: Biases, variances, and MSEs of the estimators $\widehat{\alpha}_{CLS}$ (without bias-correction), $\widetilde{\alpha}^{(q)}_{CLS}$, $q = 1, 2$, $\widetilde{\alpha}^{\text{Ana}}_{CLS}$, $\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$, and $\widetilde{\alpha}'^{\text{Ana}}_{CLS}$ in the stationary INAR(1) process under the Poisson and NB innovations (note that $\widetilde{\alpha}'^{\text{Ana}}_{YW}$ is used in a over-fitted case).

Poission case		Biases (×10)			Variances (×100)			MSEs (×100)		
$\alpha(=\vartheta)$	n	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$
0.2	100	-0.183	-0.042	-0.040	0.985	1.030	1.032	1.018	1.031	1.033
	200	-0.081	-0.009	-0.008	0.480	0.491	0.492	0.486	0.491	0.492
	300	-0.062	-0.013	-0.013	0.321	0.326	0.326	0.325	0.326	0.326
0.5	100	-0.274	-0.077	-0.076	0.796	0.824	0.821	0.871	0.830	0.826
	200	-0.127	-0.026	-0.026	0.399	0.406	0.405	0.415	0.407	0.406
	300	-0.105	-0.034	-0.037	0.270	0.274	0.273	0.281	0.275	0.274
0.8	100	-0.368	-0.144	-0.149	0.506	0.508	0.507	0.641	0.529	0.529
	200	-0.155	-0.041	-0.044	0.206	0.206	0.206	0.230	0.208	0.208
	300	-0.110	-0.029	-0.036	0.138	0.138	0.138	0.150	0.139	0.139
		$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$
0.2	100	-0.015	0.001	-0.026	1.052	0.966	1.047	1.052	0.966	1.047
	200	0.002	0.003	-0.001	0.494	0.491	0.495	0.494	0.491	0.495
	300	-0.007	-0.008	-0.009	0.327	0.327	0.327	0.327	0.327	0.328
0.5	100	-0.012	-0.012	-0.028	0.844	0.844	0.845	0.844	0.844	0.846
	200	0.002	0.002	-0.002	0.410	0.410	0.412	0.410	0.410	0.412
	300	-0.019	-0.019	-0.021	0.274	0.274	0.275	0.275	0.275	0.276
0.8	100	0.033	0.033	-0.032	0.548	0.548	0.538	0.549	0.549	0.539
	200	0.032	0.032	0.017	0.210	0.210	0.213	0.211	0.211	0.213
	300	0.012	0.012	0.004	0.139	0.139	0.141	0.139	0.139	0.141
NB case										
$\alpha(=\vartheta)$	n	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$
0.2	100	-0.178	-0.038	-0.035	0.979	1.028	1.031	1.011	1.029	1.032
	200	-0.069	0.003	0.004	0.479	0.491	0.492	0.483	0.491	0.492
	300	-0.059	-0.010	-0.010	0.317	0.323	0.323	0.320	0.323	0.323
0.5	100	-0.253	-0.056	-0.055	0.830	0.863	0.860	0.894	0.866	0.863
	200	-0.101	0.0001	0.0004	0.379	0.387	0.386	0.389	0.387	0.386
	300	-0.091	-0.020	-0.023	0.267	0.271	0.270	0.275	0.272	0.271
0.8	100	-0.352	-0.129	-0.134	0.483	0.486	0.485	0.607	0.503	0.503
	200	-0.172	-0.058	-0.061	0.211	0.211	0.211	0.240	0.215	0.215
	300	-0.109	-0.027	-0.035	0.135	0.136	0.135	0.147	0.136	0.137
		$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$
0.2	100	-0.004	0.003	-0.021	1.041	0.977	1.041	1.041	0.977	1.041
	200	0.016	0.014	0.011	0.490	0.492	0.493	0.490	0.492	0.494
	300	-0.004	-0.004	-0.006	0.323	0.324	0.323	0.323	0.324	0.324
0.5	100	0.011	0.011	-0.006	0.879	0.879	0.883	0.879	0.879	0.883
	200	0.029	0.029	0.025	0.389	0.389	0.391	0.390	0.390	0.392
	300	-0.005	-0.005	-0.007	0.271	0.271	0.272	0.271	0.271	0.272
0.8	100	0.042	0.042	-0.016	0.530	0.530	0.513	0.532	0.532	0.513
	200	0.017	0.017	-0.0004	0.213	0.213	0.217	0.213	0.213	0.217
	300	0.014	0.014	0.006	0.136	0.136	0.138	0.136	0.136	0.138

Table 6.4: Biases, variances, and MSEs of the estimators $\widehat{\alpha}_{YW}$ (without bias-correction), $\widetilde{\alpha}_{YW}^{(q)}$, $q = 1, 2$, $\widetilde{\alpha}_{YW}^{\text{Ana}}$, $\widetilde{\alpha}_{YW}^{\dagger\text{Ana}}$, and $\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$ in the stationary ADCINAR(1) process ($\vartheta = 0.9$) under the Poisson and NB innovations (note that $\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$ is used in a under-fitted case).

Poission case		Biases (×10)			Variances (×100)			MSEs (×100)		
α	n	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}_{YW}^{(1)}$	$\widetilde{\alpha}_{YW}^{(2)}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}_{YW}^{(1)}$	$\widetilde{\alpha}_{YW}^{(2)}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}_{YW}^{(1)}$	$\widetilde{\alpha}_{YW}^{(2)}$
0.2	100	-0.418	-0.155	-0.151	1.672	1.929	1.909	1.847	1.953	1.932
	200	-0.232	-0.083	-0.082	0.975	1.069	1.059	1.029	1.075	1.066
	300	-0.203	-0.089	-0.099	0.644	0.701	0.685	0.686	0.708	0.695
0.5	100	-0.585	-0.276	-0.292	1.591	1.693	1.675	1.933	1.769	1.760
	200	-0.385	-0.219	-0.229	0.929	0.963	0.956	1.077	1.011	1.008
	300	-0.240	-0.109	-0.132	0.681	0.705	0.695	0.739	0.717	0.712
0.8	100	-0.475	-0.202	-0.208	0.593	0.609	0.611	0.818	0.650	0.655
	200	-0.237	-0.097	-0.100	0.313	0.317	0.317	0.369	0.326	0.327
	300	-0.176	-0.078	-0.084	0.212	0.213	0.214	0.243	0.219	0.221
		$\widetilde{\alpha}_{YW}^{\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\dagger\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\dagger\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\dagger\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$
0.2	100	0.024	0.037	-0.253	6.356	2.017	1.814	6.357	2.018	1.878
	200	0.013	0.005	-0.146	1.269	1.142	1.016	1.269	1.142	1.037
	300	-0.027	-0.041	-0.145	0.719	0.727	0.663	0.720	0.729	0.684
0.5	100	0.020	0.013	-0.303	1.973	1.983	1.726	1.974	1.983	1.818
	200	-0.060	-0.061	-0.240	1.076	1.072	0.968	1.079	1.075	1.026
	300	-0.012	-0.012	-0.141	0.763	0.760	0.700	0.763	0.761	0.720
0.8	100	-0.014	-0.014	-0.071	0.645	0.645	0.640	0.645	0.645	0.646
	200	0.005	0.005	-0.030	0.332	0.332	0.325	0.332	0.332	0.326
	300	-0.013	-0.013	-0.037	0.221	0.221	0.217	0.221	0.221	0.219
NB case										
α	n	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}_{YW}^{(1)}$	$\widetilde{\alpha}_{YW}^{(2)}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}_{YW}^{(1)}$	$\widetilde{\alpha}_{YW}^{(2)}$	$\widehat{\alpha}_{YW}$	$\widetilde{\alpha}_{YW}^{(1)}$	$\widetilde{\alpha}_{YW}^{(2)}$
0.2	100	-0.411	-0.170	-0.166	1.572	1.800	1.786	1.741	1.828	1.814
	200	-0.248	-0.113	-0.112	0.891	0.965	0.958	0.953	0.978	0.971
	300	-0.126	-0.019	-0.029	0.689	0.747	0.731	0.705	0.748	0.731
0.5	100	-0.582	-0.277	-0.290	1.657	1.765	1.746	1.996	1.842	1.830
	200	-0.367	-0.203	-0.212	0.953	0.989	0.981	1.088	1.030	1.026
	300	-0.245	-0.116	-0.138	0.674	0.698	0.688	0.734	0.712	0.707
0.8	100	-0.469	-0.194	-0.200	0.611	0.627	0.629	0.831	0.665	0.669
	200	-0.256	-0.115	-0.118	0.320	0.324	0.324	0.385	0.337	0.338
	300	-0.174	-0.075	-0.081	0.206	0.207	0.208	0.236	0.213	0.214
		$\widetilde{\alpha}_{YW}^{\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\dagger\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\dagger\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\dagger\text{Ana}}$	$\widetilde{\alpha}_{YW}^{\prime\text{Ana}}$
0.2	100	0.043	-0.004	-0.246	4.755	1.827	1.705	4.756	1.827	1.766
	200	-0.021	-0.039	-0.162	1.036	1.013	0.929	1.036	1.014	0.955
	300	0.036	0.022	-0.067	0.776	0.780	0.708	0.777	0.781	0.713
0.5	100	0.014	0.005	-0.300	2.073	2.089	1.797	2.073	2.089	1.888
	200	-0.047	-0.050	-0.222	1.123	1.111	0.993	1.125	1.114	1.042
	300	-0.023	-0.024	-0.147	0.760	0.756	0.693	0.760	0.756	0.714
0.8	100	0.002	0.002	-0.064	0.672	0.672	0.661	0.672	0.672	0.665
	200	-0.012	-0.012	-0.049	0.338	0.338	0.333	0.339	0.339	0.335
	300	-0.009	-0.009	-0.034	0.215	0.215	0.211	0.215	0.215	0.212

Table 6.5: Biases, variances, and MSEs of the estimators $\widehat{\alpha}_{CLS}$ (without bias-correction), $\widetilde{\alpha}^{(q)}_{CLS}$, $q = 1, 2$, $\widetilde{\alpha}^{\text{Ana}}_{CLS}$, $\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$, and $\widetilde{\alpha}'^{\text{Ana}}_{CLS}$ in the stationary ADCINAR(1) process ($\vartheta = 0.9$) under the Poisson and NB innovations (note that $\widetilde{\alpha}'^{\text{Ana}}_{YW}$ is used in a over-fitted case).

Poission case		Biases (×10)			Variances (×100)			MSEs (×100)		
α	n	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$
0.2	100	-0.399	-0.152	-0.148	1.720	1.943	1.923	1.880	1.966	1.945
	200	-0.223	-0.083	-0.082	0.984	1.068	1.058	1.034	1.075	1.065
	300	-0.197	-0.089	-0.099	0.649	0.701	0.685	0.688	0.709	0.695
0.5	100	-0.532	-0.268	-0.283	1.630	1.699	1.681	1.913	1.770	1.761
	200	-0.361	-0.218	-0.228	0.943	0.967	0.961	1.073	1.015	1.013
	300	-0.224	-0.109	-0.132	0.686	0.706	0.696	0.736	0.718	0.713
0.8	100	-0.388	-0.191	-0.196	0.610	0.614	0.616	0.760	0.651	0.655
	200	-0.195	-0.094	-0.097	0.314	0.315	0.315	0.352	0.324	0.325
	300	-0.147	-0.075	-0.081	0.212	0.211	0.212	0.233	0.217	0.219
		$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$
0.2	100	0.027	0.034	-0.249	5.271	2.010	1.831	5.272	2.011	1.893
	200	0.009	0.004	-0.145	1.277	1.138	1.016	1.277	1.138	1.037
	300	-0.028	-0.042	-0.145	0.719	0.727	0.663	0.720	0.729	0.684
0.5	100	0.012	0.007	-0.292	1.958	1.967	1.735	1.958	1.967	1.820
	200	-0.063	-0.064	-0.239	1.076	1.072	0.973	1.080	1.076	1.030
	300	-0.013	-0.014	-0.141	0.762	0.760	0.701	0.762	0.760	0.721
0.8	100	-0.010	-0.010	-0.057	0.643	0.643	0.646	0.644	0.643	0.649
	200	0.005	0.005	-0.025	0.328	0.328	0.323	0.328	0.328	0.324
	300	-0.012	-0.012	-0.033	0.219	0.219	0.216	0.219	0.219	0.217
NB case										
α	n	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$	$\widehat{\alpha}_{CLS}$	$\widetilde{\alpha}^{(1)}_{CLS}$	$\widetilde{\alpha}^{(2)}_{CLS}$
0.2	100	-0.394	-0.168	-0.165	1.602	1.797	1.783	1.757	1.825	1.811
	200	-0.239	-0.113	-0.112	0.900	0.965	0.958	0.957	0.978	0.971
	300	-0.120	-0.019	-0.029	0.693	0.747	0.730	0.707	0.747	0.731
0.5	100	-0.531	-0.270	-0.284	1.705	1.779	1.760	1.987	1.852	1.841
	200	-0.343	-0.202	-0.211	0.962	0.988	0.981	1.080	1.029	1.025
	300	-0.229	-0.116	-0.138	0.680	0.699	0.689	0.732	0.713	0.708
0.8	100	-0.383	-0.184	-0.190	0.630	0.634	0.636	0.777	0.668	0.672
	100	-0.215	-0.113	-0.116	0.323	0.324	0.324	0.369	0.336	0.338
	300	-0.145	-0.072	-0.078	0.207	0.207	0.208	0.228	0.212	0.214
		$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\text{Ana}}_{CLS}$	$\widetilde{\alpha}^{\dagger\text{Ana}}_{CLS}$	$\widetilde{\alpha}'^{\text{Ana}}_{CLS}$
0.2	100	0.007	-0.007	-0.244	2.433	1.815	1.704	2.433	1.815	1.764
	200	-0.022	-0.040	-0.162	1.034	1.011	0.929	1.035	1.012	0.955
	300	0.035	0.021	-0.067	0.774	0.779	0.708	0.776	0.780	0.712
0.5	100	0.006	-0.003	-0.292	2.063	2.079	1.815	2.063	2.079	1.900
	200	-0.050	-0.052	-0.220	1.118	1.107	0.993	1.121	1.110	1.042
	300	-0.024	-0.025	-0.146	0.759	0.756	0.694	0.760	0.756	0.716
0.8	100	0.003	0.003	-0.051	0.671	0.671	0.668	0.671	0.671	0.671
	200	-0.013	-0.013	-0.046	0.337	0.337	0.333	0.337	0.337	0.335
	300	-0.008	-0.008	-0.031	0.214	0.214	0.211	0.215	0.215	0.212

Let $\delta = (\delta_1, \delta_2, \delta_3)^T = (\widehat{\alpha} - \alpha, \overline{Y} - \mu_Y, \widehat{\sigma}_Y^2 - \sigma_Y^2)^T = O_p(1/\sqrt{n})$. It is straightforward to see that

$$\begin{aligned} g_t(\widehat{\alpha}, \overline{Y}) &= g_t(\alpha, \mu_Y) - 2(Z_t - \alpha Z_{t-1})\{Z_{t-1}\delta_1 + (1-\alpha)\delta_2\} + R_1(\delta_1, \delta_2, Z_t, Z_{t-1}), \\ f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) &= f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) + \{-2\alpha Z_{t-1}^2 + (1 - 4\alpha\mu_Y)Z_{t-1}\}\delta_1 - \alpha(1 - 2\alpha\mu_Y)\delta_2 + \delta_3 \\ &\quad + R_2(\delta_1, \delta_2, Z_{t-1}), \\ f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2) &= f_{2,t-1}(\mu_Y, \sigma_Y^2) + (1 - 2\mu_Y)\delta_2 - \delta_3 - \delta_2^2, \end{aligned}$$

where the random variables $R_1(\delta_1, \delta_2, Z_t, Z_{t-1})$ and $R_2(\delta_1, \delta_2, Z_{t-1})$ are monomials in (δ_1, δ_2) of degrees 2, 3, and 4, involving $Z_{t-1}^i Z_t^j$ for $0 \le i + j \le 2 (j = 0, 1)$, hence,

$$\begin{aligned} & g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) - \vartheta\widehat{\alpha} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2) - \{g_t(\alpha, \mu_Y) - f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) - \vartheta\alpha f_{2,t-1}(\mu_Y, \sigma_Y^2)\} \\ &= [-2(Z_t - \alpha Z_{t-1})Z_{t-1} + (2\alpha - \vartheta)Z_{t-1}^2 - \{(1 - 4\alpha\mu_Y) - \vartheta(1 - 2\mu_Y)\}Z_{t-1} + \vartheta\sigma_Y^2]\delta_1 \\ &\quad - \{2(1-\alpha)(Z_t - \alpha Z_{t-1}) - \alpha(1 - 2\alpha\mu_Y) + \vartheta\alpha(1 - 2\mu_Y)\}\delta_2 - (1 - \vartheta\alpha)\delta_3 + R_3(\delta, Z_t, Z_{t-1}), \end{aligned}$$

where the random variable $R_3(\delta, Z_t, Z_{t-1})$ is a monomial in δ of degrees 2, 3, and 4, involving $Z_{t-1}^i Z_t^j$ for $0 \le i + j \le 2 (j = 0, 1)$.

Now, we define, for $t = 2, \ldots, n$,

$$\begin{aligned} W_{1,t} &= [-2(Z_t - \alpha Z_{t-1})Z_{t-1} + (2\alpha - \vartheta)Z_{t-1}^2 - \{(1 - 4\alpha\mu_Y) - \vartheta(1 - 2\mu_Y)\}Z_{t-1} + \vartheta\sigma_Y^2] f_{2,t-1}(\mu_Y, \sigma_Y^2), \\ W_{2,t} &= [\{g_t(\alpha, \mu_Y) - f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) - \vartheta\alpha f_{2,t-1}(\mu_Y, \sigma_Y^2)\}(1 - 2\mu_Y) \\ &\quad - \{2(1-\alpha)(Z_t - \alpha Z_{t-1}) - \alpha(1 - 2\alpha\mu_Y) + \vartheta\alpha(1 - 2\mu_Y)\} f_{2,t-1}(\mu_Y, \sigma_Y^2)], \\ W_{3,t} &= -[\{g_t(\alpha, \mu_Y) - f_{1,t-1}(\alpha, \mu_Y, \sigma_Y^2) - \vartheta\alpha f_{2,t-1}(\mu_Y, \sigma_Y^2)\} + (1 - \vartheta\alpha) f_{2,t-1}(\mu_Y, \sigma_Y^2)]. \end{aligned}$$

It is not difficult to see that $E(W_{i,t}) = E(W_{i,2}), i = 1, 2, 3$, by strictly stationarity of $\{Y_t\}$, where

$$E(W_{1,2}) = A\sigma_Y^2 \text{ and } E(W_{2,2}) = E(W_{3,2}) = 0. \tag{6.23}$$

After some algebra, we have

$$\{g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) - \vartheta\widehat{\alpha} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)\} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2) = M_{\vartheta,t} + \sum_{i=1}^{3} \delta_i W_{i,t} + R(\delta, Z_t, Z_{t-1}),$$

where the random variable $R(\delta, Z_t, Z_{t-1})$ is a monomial in δ of degrees 2, 3, 4, 5, 6, involving $Z_{t-1}^i Z_t^j$ for $0 \le i + j \le 4 (j = 0, 1)$. This, together with Lemma 5.6, yields

$$\frac{1}{\sqrt{n}} \sum_{t=2}^{n} \{g_t(\widehat{\alpha}, \overline{Y}) - f_{1,t-1}(\widehat{\alpha}, \overline{Y}, \widehat{\sigma}_Y^2) - \vartheta\widehat{\alpha} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)\} f_{2,t-1}(\overline{Y}, \widehat{\sigma}_Y^2)$$

$$= \frac{1}{\sqrt{n}} \sum_{t=2}^{n} M_{\vartheta,t} + \sqrt{n} \sum_{i=1}^{3} \delta_i E(W_{i,2}) + o_p(1). \tag{6.24}$$

Substituting (6.24) for the numerator of (6.22), together with (6.6) and (6.23), we have the stochastic expansion of $\widehat{\vartheta}_{2CLS}$ (we also used $\widehat{\alpha} = \alpha + O_p(n^{-1/2}) = \alpha + o_p(1)$, which is implied by (6.6) and (6.7)).

Chapter 7

Data analyses

We analyze two real datasets; IP count data and Download count data, available in Weiß (2018), and demonstrate the usefulness of two models of nonnegative integer-valued autoregressive process of the first-order (INAR(1)) and alternative dependent counting nonnegative INAR process of the first-order (ADCINAR(1)). Especially, Section 7.1 focuses on the equidispersion tests developed in Section 4.3. Also, Section 7.2 illustrates the CLS estimation (Section 6.2) for the stationary ADCINAR(1) process without the specific distributional form of the innovation.

7.1 IP count data

We first analyze the IP count data of length $n = 241$, available in Weiß (2018), whose count represents how many different IPs have registered in a 2-minute period. In view of Figure 7.1, the IP count data exhibit the first-order autoregressive (AR(1))-like autocorrelation structure (the autocorrelation at lag 1 is 0.219). According to Chapter 4, the equidispersion tests are performed by fitting the INAR(1) process in three different ways, i.e., the Whittle, Yule–Walker (YW), and conditional least squares (CLS) estimators for the parameter α and the innovation mean and variance. It is revealed that the IP count data is equidispersed since the z-values are given by $z_W = 0.396$, $z_{YW} = 0.375$, and $z_{CLS} = 0.336$. Our estimates for α, μ_ε, and σ_ε^2, together with the standard errors (SEs), are given in Table 7.1(i). Note that, assuming the Poisson marginals, Weiß (2018) obtained the maximum likelihood (ML) estimates $\widehat{\alpha}_{ML} = 0.243$ (SE = 0.062) and $\widehat{\mu}_{\varepsilon;ML} = 0.997$ (SE = 0.099) for α and μ_ε, respectively

On the other hand, according to Weiß (2018; page 175), the IP count data may contain an outlier Y_{224}; we now set $Y_{224} = 1$. The z-values for the outlier-corrected IP count data are given by $z_W = -0.707$, $z_{YW} = -0.762$, and $z_{CLS} = -0.835$, which also reveals the equidispersion. The SEs of all estimates in Table 7.1(ii) for the outlier-corrected data are smaller than those in Table 7.1(i) for the original data.

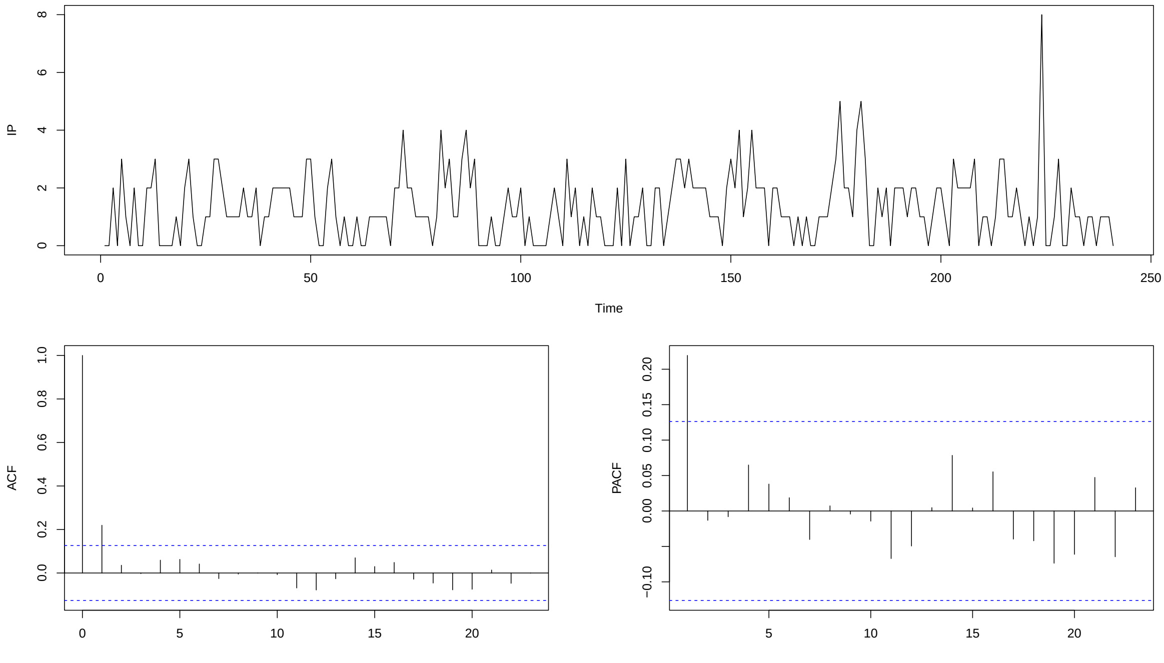

Figure 7.1: The time series plot, ACF, and PACF of the IP count data.

Table 7.1: Estimates (SEs) for α, μ_ε, and σ_ε^2.

(i) The IP count data.

	Whittle		YW		CLS	
α	0.219	(0.069)	0.219	(0.069)	0.221	(0.069)
μ_ε	1.024	(0.108)	1.027	(0.108)	1.031	(0.108)
σ_ε^2	1.095	(0.207)	1.094	(0.207)	1.091	(0.207)

(ii) The outlier-corrected data.

	Whittle		YW		CLS	
α	0.293	(0.066)	0.292	(0.066)	0.294	(0.066)
μ_ε	0.906	(0.098)	0.910	(0.098)	0.914	(0.098)
σ_ε^2	0.833	(0.105)	0.831	(0.105)	0.828	(0.104)

7.2 Download count data

We next analyze the Download count data of length $n = 267$, available in Weiß (2018), whose count represents the daily number of downloads of a TeX editor for the period from June 2006 to February 2007. In view of Figure 7.2, the Download count data exhibit an AR(1)-like autocorrelation structure (the autocorrelation at lag 1 is 0.245). Weiß (2018) fitted the data to the stationary INAR(1) model under the negative binomial (NB) innovation, random coefficient INAR(1) model under the NB innovation, and so on. Here, using the stationary ADCINAR(1) model without the specific distributional assumption on the innovation, we obtain $\widehat{\alpha}_{CLS} = 0.247$ (SE = 0.065) and $\widehat{\vartheta}^{\dagger}_{2CLS} = 0.472$ (SE = 0.084), respectively.

It is important to test whether the data are generated by the stationary INAR(1) process or the stationary ADCINAR(1) process. By Proposition 6.3(ii), one may compute the Wald-type test statistic

$$Z = \frac{\sqrt{n}(\widehat{\vartheta}_{2CLS} - \widehat{\alpha}_{CLS})}{\sqrt{\widehat{\psi}_{11} - 2\widehat{\psi}_{12} + \widehat{\psi}_{22}}}$$

for testing H: $\vartheta = \alpha$ against A: $\vartheta > \alpha$. The ADCINAR(1) model might be suitable (the z-value is 2.097). However, since the null hypothesis H: $\vartheta = \alpha$ is on the boundary of the parameter space of (α, ϑ), the detailed study of this testing problem is left for the future.

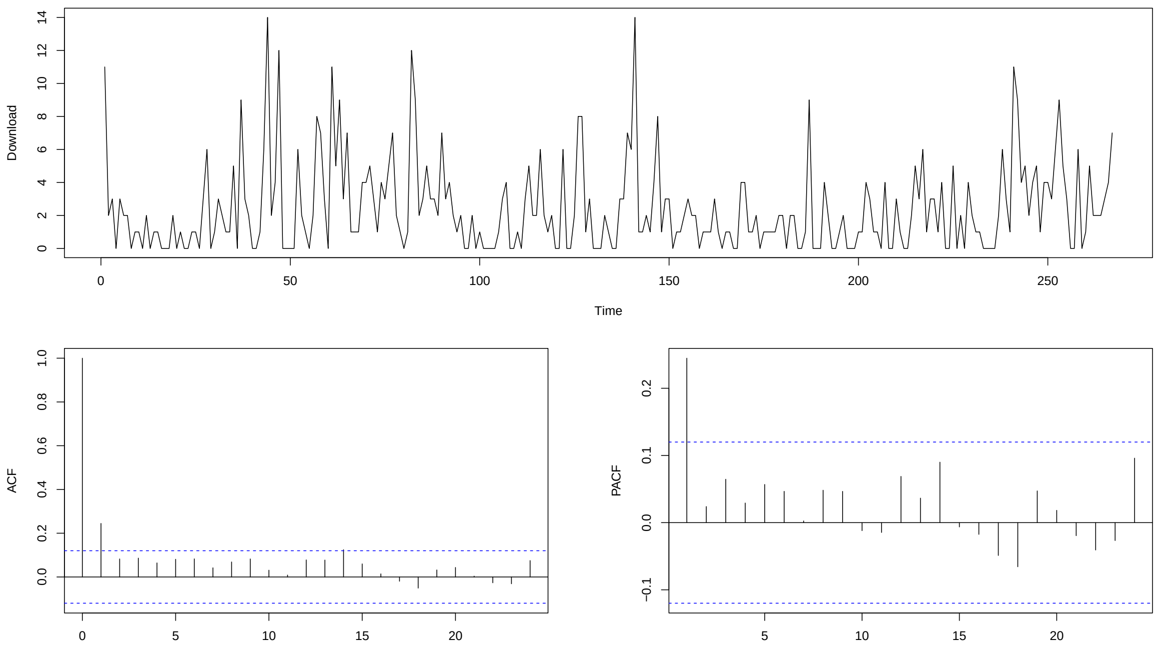

Figure 7.2: The time series plot, ACF, and PACF of the Download count data.

Chapter 8

Conclusions and future issues

Summary

We have mainly clarified the following points.

1 For the stationary nonnegative integer-valued autoregressive process of the first-order (INAR(1)) and alternative dependent counting nonnegative INAR process of the first-order (ADCINAR(1)), the third and fourth autocumulant functions have been derived explicitly, together with the structure about arbitrary higher autocumulant functions.

2 A nonparametric (lag window-type) bias-correction and an analytical bias-correction have been developed for the widely used Yule–Walker (YW) and conditional least squares (CLS) estimators in the stationary INAR(1) and ADCINAR(1) processes. The lag window-type bias-correction, which is available without computing the closed-form expression for asymptotic expansions of the biases, is practically useful since the analytical bias formula is complicated for the stationary ADCINAR(1) process.

3 The asymptotic theory about the frequency domain analysis of the stationary INAR(1) process has been presented. The Wald-type test about the equidispersion has been constructed, on the basis of the estimators for the innovation mean and variance.

Future works

The following issues would be interesting and challenging.

- We are considering an efficient and practically feasible estimation for the new parameter ϑ in the stationary ADCINAR(1) process. Note that the null hypothesis H: $\vartheta = \alpha$, which we mentioned in Section 7.2, is on the boundary of the parameter space of (α, ϑ) in the stationary ADCINAR(1) process; the detailed study of this testing problem is left for the future.

- The choice of the truncation parameter L_n is crucial for the lag window-type bias-correction developed in Section 6.3, although the optimal order $L_n = O(n^{1/(2+2q)})$ for Parzen's lag window $k(x) = 1 - |x|^q$, where $q = 1, 2$, was justified there. We leave a suitable data-driven choice of L_n for future.

- We know the recent work about a novel interesting thinning operation (and model) based on a generalized Neyman type distribution theory (see Amiri et al. (2022)). Since such a model is strictly stationary and ergodic, whose autocorrelation structures are the same as the ADCINAR(1) process, the bias-corrected YW and CLS estimators (Section 6.3) are expected to be applicable even for their model.

- Bootstrap and jackknife procedures are important techniques in Statistics. Jentsch and Weiß (2019) proved the validity of bootstrapping for the stationary INAR(p) process, so that the extension of their result to the stationary ADCINAR(1) process is left for future.

Acknowledgments

As a Ph.D student, I have spent five significant years studying at Hokkaido University. It has been a crucial period in my life. During this time, I have encountered many kind and supportive individuals and organizations that offered me substantial help, especially with my studies. I am honored and proud to be a student at the graduate school.

Firstly, I am grateful for the professors I encountered during my academic program, including Professor Yoshihide Kakizawa, Hikaru Hasegawa, Shintaro Sono, Akio Suzukawa, Shingo Takagi, and Yukiko Abe. Professor Kakizawa, as my advisor for my thesis, not only provided meticulous guidance in writing it but also detailedly introduced the basics of time series analysis and asymptotic theory in the graduate school course. Professor Hasegawa's course equipped me with foundational skills in using the R software and understanding Bayesian statistics. In Professor Sono's course, I delved into various statistical concepts. Professor Takagi's course gave me a comprehensive overview of modern econometrics. Professor Suzukawa's course taught me something that every statistician must fully understand. Professor Abe's course provided me with a valuable opportunity to explore a field beyond statistics and apply my statistical knowledge. It was an enriching experience that broadened my perspective.

I would like to express my deepest gratitude once again to my supervisor, Professor Kakizawa. Particularly in the later stages of my Ph.D, he dedicated a significant amount of his personal time to help me revise my thesis. He consistently displayed patience in addressing my frequent mistakes and habits, persistently guiding me toward improvement.

Secondly, I am grateful to the graduate school for providing a conducive learning environment. Additionally, to facilitate academic exchanges with the outside world, the graduate school offered travel expenses for each academic conferences I attended; (i) EcoSta (2023, August) and (ii) seminar on nonparametric and Bayesian analyses (2023, March).

Finally, I acknowledge the financial support from the China Scholarship Council, which ensured I could concentrate on my studies.

Bibliography

[1] Al-Osh, M. and Alzaid, A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis* **8**, 261–275.

[2] Amiri, J., Farnoosh, R. and Behzadi, M.H. (2022). A novel dependent NTA thinning operator and generalized geometric INAR(1) process with contagious disease case studies. *Communications in Statistics—Simulation and Computation* (in press).

[3] Anderson, T.W. (1971). *The Statistical Analysis of Time Series.* New York: Wiley.

[4] Bourguignon, M. and Vasconcellos, K.L.P. (2015a). First order non-negative integer valued autoregressive processes with power series innovations. *Brazilian Journal of Probability and Statistics* **29**, 71–93.

[5] Bourguignon, M. and Vasconcellos, K.L.P. (2015b). Improved estimation for Poisson INAR(1) models. *Journal of Statistical Computation and Simulation* **85**, 2421–2441.

[6] Brockwell, P.J. and Davis, R.A. (1987). *Time Series: Theory and Methods.* New York: Springer Verlag.

[7] Du, J.G. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis* **12**, 129–142.

[8] Freeland, R.K. and McCabe, B.P.M. (2005). Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Statistics and Probability Letters* **73**, 147–153.

[9] Fujikoshi, Y. and Ochi, Y. (1984). Asymptotic properties of the maximum likelihood estimate in the first order autoregressive process. *Annals of the Institute of Statistical Mathematics* **36**, 119–128.

[10] Jentsch, C. and Weiß, C.H. (2019). Bootstrapping INAR models. *Bernoulli* **25**, 2359–2408.

[11] Johnson, N.L., Kemp, A.W. and Kotz, S. (2005). *Univariate Discrete Distributions, 3rd Edition.* New Jersey: Wiley.

[12] Jung, R.C., Ronning, G., Tremayne, A.R. (2005). Estimation in conditional first order autoregression with discrete support. *Statistical Papers* **46**, 195–224.

[13] Kakizawa, Y. (1996). Third order asymptotic properties of estimators in Gaussian ARMA processes with unknown mean. *Journal of Time Series Analysis* **17**, 367–377.

[14] Karlsen, H. and Tjøstheim, D. (1988). Consistent estimates for the NEAR(2) and NLAR(2) time series models. *Journal of the Royal Statistical Society: Series B (Methodological)* **50**, 313–320.

[15] Klimko, L.A. and Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics* **6**, 629–642.

[16] McCullagh, P. (2018). *Tensor Methods in Statistics*. New York: Dover Publications Inc.

[17] McLeish, D.L. (1974). Dependent central limit theorems and invariance principles. *The Annals of Probability* **2**, 620–628.

[18] Nastić, A.S., Ristić, M.M. and Miletić Ilić, A.V. (2017). A geometric time-series model with an alternative dependent Bernoulli counting series. *Communications in Statistics—Theory and Methods* **46**, 770–785.

[19] Noack, A. (1950). A class of random variables with discrete distributions. *The Annals of Mathematical Statistics* **21**, 90–120.

[20] Olver, F.W.J., Lozier, D.W., Boisvert, R.F. and Clark, C.W. (Eds.) (2010). *NIST Handbook of Mathematical Functions.* Cambridge: Cambridge University Press.

[21] Park, Y. and Oh, C.W. (1997). Some asymptotic properties in INAR(1) processes with Poisson marginals. *Statistical Papers* **38**, 287–302.

[22] Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. *The Annals of Mathematical Statistics* **28**, 329–348.

[23] Ristić, M.M., Nastić, A.S. and Miletić Ilić, A.V. (2013). A geometric time series model with dependent Bernoulli counting series. *Journal of Time Series Analysis* **34**, 466–476.

[24] Schweer, S. and Weiß, C.H. (2016). Testing for Poisson arrivals in INAR(1) processes. *Test* **25**, 503–524.

[25] da Silva, M.E. and Oliveira, V.L. (2004). Difference equations for the higher order moments and cumulants of the INAR(1) model. *Journal of Time Series Analysis* **25**, 317–333.

[26] Steutel, F.W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability* **7**, 893–899.

[27] van der Vaart, A.W. (1998). *Asymptotic Statistics.* Cambridge: Cambridge University Press.

[28] Weiß, C.H. (2008). Thinning operations for modeling time series of counts – a survey. *Advances in Statistical Analysis* **92**, 319–341.

[29] Weiß, C.H. and Schweer, S. (2016). Bias corrections for moment estimators in Poisson INAR(1) and INARCH(1) processes. *Statistics and Probability Letters* **112**, 124–130.

[30] Weiß, C.H. (2018). *An Introduction to Discrete-Valued Time Series.* New Jersey: John Wiley & Sons.

[31] Yang, Y.Q., Wang, D.H. and Zhao, Z.W. (2018). Empirical likelihood for first-order mixed integer-valued autoregressive model. *Applied Mathematics—A Journal of Chinese Universities* **33**, 313–322.

[32] Zeng, X. (2023). Whittle likelihood estimation in INAR(1) process. *Communications in Statistics—Theory and Methods* (in press).

[33] Zeng, X. and Kakizawa, Y. (2022). Bias-correction of some estimators in the INAR(1) process. *Statistics and Probability Letters* **187**, 109503.

[34] Zeng, X. and Kakizawa, Y. (2023). Higher autocumulant functions for ADCINAR(1) process and bias-correction of some estimators. *Japanese Journal of Statistics and Data Science* **6**, 179–211.

[35] Zeng, X. and Kakizawa, Y. (2024). Two-step conditional least squares estimation in ADCINAR(1) process, revisited. *Statistics and Probability Letters* **206**, 110003.

[36] Zhang, H. and Wang, D. (2015). Inference for random coefficient INAR(1) process based on frequency domain analysis. *Communications in Statistics—Simulation and Computation* **44**, 1078–1100.