Filomat 38:32 (2024), 11477–11499
https://doi.org/10.2298/FIL2432477A

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

On the Marshall-Olkin extended Gamma Lindley autoregressive process

Mariem Ammar[a,*], Imen Boutouria[a], Afif Masmoudi[a]

[a] *University of Sfax, Faculty of Sciences of Sfax, Laboratory of Probability and Statistics, Sfax, Tunisia*

Abstract. In this research paper, we generalize the Gamma Lindley distribution according to the Marshall-Olkin transformation, in order to enhance data modeling flexibility. The r^{th} moment of this distribution is derived. Hence, we introduce a first autoregressive process $(X_n)_{n\in\mathbb{N}}$ with minification structure using the proposed model. We exhibit some of its statistical properties, in particular we compute the first order autocovariance function, the conditional mean and the conditional cumulative distribution function of X_{n+1} given X_n. We prove that $\mathbb{P}(X_{n+1} = X_n) > 0$, then we derive explicitly the joint probability distribution of $(X_n,\ X_{n+1})$. Consequently, we estimate the unknown parameters of the proposed process using maximum likelihood, conditional least squares and method of moments estimation. Then, a simulation study is conducted to compare the performance of these approaches. An application to gold price data is illustrated to show the flexibility of this process. We develop an algorithm for predicting time series real data by employing the suggested autoregressive process. Results point out that the proposed autoregressive model offers a statistically superior fit for the real data. Furthermore, it is a good predictive model compared to the standard form for the first order one.

1. Introduction

Modeling data by generalized distributions is still very common nowadays. Several researchers proposed new generalizations for lifetime distributions used in various fields such as finance, medicine and engineering. For example, Laribi, Masmoudi, and Boutouria [12, 13] derived the Generalized Gamma Lindley distribution to enhance the capability of analyzing various forms of lifetime data. Ghitany *et al.* [7] used the Lindley distribution to illustrate an application to waiting time data in a bank. Rajitha and Akhilnath [23] generalized the Lindley distribution using the power exponentiated family of distributions to model the data of the covid-19 case fatality ratio. Ghitany *et al.* [6] proved that the power Lindley distribution is significant to address stress-strength reliability modeling. El-Morshedy and Eliwa [4] defined the odd flexible Weibull-H family of distributions and provided its applications in complete and upper record data. Additionally, Marshall and Olkin [17] introduced an important method of incorporating a new parameter to an existing distribution. The resulting model, known as Marshall-Olkin extended distribution, involves

2020 *Mathematics Subject Classification*. Primary 60E05; Secondary 62G05.
Keywords. Autocovariance function, Autoregressive minification process, Gamma Lindley distribution, Inference, Marshall-Olkin family, Time series modeling.
Received: 16 December 2023; Revised: 14 May 2024; Accepted: 27 September 2024
Communicated by Biljana Popović
* Corresponding author: Mariem Ammar
Email addresses: mariem.ammar.etud@fss.usf.tn (Mariem Ammar), imen.boutouria@gmail.com (Imen Boutouria), afif.masmoudi@fss.usf.tn (Afif Masmoudi)

the original distribution as a special case and provides greater flexibility to model different types of real data. Let $f_Y(x)$, $F_Y(x)$, $\overline{F}_Y(x) = 1 - F_Y(x)$ are respectively the probability density function (pdf), the cumulative distribution function (cdf) and the survival function (sf) of a random variable Y following the original model. According to the technique of Marshall-Olkin, the new survival function of a random variable X is expressed as

$$\overline{F}_X(x;\, p) = \frac{p\overline{F}_Y(x)}{1 - \overline{p}\overline{F}_Y(x)}, \qquad x \in \mathbb{R}, \tag{1}$$

where $p > 0$ and $\overline{p} = 1 - p$. The corresponding probability density function is given by

$$f_X(x;\, p) = \frac{p f_Y(x)}{(1 - \overline{p}\overline{F}_Y(x))^2}, \qquad x \in \mathbb{R}. \tag{2}$$

The use of Marshall-Olkin transformation has been a great interest among many researchers. For instance, Eghwerido *et al.* [3] proposed the Marshall-Olkin Gompertz distribution and explored its practical applications. Ghitani *et al.* [5] derived the Marshall-Olkin extended Weibull distribution and examined its application for censored data. Krishna *et al.* [10] studied the applications of the Marshall-Olkin fréchet distribution in acceptance sampling, reliability analysis and time series modeling. Zeghdoudi and Nedjar, recently introduced a new distribution, called Gamma Lindley (GaL) distribution [20, 24], as a mixture of Gamma(2, θ) distribution and Lindley(θ) distribution (see Lindley [14]) with mixing probabilities $\frac{\alpha - 1}{\alpha}$ and $\frac{1}{\alpha}$, respectively. They proved that this distribution is flexible for modeling real data in various fields and it is more suitable than Exponential, Weibull, Gamma, two parameter Lindley distributions for some applications. The usefulness of the Gamma Lindley distribution allows for the derivation of numerous generalizations using a variety of methodology (see [2, 19, 21]).
Motivated by the idea of Marshall-Olkin and to increase the flexibility of the Gamma Lindley distribution, we generalize it according to this technique.
Furthermore, modeling time series data is indeed crucial in different fields, especially finance, economics, environmental science. As a result of numerous practical issues of modeling real data instances, several time series models with non-Gaussian marginal are introduced. Accordingly, some attractive models with Exponential [16], Lindley [1] and Gamma Lindley are discussed [18]. Autoregressive modeling is one of the interesting techniques applied in time series analysis. This technique predicts future values based on past observations, when there is some correlation between values.
Autoregressive models were constructed behind the idea that the present value of a series $(X_n)_{n\in\mathbb{N}^*}$ can be explained through a function of p past values $X_{n-1},\ X_{n-2},\ X_{n-3}, ...,\ X_{n-p}$.
The standard form of an autoregressive model of order p, denoted by AR(p), is given by

$$X_n = a_0 + \sum_{i=1}^{p} a_i X_{n-i} + \epsilon_n$$

with ϵ_n is a white noise and $(a_i)_{0\le i\le p}$ are $p + 1$ fixed parameters. The minification process [15] can be used to model catastrophic events such as earthquakes, storms and stock market crashes. A first order autoregressive minification process having the following general structure,

$$X_n = \begin{cases} \epsilon_n & \text{with} \quad \text{probability} \quad p \\ \min(X_{n-1}, \epsilon_n) & \text{with} \quad \text{probability} \quad 1 - p \end{cases} \tag{3}$$

with $0 < p < 1$ and $\{\epsilon_n\}$ is a sequence of independent and identically distributed random variables independent of $\{X_i,\ i < n\}$. This process was introduced with a marginal distribution that belonging to the family of Marshall-Olkin distributions by many researchers. Within this context, Krishnan and George [11] defined the minification process with Marshall-Olkin Weibull Truncated Negative Binomial Distribution as marginal. Jose *et al.* [9] constructed a minification process with Marshall-Olkin beta distribution. Further, Jaykumar and Mathew [8] developed an application in modeling time series exchange rate data using the Marshall-Olkin semi-Burr autoregressive process.

In this paper, we introduce a minification process with the Marshall-Olkin Gamma Lindley distribution called the Marshall–Olkin Gamma Lindley autoregressive process. The latter is motivated by the following: the mathematical properties of this process are relatively simple in form and may be used as quick approximations in many cases. This new process has several advantages, including a number of parameters that can be used to model a variety of phenomena in diverse fields such as actuarial science and economics. Moreover, it can provide adequate fits to many data sets and can be used effectively in analyzing many real lifetime data sets.

The main goal of this paper is to introduce a new generalization of the Gamma Lindley distribution [24] based on the Marshall-Olkin transformation in order to develop an autoregressive process having the minification structure (3) and to show that this model can well predict time series data in diverse disciplines.

The paper is outlined as follows. In Section 2, we generalize the Gamma Lindley distribution using the Marshall-Olkin technique in order to obtain a more flexible model. In Section 3, we develop the autoregressive process of first order with minification structure using the proposed generalization and we explore some of its statistical properties. In Section 4, we conduct a statistical inference for the unknown parameters of this process using different methods of estimation, namely maximum likelihood, conditional least squares and method of moments. In Section 5, we illustrate an application of the proposed process to real data set. Finally, Section 6 is dedicated to concluding remarks and perspectives for future research.

2. Marshall-Olkin Extended Gamma Lindley distribution

Due to the ability of the Gamma Lindley distribution in modeling time series data, as well as the remarkable flexibility of the Marshal-Olkin family in autoregressive models as marginal. We extend the GaL distribution relying on the Marshall-Olkin method. The probability density function and the survival function of a random variable Y follows a Gal distribution are given as follows

$$f_Y(x;\ \alpha,\ \theta) = \frac{\theta^2((\alpha+\alpha\theta-\theta)x+1)e^{-\theta x}}{\alpha(1+\theta)}, \quad x>0,\ \alpha>\frac{\theta}{1+\theta},\ \theta>0,$$

$$\overline{F}_Y(x;\ \alpha,\ \theta) = \frac{((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)}, \quad x>0,\ \alpha>\frac{\theta}{1+\theta},\ \theta>0. \tag{4}$$

By substituting Equation (4) into Equation (1), we obtain the new survival function of the random variable X as follows

$$\overline{F}_X(x;\ \alpha,\ p,\ \theta) = \frac{p((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)-\overline{p}((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}, \quad x>0. \tag{5}$$

The new generalisation entitled Marshall-Olkin Extended Gamma Lindley (MOEGaL) with parameters α, p and θ, where $\theta>0$ is the scale parameter, $\alpha>\dfrac{\theta}{1+\theta}$ and $p>0$ are the shape parameters.

The corresponding probability density function is expressed as below

$$f_X(x;\ \alpha,\ p,\ \theta) = \frac{\alpha(1+\theta)p\theta^2((\theta\alpha+\alpha-\theta)x+1)e^{-\theta x}}{(\alpha(1+\theta)-\overline{p}((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x})^2}, \quad x>0.$$

This distribution has three parameters, providing significant flexibility. It is well suited for modeling real data. Now, we prove that this extended model is geometrically extremely stable.

Theorem 2.1. *Let N be a geometric random variable with parameter $p \in (0,1)$, and let $\{Y_i, i \geq 1\}$ be a sequence of independently and identically random variables distributed as GaL(α, θ) distribution with $\alpha > \frac{\theta}{1+\theta}$, $\theta > 0$. Assuming that N is independent of Y_i, $i \geq 1$, then*

(i) $U_N = \min(Y_1, Y_2, ..., Y_N)$ *has a Marshall-Olkin Extended Gamma Lindley distribution with parameters α, p, θ.*

(ii) $V_N = \max(Y_1, Y_2, ..., Y_N)$ *has a Marshall-Olkin Extended Gamma Lindley distribution with parameters α, $\frac{1}{p}$, θ.*

Proof. *Using the law of total probability, the survival function of U_N is given by*
(i)

$$\begin{aligned}\overline{F}_{U_N}(x;\,\alpha,\,p,\,\theta) &= \mathbb{P}(U_N > x),\\ &= \sum_{n=1}^{\infty} \mathbb{P}(U_n > x)\mathbb{P}(N = n).\end{aligned}$$

Since, N follows the geometric distribution with parameter p $\left(\mathbb{P}(N = n) = p(1-p)^{n-1},\ n = 1,\ 2, ...\right)$, *therefore*

$$\begin{aligned}\overline{F}_{U_N}(x;\,\alpha,\,p,\,\theta) &= \sum_{n=1}^{\infty} \overline{F}^{n}_{Y}(x;\,\alpha,\,\theta)(1-p)^{n-1}p,\\ &= \frac{p\overline{F}_Y(x;\,\alpha,\,\theta)}{1-(1-p)\overline{F}_Y(x;\,\alpha,\,\theta)},\\ &= \frac{p((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)-\overline{p}((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}},\end{aligned}$$

which is the survival function of Marshall-Olkin Extended Gamma Lindley distribution with parameters α, p, θ.

(ii) *The cumulative function of V_N is expressed as*

$$\begin{aligned}F_{V_N}(x;\,\alpha,\,p,\,\theta) &= \mathbb{P}(V_N \leq x),\\ &= \sum_{n=1}^{\infty} \mathbb{P}(V_n \leq x)\mathbb{P}(N = n),\\ &= \sum_{n=1}^{\infty} F^{n}_{Y}(x;\,\alpha,\,\theta)(1-p)^{n-1}p,\\ &= \frac{pF_Y(x;\,\alpha,\,\theta)}{1-(1-p)F_Y(x;\,\alpha,\,\theta)}.\end{aligned}$$

From this it is follows that the survival function of the random variable V_N is

$$\begin{aligned}
\overline{F}_{V_N}(x;\ \alpha,\ p,\ \theta) &= \frac{\overline{F}_Y(x;\ \alpha,\ \theta)}{p+(1-p)\overline{F}_Y(x;\ \alpha,\ \theta)},\\
&= \frac{((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{p\alpha(1+\theta)+(1-p)((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}},\\
&= \frac{\frac{1}{p}((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)-(1-\frac{1}{p})((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}},
\end{aligned}$$

which implies that V_N follows the Marshall-Olkin Extended Gamma Lindley distribution with parameters α, $\frac{1}{p}$, θ. □

The following proposition establishes the r^{th} moment of $MOEGaL(\alpha,\ p,\ \theta)$ distribution.

Proposition 2.2. *Let $(Y_n)_{n\geq 1}$ be a sequence of independently and identically random variables distributed as $GaL(\alpha,\ \theta)$ distribution with $\alpha > \frac{\theta}{1+\theta}$, $\theta > 0$ and let $Z_{j+1} = \min(Y_1,\ Y_2, ..., Y_{j+1})$. Assume that X be a random variable following $MOEGaL(\alpha,\ p,\ \theta)$ distribution with $0 < p \leq 1$, then its r^{th} moment is expressed as*

$$\mu'_r = \mathbb{E}(X^r) = \sum_{j=0}^{\infty} p\overline{p}^j\mathbb{E}(Z^r_{j+1}), \qquad r = 1, 2, ...$$

Proof. Before proceeding with the computation, the initial focus is on demonstrating the existence of the r^{th} moment of $MOEGaL(\alpha,\ p,\ \theta)$ distribution.

$$\begin{aligned}
\mathbb{E}(X^r) &= \int_0^\infty \mathbb{P}(X^r > x)\,dx, \qquad r = 1,\ 2, ...\\
&= \int_0^\infty \frac{p\overline{F}_Y(x^{1/r};\ \alpha,\ \theta)}{1-\overline{p}\overline{F}_Y(x^{1/r};\ \alpha,\ \theta)}\,dx,\\
&= \int_0^\infty \frac{p\overline{F}_Y(x^{1/r};\ \alpha,\ \theta)}{p+\overline{p}F_Y(x^{1/r};\ \alpha,\ \theta)}\,dx,\\
&\leq \int_0^\infty \overline{F}_Y(x^{1/r};\ \alpha,\ \theta)\,dx,\\
&\leq \int_0^\infty \mathbb{P}(Y > x^{1/r})\,dx,\\
&\leq \mathbb{E}(Y^r).
\end{aligned}$$

In addition, the Gamma Lindley distribution is a mixture of $Gamma(2,\ \theta)$ and $Lindley(\theta)$, with mixing probabilities $\frac{\alpha-1}{\alpha}$ and $\frac{1}{\alpha}$, respectively.

$$\begin{aligned}
\mathbb{E}(Y^r) &= \int_0^\infty x^r f_Y(x;\ \alpha,\ \theta)\,dx,\\
&= \frac{\alpha-1}{\alpha}\int_0^\infty \theta^2 x^{r+1}e^{-\theta x}\,dx + \frac{1}{\alpha}\int_0^\infty \frac{\theta^2}{\theta+1}x^r(1+x)e^{-\theta x}dx,\\
&= \frac{(\alpha-1)(r+1)!}{\alpha\theta^r} + \frac{r!(\theta+r+2)}{\alpha\theta^r(\theta+1)} \quad < \infty. \qquad (6)
\end{aligned}$$

By proving the integrability, we affirm the convergence. Then, we compute the r^{th} moment as follows

$$\begin{aligned} \mathbb{E}(X^r) &= \int_0^\infty \mathbb{P}(X^r > x)\, dx, \\ &= \int_0^\infty \frac{p\overline{F}_Y(x^{1/r};\ \alpha,\ \theta)}{1-\overline{p}\overline{F}_Y(x^{1/r};\ \alpha,\ \theta)}\, dx. \end{aligned}$$

Since $0 < \overline{p}\overline{F}_Y(x^{1/r};\ \alpha,\ \theta) < 1$ for $x > 0$, using the generalized binomial formula given by

$$(1-z)^{-s} = \sum_{j=0}^{\infty} \binom{j+s-1}{j} z^j, \quad |z|< 1, \ \ s > 0, \tag{7}$$

we get for s=1

$$\begin{aligned} \mathbb{E}(X^r) &= \int_0^\infty p\overline{F}_Y(x^{1/r};\ \alpha,\ \theta)\left(1-\overline{p}\overline{F}_Y(x^{1/r};\ \alpha,\ \theta)\right)^{-1}\, dx, \\ &= \int_0^\infty \sum_{j=0}^{\infty} p\overline{p}^j\overline{F}_Y^{j+1}(x^{1/r};\ \alpha,\ \theta)\, dx, \\ &= \int_0^\infty \sum_{j=0}^{\infty} p\overline{p}^j\mathbb{P}(\min(Y_1, Y_2, ..., Y_{j+1}) > x^{1/r})\, dx. \end{aligned}$$

Relying on the Theorem of Fubini-Tonelli, we get

$$\begin{aligned} \mathbb{E}(X^r) &= \sum_{j=0}^{\infty} p\overline{p}^j \int_0^\infty \mathbb{P}(\min(Y_1, Y_2, ..., Y_{j+1}) > x^{1/r})\, dx, \\ &= \sum_{j=0}^{\infty} p\overline{p}^j\mathbb{E}((\min(Y_1, Y_2, ..., Y_{j+1}))^r). \end{aligned}$$

Denote $Z_{j+1} = \min(Y_1, Y_2, ..., Y_{j+1})$, therefore

$$\mathbb{E}(X^r) = \sum_{j=0}^{\infty} p\overline{p}^j\mathbb{E}(Z_{j+1}^r),$$

where, $\mathbb{E}(Z_{j+1}^r) = \int_0^\infty x^r(j+1)\overline{F}_Y^j(x;\ \alpha,\ \theta)f_Y(x;\ \alpha,\ \theta)\, dx$ can be computed numerically using Monte Carlo method. □

In particular, the first two moments of X are

$$\mu_1' = \mathbb{E}(X) = \sum_{j=0}^{\infty} p\overline{p}^j\mathbb{E}(Z_{j+1}). \tag{8}$$

and

$$\mu_2' = \mathbb{E}(X^2) = \sum_{j=0}^{\infty} p\overline{p}^j\mathbb{E}(Z_{j+1}^2). \tag{9}$$

The k^{th} central moment of the $MOEGaL(\alpha,\ p,\ \theta)$ distribution is given by:

$$\mu_k = \mathbb{E}[(X-\mu_1')^k] = \sum_{r=0}^{k} \binom{k}{r} \mu_r'(-\mu_1')^{k-r}.$$

So, the kurtosis and skewness coefficients are expressed as $\gamma_1 = \frac{\mu_4}{\mu_2^2}$ and $\gamma_2 = \frac{\mu_3}{\sqrt{\mu_2^3}}$. The moments of $MOEGaL(\alpha,\ p,\ \theta)$ distribution for different values of α, p and θ using Monte Carlo method are displayed in Table (1).

Parameters	$\mathbb{E}(X)$	Variance	Kurtosis	Skewness
$\alpha = 2, p = 0.6, \theta = 2$	0.662	0.378	7.811	1.815
$\alpha = 1.5, p = 0.8, \theta = 5$	0.260	0.060	8.295	1.854
$\alpha = 5, p = 0.4, \theta = 3$	0.413	0.141	8.694	1.960
$\alpha = 3.2, p = 0.2, \theta = 0.8$	1.127	1.440	13.685	2.647

Table 1: Moments of MOEGaL distribution for various values of α, p and θ using Monte Carlo method.

From Table (1), we can see that MOEGaL AR(1) distribution is positively skewed and leptokurtic. In addition, when θ is increasing, the mean and variance are decreasing.

3. Autoregressive time series modeling

In the current section, we develop minification autoregressive process of order one, where MOEGaL is the stationary marginal distribution. This process is denoted MOEGaL AR(1) process and some of its statistical properties are investigated.

3.1. Construction of the MOEGaL AR(1) process

Let $\{X_n,\ n \in \mathbb{N}^*\}$, be a first order autoregressive ($AR(1)$) model

$$X_n = \begin{cases} Y_n & \text{with} \quad \text{probability} \quad p \\ \min(X_{n-1}, Y_n) & \text{with} \quad \text{probability} \quad 1-p \end{cases} \tag{10}$$

where $p \in (0,1)$ and $\{Y_n\}$ is a sequence of independent and identically distributed random variables independent of $\{X_i,\ i < n\}$. This process is stationary Markovian with Marshall-Olkin distribution as marginal.
The following theorem characterizes the first order autoregressive process using Marshall-Olkin Extended Gamma Lindley distribution (MOEGaL AR(1)).

Theorem 3.1. *In autoregressive process with structure given by* (10)*, $\{X_n\}$ is stationary Markovian with $MOEGaL(\alpha,\ p,\ \theta)$ distribution if and only if $\{Y_n\}$ is distributed as $GaL(\alpha,\ \theta)$ distribution and $X_0 \overset{d}{=} MOEGaL(\alpha,\ p,\ \theta)$.*

Proof. $'' \Rightarrow ''$

Based on (10), the survival function of X_n is expressed as

$$\begin{aligned} \overline{F}_{X_n}(x;\ \alpha,\ p,\ \theta) &= \mathbb{P}(X_n > x), \\ &= p\mathbb{P}(Y_n > x) + (1-p)\mathbb{P}(Y_n > x)\mathbb{P}(X_{n-1} > x), \\ &= p\overline{F}_{Y_n}(x;\ \alpha,\ \theta) + (1-p)\overline{F}_{X_{n-1}}(x;\ \alpha,\ p,\ \theta)\overline{F}_{Y_n}(x;\ \alpha,\ \theta). \end{aligned}$$

If $\{X_n\}$ is stationary with $MOEGaL(\alpha,\ p,\ \theta)$ distribution, then

$$\begin{aligned}\overline{F}_{Y_n}(x;\ \alpha,\ \theta) &= \frac{\overline{F}_{X_n}(x;\ \alpha,\ p,\ \theta)}{p+(1-p)\overline{F}_{X_n}(x;\ \alpha,\ p,\ \theta)},\\ &= \frac{1}{\dfrac{p}{\overline{F}_{X_n}(x;\ \alpha,\ p,\ \theta)}+1-p},\\ &= \frac{1}{p\left(\dfrac{\alpha(1+\theta)-(1-p)((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{p((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}\right)+(1-p)},\\ &= \frac{((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)}.\end{aligned}$$

Then Y_n follows the Gamma Lindley distribution with parameters $\theta > 0$ and $\alpha > \dfrac{\theta}{1+\theta}$.
$'' \Leftarrow ''$

$$\begin{aligned}\overline{F}_{X_1}(x;\ \alpha,\ p,\ \theta) &= p\overline{F}_{Y_1}(x;\ \alpha,\ \theta)+(1-p)\overline{F}_{X_0}(x;\ \alpha,\ p,\ \theta)\overline{F}_{Y_1}(x;\ \alpha,\ \theta),\\ &= \overline{F}_{Y_1}(x;\ \alpha,\ \theta)\left(p+(1-p)\overline{F}_{X_0}(x;\ \alpha,\ p,\ \theta)\right),\\ &= \frac{((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)}\left(p+\frac{(1-p)p((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)-(1-p)((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}\right),\\ &= \frac{p((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)-(1-p)((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}.\end{aligned}$$

Therefore by induction, we can establish that $X_n \overset{d}{=} MOEGaL(\alpha,\ p,\ \theta)$. Which implies that the process $\{X_n\}$ is stationary with Marshall Olkin Extended Gamma Lindley distribution. □

3.2. Statistical properties of the MOEGaL AR(1) process

In this part, we explore some of the important properties of the MOEGaL AR(1) process such as the first order autocovariance function, the joint survival function and the joint probability density function of $(X_n,\ X_{n+1})$.

3.2.1. Autocovariance function of the MOEGaL AR(1) process

We provide all the expressions for the first order autocovariance function of the MOEGaL AR(1) process in the following proposition.
More precisely, we compute $I_1 = \mathbb{E}(X_n^2 \mathbb{1}_{(X_n>Y_{n+1})})$, $I_2 = \mathbb{E}(Y_{n+1}\mathbb{1}_{(X_n=x>Y_{n+1})}|X_n = x)$, $I_3 = \mathbb{E}(X_n Y_{n+1}\mathbb{1}_{(X_n>Y_{n+1})})$ and $I_4 = \mathbb{E}(\min(X_n, Y_{n+1}))$.

Proposition 3.2. *Assume that the process* $(X_n)_{n\in\mathbb{N}}$ *satisfies* (10), *then*

(i)

$$\begin{aligned}I_1 &= \mathbb{E}(X_n^2\mathbb{1}_{(X_n>Y_{n+1})}),\\ &= \mu_2' - \sum_{j=0}^{\infty}\frac{(j+1)p(1-p)^j}{j+2}\mathbb{E}(Z_{j+2}^2).\end{aligned}$$

(ii)

$$\begin{aligned} I_2 &= \mathbb{E}(Y_{n+1}\mathbb{1}_{(X_n=x>Y_{n+1})}|X_n = x), \\ &= (-\theta + \frac{\theta^2}{\alpha(1+\theta)})x^2e^{-\theta x} + (2\theta - \frac{\theta^2}{\alpha(1+\theta)})(-\frac{x}{\theta}e^{-\theta x} - \frac{1}{\theta^2}e^{-\theta x} + \frac{1}{\theta^2}). \end{aligned}$$

(iii)

$$\begin{aligned} I_3 &= \mathbb{E}(X_nY_{n+1}\mathbb{1}_{(X_n>Y_{n+1})}), \\ &= (\frac{2}{\theta} - \frac{1}{\alpha(1+\theta)})\mu_1' + (-\theta + \frac{\theta^2}{\alpha(1+\theta)})\int_0^\infty x^3e^{-\theta x}f_{X_n}(x;\ \alpha,\ p,\ \theta)dx \\ &\quad +(2\theta - \frac{\theta^2}{\alpha(1+\theta)})(\int_0^\infty (\frac{-x^2}{\theta} - \frac{x}{\theta^2})e^{-\theta x}f_{X_n}(x;\ \alpha,\ p,\ \theta)dx. \end{aligned}$$

(iv)

$$I_4 = \mathbb{E}(\min(X_n, Y_{n+1})) = \frac{\mathbb{E}(X_{n+1}) - p\mathbb{E}(Y_{n+1})}{1-p}.$$

Proof.

(i)

$$\begin{aligned} \mathbb{E}(X_n^2\mathbb{1}_{(X_n>Y_{n+1})}) &= \int_0^\infty \int_{\{x>y>0\}} x^2 f_{X_n}(x;\ \alpha,\ p,\ \theta)f_{Y_{n+1}}(x;\ \alpha,\ \theta)dx\,dy, \\ &= \int_0^\infty x^2 f_{X_n}(x;\ \alpha,\ p,\ \theta)F_{Y_{n+1}}(x;\ \alpha,\ \theta)dx, \\ &= \mu_2' - \int_0^\infty x^2 \frac{pf_{Y_n}(x;\ \alpha,\ \theta)}{(1-\overline{p}\overline{F}_{Y_n}(x;\ \alpha,\ \theta))^2}\overline{F}_{Y_{n+1}}(x;\ \alpha,\ \theta)dx. \end{aligned}$$

Since $0 < (1-p)\overline{F}_{Y_n}(x;\ \alpha,\ \theta) < 1$ and using the fact that the generalized binomial formula (7) for $s = 2$, we get

$$\begin{aligned} \mathbb{E}(X_n^2\mathbb{1}_{(X_n>Y_{n+1})}) &= \mu_2' - \int_0^\infty \sum_{j=0}^\infty (j+1)p(1-p)^j x^2 f_{Y_n}(x;\ \alpha,\ \theta)\overline{F}_{Y_n}^{j+1}(x;\ \alpha,\ \theta)dx, \\ &= \mu_2' - \sum_{j=0}^\infty \frac{(j+1)p(1-p)^j}{j+2}\mathbb{E}(Z_{j+2}^2). \end{aligned}$$

(ii) Using Theorem (3.1) and an integration by parts, we get

$$\begin{aligned} \mathbb{E}(Y_{n+1}\mathbb{1}_{(X_n=x>Y_{n+1})}|X_n = x) &= \int_{\{x>y>0\}} yf_{Y_{n+1}|X_n=x}(y;\ \alpha,\ \theta)dy, \\ &= \int_0^x \frac{y\theta^2((\alpha+\alpha\theta-\theta)y+1)e^{-\theta y}}{\alpha(1+\theta)}dy, \\ &= \frac{\theta^2(\alpha+\alpha\theta-\theta)}{\alpha(1+\theta)}\int_0^x y^2e^{-\theta y}dy + \frac{\theta^2}{\alpha(1+\theta)}\int_0^x ye^{-\theta y}dy, \\ &= (-\theta + \frac{\theta^2}{\alpha(1+\theta)})x^2e^{-\theta x} + (2\theta - \frac{\theta^2}{\alpha(1+\theta)})(-\frac{x}{\theta}e^{-\theta x} - \frac{1}{\theta^2}e^{-\theta x} + \frac{1}{\theta^2}). \end{aligned} \tag{11}$$

(iii) Using some properties of conditional means, we obtain

$$\begin{aligned}
\mathbb{E}(X_nY_{n+1}\mathbb{1}_{(X_n>Y_{n+1})}) &= \mathbb{E}(\mathbb{E}(X_nY_{n+1}\mathbb{1}_{(X_n>\epsilon_{n+1})}|X_n)),\\
&= \mathbb{E}(X_n\mathbb{E}(Y_{n+1}\mathbb{1}_{(X_n>Y_{n+1})}|X_n)),\\
&= \mathbb{E}(X_n((-\theta+\frac{\theta^2}{\alpha(1+\theta)})X_n^2e^{-\theta X_n}+(2\theta-\frac{\theta^2}{\alpha(1+\theta)})(\frac{-X_n}{\theta}e^{-\theta X_n}-\frac{1}{\theta^2}e^{-\theta X_n}+\frac{1}{\theta^2}))),\\
&= (\frac{2}{\theta}-\frac{1}{\alpha(1+\theta)})\mu_1'+(-\theta+\frac{\theta^2}{\alpha(1+\theta)})\int_0^\infty x^3e^{-\theta x}f_{X_n}(x;\ \alpha,\ p,\ \theta)dx\\
&\quad+(2\theta-\frac{\theta^2}{\alpha(1+\theta)})\int_0^\infty(\frac{-x^2}{\theta}-\frac{x}{\theta^2})e^{-\theta x}f_{X_n}(x;\ \alpha,\ p,\ \theta)dx.
\end{aligned}$$

In order to compute numerically each of these integrals, Monte Carlo method can be employed.

(iv)
From (10), we obtain $\mathbb{E}(X_{n+1}) = p\mathbb{E}(Y_{n+1}) + (1-p)\mathbb{E}(\min(X_n, Y_{n+1}))$. Therefore

$$\mathbb{E}(\min(X_n, Y_{n+1})) = \frac{\mathbb{E}(X_{n+1}) - p\mathbb{E}(Y_{n+1})}{1-p}.$$

□

Now, we compute the first order autocovariance function of the MOEGaL AR(1) process based on the Proposition (3.2).

Proposition 3.3. *The autocovariance between the random variables X_{n+1} and X_n of MOEGaL AR(1) process is given by*

$$\mathrm{Cov}(X_{n+1}, X_n) = (1-p)\left(\mu_2' - I_1 + I_3 - I_4\mu_1'\right),$$

with I_1, I_3, I_4 are the expressions provided in the above proposition and μ_1', μ_2' are the first two moments of MOEGaL distribution given by (8) *and* (9) *respectively.*

Proof. From (10), one obtains

$$\mathrm{Cov}(X_{n+1}, X_n) = p\mathrm{Cov}(X_n, Y_{n+1}) + (1-p)\mathrm{Cov}(\min(X_n, Y_{n+1}), X_n).$$

As X_n and Y_{n+1} are independent, thus

$$\begin{aligned}
\mathrm{Cov}(X_{n+1}, X_n) &= (1-p)\mathrm{Cov}(\min(X_n, Y_{n+1}), X_n),\\
&= (1-p)\left(\mathbb{E}(\min(X_n, Y_{n+1})X_n) - \mathbb{E}(\min(X_n, Y_{n+1}))\mathbb{E}(X_n)\right),\\
&= (1-p)\left(\mathbb{E}(X_n^2\mathbb{1}(X_n < Y_{n+1})) + \mathbb{E}(X_nY_{n+1}\mathbb{1}(X_n > Y_{n+1})) - \mathbb{E}(\min(X_n, Y_{n+1}))\mathbb{E}(X_n)\right),\\
&= (1-p)\left(\mathbb{E}(X_n^2) - \mathbb{E}(X_n^2\mathbb{1}(X_n > Y_{n+1})) + \mathbb{E}(X_nY_{n+1}\mathbb{1}(X_n > Y_{n+1})) - \mathbb{E}(\min(X_n, Y_{n+1}))\mathbb{E}(X_n)\right).
\end{aligned}$$

□

In the below proposition, we compute the conditional mean and the conditional cumulative distribution function of X_{n+1} given $X_n = x$.

Proposition 3.4. *Let (X_n) be the MOEGaL AR(1) process, then the conditional mean and the conditional cumulative distribution function of X_{n+1} given $X_n = x$ are expressed as follows*

(i)

$$\mathbb{E}(X_{n+1}|X_n = x) = p\left(\frac{2(\alpha-1)}{\alpha\theta} + \frac{(\theta+3)}{\alpha\theta(\theta+1)}\right) + (1-p)x\left(\frac{((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)}\right)$$
$$+(1-p)\left((-\theta + \frac{\theta^2}{\alpha(1+\theta)})x^2 e^{-\theta x} + (2\theta - \frac{\theta^2}{\alpha(1+\theta)})(-\frac{x}{\theta}e^{-\theta x} - \frac{1}{\theta^2}e^{-\theta x} + \frac{1}{\theta^2})\right). \quad (12)$$

(ii)

$$F_{X_{n+1}|X_n=x_1}(x_2) = \frac{((\theta\alpha+\alpha-\theta)(\theta x_2+1)+\theta)e^{-\theta x_2}}{\alpha(1+\theta)}\mathbb{1}_{x_1>x_2} + \left((1-p) + p\frac{((\theta\alpha+\alpha-\theta)(\theta x_2+1)+\theta)e^{-\theta x_2}}{\alpha(1+\theta)}\right)\mathbb{1}_{x_1\le x_2}. \quad (13)$$

Proof.

(i) Based on (10), the conditional mean is given by

$$\begin{aligned}\mathbb{E}(X_{n+1}|X_n = x) &= p\mathbb{E}(Y_{n+1}) + (1-p)\mathbb{E}(\min(X_n, Y_{n+1})|X_n = x),\\ &= p\mathbb{E}(Y_{n+1}) + (1-p)\mathbb{E}\left(x\mathbb{1}_{(Y_{n+1}\ge x)} + Y_{n+1}\mathbb{1}_{(x>Y_{n+1})}|X_n = x\right).\end{aligned}$$

As X_n and Y_{n+1} are independent, thus

$$\begin{aligned}\mathbb{E}(X_{n+1}|X_n = x) &= p\mathbb{E}(Y_{n+1}) + (1-p)\left(x\mathbb{P}(Y_{n+1}\ge x) + \mathbb{E}(Y_{n+1}\mathbb{1}_{(x>Y_{n+1})})\right),\\ &= p\mathbb{E}(Y_{n+1}) + (1-p)\left(x\overline{F}_{Y_{n+1}}(x;\ \alpha,\ \theta) + \mathbb{E}(Y_{n+1}\mathbb{1}_{(x>Y_{n+1})})\right).\end{aligned}$$

Where, $\mathbb{E}(Y_{n+1})$, $\overline{F}_{Y_{n+1}}(x;\ \alpha,\ \theta)$ and $\mathbb{E}(Y_{n+1}\mathbb{1}_{(x>Y_{n+1})})$ are given respectively by the Equations (6), (4) and (11), therefore

$$\mathbb{E}(X_{n+1}|X_n = x) = p\left(\frac{2(\alpha-1)}{\alpha\theta} + \frac{(\theta+3)}{\alpha\theta(\theta+1)}\right) + (1-p)x\left(\frac{((\theta\alpha+\alpha-\theta)(\theta x+1)+\theta)e^{-\theta x}}{\alpha(1+\theta)}\right)$$
$$+(1-p)\left((-\theta + \frac{\theta^2}{\alpha(1+\theta)})x^2 e^{-\theta x} + (2\theta - \frac{\theta^2}{\alpha(1+\theta)})(-\frac{x}{\theta}e^{-\theta x} - \frac{1}{\theta^2}e^{-\theta x} + \frac{1}{\theta^2})\right).$$

(ii) Relying on (10), the conditional cumulative distribution function of X_{n+1} given $X_n = x$ is expressed as

$$\begin{aligned}F_{X_{n+1}|X_n=x_1}(x_2) &= \mathbb{P}(X_{n+1}\le y|X_n = x_1),\\ &= (1-p)\mathbb{P}(\min(X_n, Y_{n+1})\le x_2|X_n = x_1) + p\mathbb{P}(Y_{n+1}\le x_2|X_n = x_1).\end{aligned}$$

Since, X_n and Y_{n+1} are independent, hence

$$\begin{aligned}F_{X_{n+1}|X_n=x_1}(x_2) &= (1-p)\mathbb{P}(\min(x_1, Y_{n+1})\le x_2|X_n = x_1) + p\mathbb{P}(Y_{n+1}\le x_2),\\ &= (1-p)(1-\mathbb{P}(\min(x_1, Y_{n+1}) > x_2)) + pF_{Y_{n+1}}(x_2;\ \alpha,\ \theta),\\ &= (1-p)(1-\mathbb{P}(x_1 > x_2,\ Y_{n+1} > x_2)) + pF_{Y_{n+1}}(x_2;\ \alpha,\ \theta).\end{aligned}$$

Therefore

$$F_{X_{n+1}|X_n=x_1}(x_2;\ \alpha,\ p,\ \theta) = \begin{cases} F_{Y_{n+1}}(x_2;\ \alpha,\ \theta), & \text{if } x_1 > x_2,\\ 1-p+pF_{Y_{n+1}}(x_2;\ \alpha,\ \theta), & \text{if } x_1 \le x_2.\end{cases}$$

□

3.2.2. The joint probability distribution of $(X_n,\ X_{n+1})$

In this part, we derive the joint probability distribution of the couple of random variables $(X_n,\ X_{n+1})$ due to its importance in statistical analysis and applications. We begin first by computing its survival function for MOEGaL AR(1) process in the following theorem.

Theorem 3.5. *Assume that the process* $(X_n)_{n\in\mathbb{N}}$ *satisfies* (16), *then the joint survival function of the couple of random variables* $(X_n,\ X_{n+1})$ *is given by*

$$S(x_1,\ x_2;\ \alpha,\ p,\ \theta) = p\overline{F}_Y(x_1;\ \alpha,\ \theta)\overline{F}_X(x_2;\ \alpha,\ p,\ \theta) + (1-p)\overline{F}_Y(x_1;\ \alpha,\ \theta)S_0(x_1,\ x_2;\ \alpha,\ p,\ \theta), \tag{14}$$

where

$\alpha > \dfrac{\theta}{1+\theta},\ 0 < p \le 1,\ \theta > 0$ *and* $S_0(x_1, x_2;\ \alpha,\ p,\ \theta) = \mathbb{P}(X_n > \max(x_1, x_2)) = \begin{cases} \overline{F}_X(x_1;\ \alpha,\ p,\ \theta), & x_1 \ge x_2, \\ \overline{F}_X(x_2;\ \alpha,\ p,\ \theta), & x_1 < x_2. \end{cases}$

Proof. Using (16) and the stationary property of $(X_n)_{n\in\mathbb{N}}$, we have

$$\begin{aligned} S(x_1,\ x_2;\ \alpha,\ p,\ \theta) &= \mathbb{P}(X_{n+1} > x_1, X_n > x_2), \\ &= p\mathbb{P}(Y_{n+1} > x_1, X_n > x_2) + (1-p)\mathbb{P}(\min(X_n, Y_{n+1}) > x_1, X_n > x_2), \\ &= p\mathbb{P}(Y_{n+1} > x_1)\mathbb{P}(X_n > x_2) + (1-p)\mathbb{P}(Y_{n+1} > x_1)\mathbb{P}(X_n > x_1, X_n > x_2), \\ &= p\mathbb{P}(Y_{n+1} > x_1)\mathbb{P}(X_n > x_2) + (1-p)\mathbb{P}(Y_{n+1} > x_1)\mathbb{P}(X_n > \max(x_1, x_2)), \\ &= p\overline{F}_Y(x_1;\ \alpha,\ \theta)\overline{F}_X(x_2;\ \alpha,\ p,\ \theta) + (1-p)\overline{F}_Y(x_1;\ \alpha,\ \theta)S_0(x_1, x_2;\ \alpha,\ p,\ \theta), \end{aligned}$$

where

$$S_0(x_1, x_2;\ \alpha,\ p,\ \theta) = \mathbb{P}(X_n > \max(x_1, x_2)) = \begin{cases} \overline{F}_X(x_1;\ \alpha,\ p,\ \theta), & x_1 \ge x_2, \\ \overline{F}_X(x_2;\ \alpha,\ p,\ \theta), & x_1 < x_2. \end{cases}$$

□

Note that the joint survival function of random variables X_n and X_{n+1} is not continuous. We have

$$\mathbb{P}(X_n = X_{n+1}) = (1-p)\mathbb{P}(Y_n \ge X_{n+1}) = (1-p)\int_0^{+\infty} \mathbb{P}(Y_n \ge x) f_{X_{n+1}}(x;\ \alpha,\ p,\ \theta)\, dx = \frac{1-p+p\log(p)}{1-p} > 0.$$

In the following theorem, we prove that the joint probability distribution of $(X_n,\ X_{n+1})$ for MOEGaL AR(1) process is the sum of an absolutely continuous and singular measures.

Theorem 3.6. *Assume that* $(X_n)_{n\in\mathbb{N}}$ *satisfies* (16), *then the joint probability distribution of* $(X_n,\ X_{n+1})$ *is given by*

$$P_{(X_n, X_{n+1})}(dx_1, dx_2) = f(x_1,\ x_2;\ \alpha,\ p,\ \theta)dx_1dx_2 + f_0(x_1;\ \alpha,\ p,\ \theta)dx_1 \otimes \delta_{x_1}(dx_2), \tag{15}$$

where,

$f(x_1, x_2;\ \alpha,\ p,\ \theta) = pf_Y(x_1;\ \alpha,\ \theta)f_X(x_2;\ \alpha,\ p,\ \theta) + (1-p)f_Y(x_1;\ \alpha,\ \theta)f_X(x_2;\ \alpha,\ p,\ \theta)\mathbb{1}_{x_1<x_2}$,

$f_0(x_1;\ \alpha,\ p,\ \theta) = (1-p)f_X(x_1;\ \alpha,\ p,\ \theta)\overline{F}_Y(x_1;\ \alpha,\ \theta)$ *and* δ_{x_1} *denotes the Dirac measure at* x_1.

Proof. To find the absolutely continuous part of $P_{(X_n,X_{n+1})}(dx_1,dx_2)$, we compute

$$\begin{aligned} f(x_1,x_2;\ \alpha,\ p,\ \theta) &= \frac{\partial^2 S(x_1,x_2;\ \alpha,\ p,\ \theta)}{\partial x_1 \partial x_2}, \\ &= p f_Y(x_1;\ \alpha,\ \theta) f_X(x_2;\ \alpha,\ p,\ \theta) + (1-p) f_Y(x_1;\ \alpha,\ \theta) f_X(x_2;\ \alpha,\ p,\ \theta) \mathbb{1}_{x_1<x_2}. \end{aligned}$$

Since that $\int_0^\infty \int_0^\infty f(x_1,x_2;\ \alpha,\ p,\ \theta)\ dx_1 dx_2 < 1$, it is clear that the joint probability distribution of X_n and X_{n+1} can be written as a mixture of an absolute continuous and singular measures.
In order to prove (15), we will demonstrate that

$$\int_{x_1}^\infty \int_{x_2}^\infty f(u_1,u_2;\ \alpha,\ p,\ \theta) du_1 du_2 + f_0(u_1;\ \alpha,\ p,\ \theta) du_1 \otimes \delta_{u_1}(du_2) = S(x_1,\ x_2;\ \alpha,\ p,\ \theta),$$

with $S(x_1,\ x_2;\ \alpha,\ p,\ \theta)$ is given by (14). We distinguish two cases:
Case $1 : x_1 \geq x_2$

$$\begin{aligned} &\int_{x_1}^\infty \int_{x_2}^\infty f(u_1,\ u_2;\ \alpha,\ p,\ \theta) \mathrm{d}u_1 \mathrm{d}u_2 \\ &+ \int_{x_1}^\infty \int_{x_2}^\infty f_0(u_1;\ \alpha,\ p,\ \theta) du_1 \otimes \delta_{u_1}(du_2) = \int_{x_1}^\infty \int_{x_2}^\infty p f_Y(u_1;\ \alpha,\ \theta) f_X(u_2;\ \alpha,\ p,\ \theta) du_1 du_2 \\ &\quad + \int_{x_1}^\infty \int_{x_2}^\infty (1-p) f_Y(u_1;\ \alpha,\ \theta) f_X(u_2;\ \alpha,\ p,\ \theta) \mathbb{1}_{u_1<u_2} du_1 du_2 \\ &\quad + \int_{x_1}^\infty (1-p) f_X(u_1;\ \alpha,\ p,\ \theta) \overline{F}_Y(u_1;\ \alpha,\ \theta) \left(\int_{x_2}^\infty \delta_{u_1}(du_2)\right) du_1. \end{aligned}$$

Using the fact that,

$$\int_{x_2}^\infty \delta_{u_1}(du_2) = \delta_{u_1}([x_2,+\infty[)$$

one gets,

$$\begin{aligned} &\int_{x_1}^\infty \int_{x_2}^\infty f(u_1,\ u_2;\ \alpha,\ p,\ \theta) \mathrm{d}u_1 \mathrm{d}u_2 \\ &+ \int_{x_1}^\infty \int_{x_2}^\infty f_0(u_1;\ \alpha,\ p,\ \theta) du_1 \otimes \delta_{u_1}(du_2) = p\overline{F}_Y(x_1;\ \alpha,\ \theta) \overline{F}_X(x_2;\ \alpha,\ p,\ \theta) \\ &\quad +(1-p) \int_{x_1}^\infty \overline{F}_X(u_1;\ \alpha,\ p,\ \theta) f_Y(u_1;\ \alpha,\ \theta) du_1 \\ &\quad +(1-p) \int_{x_1 \vee x_2}^\infty f_X(u_1;\ \alpha,\ p,\ \theta) \overline{F}_Y(u_1;\ \alpha,\ \theta) du_1, \\ &= p\overline{F}_Y(x_1;\ \alpha,\ \theta) \overline{F}_X(x_2;\ \alpha,\ p,\ \theta) \\ &\quad +(1-p) \int_{x_1}^\infty \overline{F}_X(u_1;\ \alpha,\ p,\ \theta) f_Y(u_1;\ \alpha,\ \theta) du_1 \\ &\quad +(1-p) \int_{x_1}^\infty f_X(u_1;\ \alpha,\ p,\ \theta) \overline{F}_Y(u_1;\ \alpha,\ \theta) du_1. \end{aligned}$$

Using integration by parts, one obtains,

$$
\begin{aligned}
(1-p)\int_{x_1}^{\infty} f_X(u_1;\ \alpha,\ p,\ \theta)\overline{F}_Y(u_1;\ \alpha,\ \theta)du_1 &= (1-p)[-\overline{F}_Y(x_1;\ \alpha,\ \theta)F_X(x_1,\ \alpha,\ p,\ \theta) \\
&\quad + \int_{x_1}^{\infty} F_X(u_1;\ \alpha,\ p,\ \theta)f_Y(u_1;\ \alpha,\ \theta)du_1], \\
&= -(1-p)\overline{F}_Y(x_1;\ \alpha,\ \theta)(1-\overline{F}_X(x_1;\ \alpha,\ p,\ \theta)) \\
&\quad +(1-p)\int_{x_1}^{\infty} f_Y(u_1;\ \alpha,\ \theta)du_1 \\
&\quad -(1-p)\int_{x_1}^{\infty} \overline{F}_X(u_1;\ \alpha,\ p,\ \theta)f_Y(u_1;\ \alpha,\ \theta)du_1, \\
&= (1-p)\overline{F}_Y(x_1,\ \alpha,\ \theta)\overline{F}_X(x_1;\ \alpha,\ p,\ \theta) \\
&\quad -(1-p)\int_{x_1}^{\infty} \overline{F}_X(u_1;\ \alpha,\ p,\ \theta)f_Y(u_1;\ \alpha,\ \theta)du_1.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\int_{x_1}^{\infty}\int_{x_2}^{\infty} f(u_1,\ u_2;\ \alpha,\ p,\ \theta)du_1du_2 + f_0(u_1;\ \alpha,\ p,\ \theta)du_1 \otimes \delta_{u_1}(du_2) &= p\overline{F}_Y(x_1;\ \alpha,\ \theta)\overline{F}_X(x_2;\ \alpha,\ p,\ \theta) \\
&\quad +(1-p)\overline{F}_Y(x_1;\ \alpha,\ \theta)\overline{F}_X(x_1;\ \alpha,\ p,\ \theta), \\
&= S(x_1, x_2;\ \alpha,\ p,\ \theta).
\end{aligned}
$$

Case 2 : $x_1 < x_2$

$$
\begin{aligned}
&\int_{x_1}^{\infty}\int_{x_2}^{\infty} f(u_1,\ u_2;\ \alpha,\ p,\ \theta)du_1du_2 \\
&+\int_{x_1}^{\infty}\int_{x_2}^{\infty} f_0(u_1;\ \alpha,\ p,\ \theta)du_1 \otimes \delta_{u_1}(du_2) = \int_{x_1}^{\infty}\int_{x_2}^{\infty} pf_Y(u_1;\ \alpha,\ \theta)f_X(u_2;\ \alpha,\ p,\ \theta) \\
&\quad + \int_{x_1}^{\infty}\int_{x_2}^{\infty}(1-p)f_Y(u_1;\ \alpha,\ \theta)f_X(u_2;\ \alpha,\ p,\ \theta)\mathbb{1}_{u_1<u_2}du_1du_2 \\
&\quad + \int_{x_1}^{\infty}(1-p)f_X(u_1;\ \alpha, p, \theta)\overline{F}_Y(u_1;\ \alpha, \theta)(\int_{x_2}^{\infty}\delta_{u_1}(du_2))du_1, \\
&= p\overline{F}_Y(x_1;\ \alpha,\ \theta)\overline{F}_X(x_2;\ \alpha,\ p,\ \theta) \\
&\quad +(1-p)\int_{x_2}^{\infty} f_X(u_2;\ \alpha,\ p,\ \theta)[F_Y(u_2;\ \alpha,\ \theta)-F_Y(x_1;\ \alpha,\ \theta)]du_2 \\
&\quad +(1-p)\int_{x_2}^{\infty} f_X(u_1;\ \alpha,\ p,\ \theta)\overline{F}_Y(u_1;\ \alpha,\ \theta)du_1.
\end{aligned}
$$

Since $F_Y(x;\ \alpha,\ \theta) = 1-\overline{F}_Y(x;\ \alpha,\ \theta)$, hence we can write that,

$$
\begin{aligned}
(1-p)\int_{x_2}^{\infty} f_X(u_2;\ \alpha,\ p,\ \theta)[F_Y(u_2;\ \alpha,\ \theta)-F_Y(x_1;\ \alpha,\ \theta)]du_2 &= (1-p)\int_{x_2}^{\infty} f_X(u_2;\ \alpha,\ p,\ \theta)(1-\overline{F}_Y(u_2;\ \alpha,\ \theta))du_2 \\
&\quad -(1-p)(1-\overline{F}_Y(x_1;\ \alpha,\ \theta))\int_{x_2}^{\infty} f_X(u_2;\ \alpha,\ p,\ \theta)du_2.
\end{aligned}
$$

Then, we deduce the desired result. □

4. Statistical inference for the MOEGaL AR(1) process

In this section, we develop statistical inference for the unknown parameters of the MOEGaL AR(1) process using Maximum Likelihood (ML), Conditional Least Squares (CLS) and Method of Moments (MM) estimation methods. Additionally, we conducted a simulation study to evaluate the obtained estimators resulting from these approaches. The experiments are performed with PYTHON software.

4.1. Maximum likelihood estimation method

In this subsection, we consider maximum likelihood estimation methods to estimate the unknown parameters of the MOEGaL AR(1) process.
Based on (15), the distribution of $(X_n,\ X_{n+1})$ is dominated by the following measure

$$v(dx_1,\ dx_2) = \mathbb{1}_{x_1<x_2}dx_1dx_2 + \mathbb{1}_{x_1>x_2}dx_1dx_2 + \mathbb{1}_{x_1=x_2}dx_1 \otimes \delta_{x_1}(dx_2).$$

The joint probability density function of $(X_n,\ X_{n+1})$ can be written as below:

$$f_{(X_n,\ X_{n+1})}(x_1,\ x_2;\ \alpha,p,\theta) = \begin{cases} f_Y(x_1,\ \alpha,\ \theta)f_X(x_2,\ \alpha,\ p,\ \theta) & \text{if} \quad x_1 < x_2 \\ pf_Y(x_1,\ \alpha,\ \theta)f_X(x_2,\ \alpha,\ p,\ \theta) & \text{if} \quad x_1 > x_2 \\ (1-p)f_X(x_1,\ \alpha,\ p,\ \theta)\overline{F}_Y(x_1,\ \alpha,\ \theta) & \text{if} \quad x_1 = x_2 \end{cases} \tag{16}$$

Let $(X_1,\ X_2,...,X_n)$ be a MOEGaL AR(1) time series, $D = \{1,...,n-1\}$, $D_1 = \{i : i \in D,\ x_i < x_{i+1}\}$, $D_2 = \{i : i \in D,\ x_i > x_{i+1}\}$ and $D_3 = \{i : i \in D,\ x_i = x_{i+1}\}$. Then due to conditioning approach and the Markov property, the joint probability density function of $(X_1,\ X_2,...,X_n)$ is given by

$$\begin{aligned} f_{(X_1,\ X_2,...,X_n)}(x_1,...,x_n;\alpha,\ p,\ \theta) &= \frac{f_{(X_{n-1},\ X_n)}(x_{n-1},\ x_n;\ \alpha,p,\theta)}{f_{X_{n-1}}(x_{n-1},\ \alpha,p,\theta)} \times ... \times \frac{f_{(X_1,\ X_2)}(x_1,\ x_2;\ \alpha,p,\theta)}{f_{X_1}(x_1,\ \alpha,p,\theta)} \times f_{X_1}(x_1;\ \alpha,p,\theta), \\ &= \frac{\prod_{i=1}^{n-1} f_{(X_i,\ X_{i+1})}(x_i,\ x_{i+1};\ \alpha,\ p,\ \theta)}{\prod_{i=2}^{n-1} f_{X_i}(x_i;\ \alpha,\ p,\ \theta)}. \end{aligned} \tag{17}$$

Based on (17), the log-likelihood function of $X_1 = x_1,\ X_2 = x_2,...,X_n = x_n$ is expressed as follows,

$$\begin{aligned} \ell(x_1,...,x_n;\ \alpha,\ p,\ \theta) &= \sum_{i\in D_1\cup D_2\cup D_3} \ln(f_{(X_i,\ X_{i+1})}(x_i,\ x_{i+1};\ \alpha,\ p,\ \theta)) - \sum_{2}^{n-1} \ln(f_{X_i}(x_i;\ \alpha,\ p,\ \theta)), \\ &= \sum_{i\in D_1} \ln(f_Y(x_i;\ \alpha,\ \theta)f_X(x_{i+1};\ \alpha,\ p,\ \theta)) + \sum_{i\in D_2} \ln(pf_Y(x_i;\ \alpha,\theta)f_X(x_{i+1};\ \alpha,\ p,\ \theta)) \\ &\quad + \sum_{i\in D_3} \ln((1-p)f_X(x_i;\ \alpha,\ p,\ \theta)\overline{F}_Y(x_i;\ \alpha,\ \theta)) - \sum_{2}^{n-1} \ln(f_{X_i}(x_i;\ \alpha,\ p,\ \theta)). \end{aligned}$$

The maximum likelihood estimators $\widehat{\alpha}$, $\widehat{p}$ and $\widehat{\theta}$ of α, p and θ respectively are obtained by maximizing numerically the log-likelihood function using the gradient descent method. The optimization algorithm is described as follows

Algorithm 1 The maximum likelihood estimation method using Gradient Descent algorithm

Problem: Estimate the unknown parameters $\Theta = (\alpha,\ p,\ \theta)$ of the MOEGaL AR(1) process.
Input: Precision ϵ, step size h, $x_1,\ x_2, ..., x_n$ are drawn from MOEGaL AR(1) process.
Output: The estimates $\widehat{\Theta}_{ML} = (\widehat{\alpha}_{ML},\ \widehat{p}_{ML},\ \widehat{\theta}_{ML})$.

$D = \{1, ..., n-1\}$, $D_1 = \{i : i \in D,\ x_i < x_{i+1}\}$, $D_2 = \{i : i \in D,\ x_i > x_{i+1}\}$ and $D_3 = \{i : i \in D,\ x_i = x_{i+1}\}$.
Initialize Θ^0
$k \leftarrow 0$
repeat
  $\Theta^{k+1} = \Theta^k + h\nabla\ell(x_1,\ x_2, ..., x_n; \Theta^k)$
  $k \leftarrow k+1$
until $||\Theta^{k+1} - \Theta^k|| < \epsilon$
return $\widehat{\Theta}_{ML} = \Theta^k$

4.2. *Conditional least squares*

Let $\{X_i,\ i = 1, ..., n\}$ be a MOEGaL AR(1) time series. The conditional least squares estimators $\widehat{\alpha}_{CLS}$, $\widehat{p}_{CLS}$ and $\widehat{\theta}_{CLS}$ of the unknown parameters α, p and θ are obtained by minimizing numerically the following cost function

$$\psi(\alpha,\ p,\ \theta) = \sum_{i=1}^{n-1} \left(X_{i+1} - \mathbb{E}(X_{i+1}|X_i = x)\right)^2,$$

where $\mathbb{E}(X_{i+1}|X_i = x)$ is given by Equation (12). The algorithm of obtaining the estimators $\widehat{\alpha}_{CLS}$, $\widehat{p}_{CLS}$ and $\widehat{\theta}_{CLS}$ using gradient descent method is described as below

Algorithm 2 The Conditional least squares estimation method using Gradient Descent algorithm

Problem: Estimate the unknown parameters $\Theta = (\alpha,\ p,\ \theta)$ of the MOEGaL AR(1) process.
Input: Precision ϵ, step size h, $x_1,\ x_2, ..., x_n$ are drawn from MOEGaL AR(1) process.
Output: The estimates $\widehat{\Theta}_{CLS} = (\widehat{\alpha}_{CLS},\ \widehat{p}_{CLS},\ \widehat{\theta}_{CLS})$.

Initialize Θ^0
$k \leftarrow 0$
repeat
  $\Theta^{k+1} = \Theta^k - h\nabla\psi(\Theta^k)$
  $k \leftarrow k+1$
until $||\Theta^{k+1} - \Theta^k|| < \epsilon$
return $\widehat{\Theta}_{CLS} = \Theta^k$

4.3. *Method of moments*

Let $\{X_i,\ i = 1, ..., n\}$ be a MOEGaL AR(1) time series. We consider first the estimation of the unknown parameter p. The estimator $\widehat{p}_{MM}$ of p is the solution of the following equation

$$g_1(p) = \frac{1 - p + p\log(p)}{1-p} - \frac{1}{n-1}\sum_{i=1}^{n-1} \mathbb{1}_{(X_i = X_{i+1})} = 0. \quad (18)$$

Then, the estimators $\widehat{\alpha}_{MM}$ and $\widehat{\theta}_{MM}$ of the unknown parameters α and θ are derived by using Monte Carlo simulation and solving the following non-linear system of equations numerically through the Newton-

Raphson method

$$\begin{cases} g_2(\alpha,\ \widehat{p}_{MM},\ \theta) = \dfrac{1}{n-1}\sum_{i=1}^{n-1} X_i - \mathbb{E}(X_n) = 0. \\ \\ g_3(\alpha,\ \widehat{p}_{MM},\ \theta) = \dfrac{1}{n-1}\sum_{i=1}^{n-1} X_i^2 - \mathbb{E}(X_n^2) = 0. \end{cases} \tag{19}$$

Where $\mathbb{E}(X_n)$ and $\mathbb{E}(X_n^2)$ are given by Equations (8) and (9) respectively. We apply the Algorithm 3 in PYTHON software, we get the estimators, $\widehat{\alpha}_{MM}$, $\widehat{p}_{MM}$ and $\widehat{\theta}_{MM}$.

Algorithm 3 Method of moments using Newton-Raphson algorithm

Problem: Estimate the unknown parameters p and $\Theta_1 = (\alpha,\ \theta)$ of the MOEGaL AR(1) process.
Input: Precision ϵ, $x_1,\ x_2, ..., x_n$ are drawn from MOEGaL AR(1) process.
Output: The estimates $\widehat{\alpha}_{MM}$, $\widehat{p}_{MM}$ and $\widehat{\theta}_{MM}$

Step 1
Initialize p^0
$k \leftarrow 0$
repeat
  $p^{k+1} = p^k - \dfrac{g_1(\widehat{p})}{g_1'(\widehat{p})}$
  $k \leftarrow k+1$
until $|p^{k+1} - p^k| < \epsilon$
return $\widehat{p}_{MM} = p^k$
Step 2
Initialize Θ_1^0
$k \leftarrow 0$
repeat
  $\Theta_1^{k+1} = \Theta_1^k - (\nabla(g_2, g_3)(\Theta_1^k,\ \widehat{p}_{MM}))^{-1}\begin{pmatrix} g_2(\Theta_1^k,\ \widehat{p}_{MM}) \\ g_3(\Theta_1^k,\ \widehat{p}_{MM}) \end{pmatrix}$
  $k \leftarrow k+1$
until $\|\Theta_1^{k+1} - \Theta_1^k\| < \epsilon$
return $\widehat{\Theta}_{1MM} = \Theta_1^k$

4.4. Simulation study

In this part, we investigate the sample path behavior of the MOEGaL AR(1) process. Additionally, we conduct a simulation study to evaluate the accuracy of the estimates obtained from the maximum likelihood, conditional least squares and moments estimation methods. For this purpose, we generate $N = 50$ samples of size $n = 100,\ 200,\ 250$ from the proposed process with true parameter values (2, 0.8, 1.5), (1.5, 0.7, 1) and (2.5, 0.3, 0.6). Figure 1 shows that the sample path behavior appears distinct and adjustable through the parameters α, p and θ, this adds a lot of value to the model.

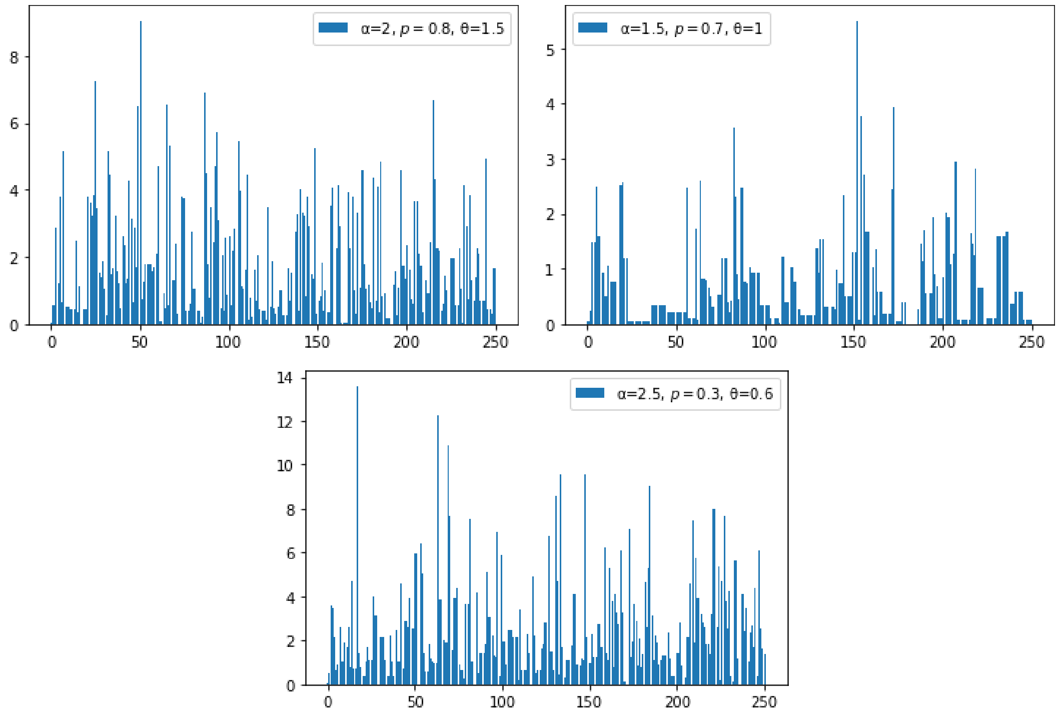


Figure 1: Sample path behavior of MOEGaL AR(1) process for various values of α, p and θ.

We compute the root mean squared error of the estimates by, $RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\widehat{\lambda_k} - \lambda)^2}$ where $\widehat{\lambda_k}$ is the estimator of λ from the k^{th} sample with $\lambda = \alpha$, $\lambda = p$ or $\lambda = \theta$. The obtained estimated values for different approaches are highlighted in the following Tables

n	RMSE	ML	MM	CLS
100	$RMSE(\alpha, \widehat{\alpha})$	0.063	0.093	0.081
	$RMSE(p, \widehat{p})$	0.053	0.086	0.074
	$RMSE(\theta, \widehat{\theta})$	0.055	0.087	0.075
200	$RMSE(\alpha, \widehat{\alpha})$	0.0042	0.052	0.031
	$RMSE(p, \widehat{p})$	0.0034	0.042	0.023
	$RMSE(\theta, \widehat{\theta})$	0.0040	0.052	0.022
250	$RMSE(\alpha, \widehat{\alpha})$	0.00013	0.010	0.0032
	$RMSE(p, \widehat{p})$	0.00064	0.0066	0.0081
	$RMSE(\theta, \widehat{\theta})$	0.00061	0.011	0.0022

Table 2: RMSE of estimated values of $\alpha = 2$, $p = 0.8$, $\theta = 1.5$ by ML, CLS and MM approaches.

n	RMSE	ML	MM	CLS
100	$RMSE(\alpha, \widehat{\alpha})$	0.059	0.093	0.081
	$RMSE(p, \widehat{p})$	0.059	0.088	0.079
	$RMSE(\theta, \widehat{\theta})$	0.056	0.081	0.071
200	$RMSE(\alpha, \widehat{\alpha})$	0.0042	0.035	0.023
	$RMSE(p, \widehat{p})$	0.0043	0.022	0.012
	$RMSE(\theta, \widehat{\theta})$	0.0097	0.051	0.024
250	$RMSE(\alpha, \widehat{\alpha})$	0.00043	0.012	0.0031
	$RMSE(p, \widehat{p})$	0.00071	0.0085	0.0022
	$RMSE(\theta, \widehat{\theta})$	0.00088	0.014	0.0024

Table 3: RMSE of estimated values of $\alpha = 1.5$, $p = 0.7$, $\theta = 1$ by ML, CLS and MM approaches.

n	RMSE	ML	MM	CLS
100	$RMSE(\alpha, \widehat{\alpha})$	0.041	0.081	0.062
	$RMSE(p, \widehat{p})$	0.043	0.084	0.063
	$RMSE(\theta, \widehat{\theta})$	0.061	0.086	0.078
200	$RMSE(\alpha, \widehat{\alpha})$	0.0032	0.041	0.032
	$RMSE(p, \widehat{p})$	0.0021	0.039	0.024
	$RMSE(\theta, \widehat{\theta})$	0.0039	0.042	0.036
250	$RMSE(\alpha, \widehat{\alpha})$	0.00010	0.0083	0.0015
	$RMSE(p, \widehat{p})$	0.00034	0.0092	0.0031
	$RMSE(\theta, \widehat{\theta})$	0.00042	0.015	0.0019

Table 4: RMSE of estimated values of $\alpha = 2.5$, $p = 0.3$, $\theta = 0.6$ by ML, CLS and MM approaches.

Relying on (2), (3) and (4), it is inferred that all estimators are consistent and asymptotically unbiased since the RMSE tend to zero when the sample size n increases. It is also noteworthy that the ML generates the best results and it outperforms MM and CLS methods in terms of RMSE criterion for different parameter values of α, p and θ. From this perspective, it is quite reasonable to use the maximum likelihood estimation method so as to estimate the unknown parameters of the MOEGaL AR(1) process.

5. Application

In this section, we illustrate an application of the MOEGaL AR(1) process with a real data set in order to study the usability and significance of the proposed process.

5.1. Data

We consider the gold price data of Japan for 100 months in Japanese yen from 31 January 2011 to 30 April 2019 (`https://www.kaggle.com/datasets/odins0n/monthly-gold-prices/`). The main idea of this data analysis is to see how the proposed model works in practice and also how it works in predictive modeling. To normalize the data, we divide each value by 10^5. The real data set is presented in Figure 2. From the plots drawn in Figure 3, the autocorrelation function decays exponentially and the partial autocorrelation function is significant at lag 1. The stationarity of the gold data set is validated through the Dickey-Fuller test, with corresponding p-value = $6.54 \times 10^{-5} < 0.05$.

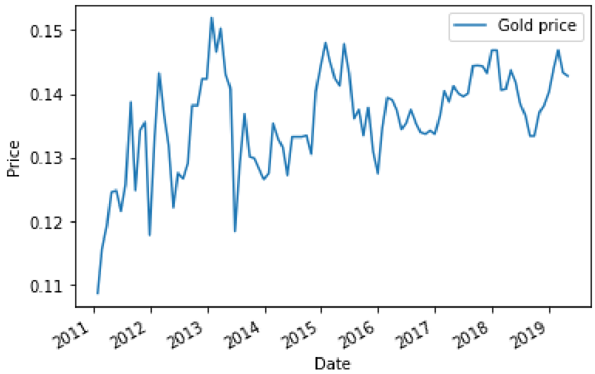
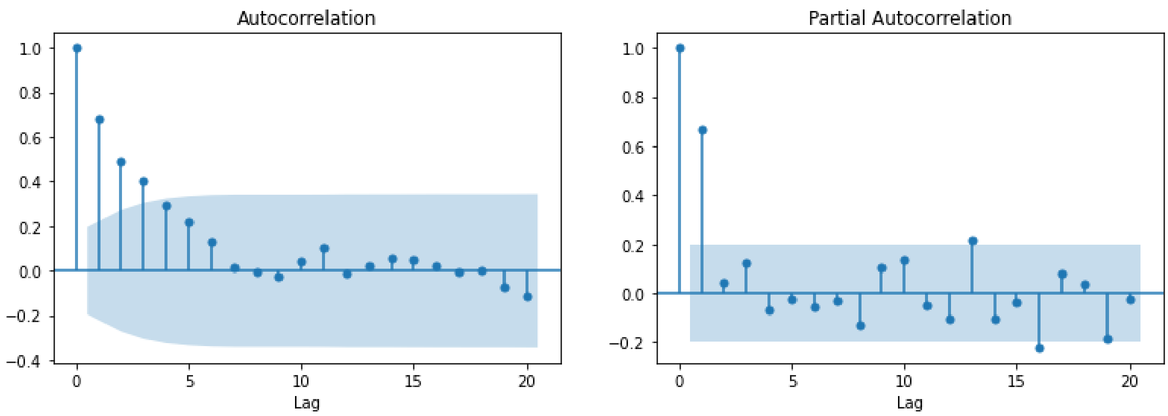


Figure 2: Gold price data in Japan from 31 January 2011 to 30 April 2019



(a) Autocorrelation function (b) Partial autocorrelation function

Figure 3: Autocorrelation function, Partial autocorrelation function plots of the real data

5.2. *Goodness of fit*

In this part, we check how well the real data fit MOEGaL AR(1) process. Referring to [22], we apply the goodness of fit test in order to examine the performance of this model. Therefore, we consider the null hypothesis

$$H_0 : X_{i+1}|X_i \text{ follows } F(.|\alpha,\ p,\ \theta, X_i), \quad i = 0, ..., n-1$$

where $F(.|\alpha,\ p,\ \theta,\ X_i)$ is given by (13).
Under the null hypothesis, we consider the following deviation process

$$U_n(x, y) = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \mathbb{1}_{X_i \leq x}[\mathbb{1}_{X_{i+1} \leq y} - F_{X_{i+1}|X_i}(y)].$$

The basis for a test of H_0 is expressed by the supremum deviation

$$S_n = \sup_{x,y} |U_n(x, y)|.$$

We undertake the following steps. Firstly, we compute $\widehat{\alpha}$, $\widehat{p}$ and $\widehat{\theta}$ the ML estimates of α, p, θ, as well as U_n and S_n from the real data. Secondly, starting with initial value X_0^*, we generate X_{i+1}^* from $F(.|\widehat{\alpha}, \widehat{p}, \widehat{\theta}, X_i^*)$,

$0 \le i \le n-1$. Thirdly, based on the bootstrap pseudo series $(X_i^*)_{1\le i\le n}$, let $U_n^*(x,y)$ be indicated as $U_n(x,y)$ by replacing $\widehat{\alpha}$, $\widehat{p}$, $\widehat{\theta}$ with $\widehat{\alpha}^*$, $\widehat{p}^*$, $\widehat{\theta}^*$. Moreover, we calculate the Bootstrap statistics of S_n which is given by $S_n^* = \sup_{x,y} |U_n^*(x,y)|$. Finally, we repeat the above steps B times in order to obtain the ordered statistics $S_{(1)}^* \le S_{(2)}^* \le ... \le S_{(B)}^*$. Then, for a level of significance $\alpha \in (0,1)$, we calculate $t_{1-\alpha}^*$ the empirical $1-\alpha$ quantile. As a matter of fact, we reject H_0 if $S_n > t_{1-\alpha}^*$.
For $\alpha = 0.05$ the computed test statistics S_n for MOEGaL model is equal to 0.12 with an associated p-value of 0.96. However, for the standard AR(1) model the calculated test statistics S_n is equal to 0.31 with an associated p-value of 0.24. Consequently, these results indicate that the MOEGaL model offers a strong and superior fit for the gold price data compared to the standard AR(1) model.

5.3. Predictions

We evaluate the performance of the proposed autoregressive model and we compare it with the standard form for the AR(1) process, given by $X_n = a_0 + a_1 X_{n-1} + \epsilon_n$, where ϵ_n is a white noise, in terms of predictions. We split the real data into $\{x_1, x_2, ..., x_{80}\}$ the training data and $\{x_{81}, x_{82}, ..., x_{100}\}$ the test data. We estimate the unknown parameters of both models using maximum likelihood estimation method on the training data. The obtained estimators are illustrated in Table (5).

Model	$\widehat{\alpha}$	$\widehat{p}$	$\widehat{\theta}$	$\widehat{a_0}$	$\widehat{a_1}$
MOEGaL AR(1)	2.15	0.72	1.49		
Standard form for the AR(1)				0.048	0.645

Table 5: The parameter estimates of the MOEGaL AR(1) and the Standard form for the AR(1)

Now, we check the performance of models on the test data. The prediction algorithm is summarized as follows:

Algorithm 4 Algorithm for predicting future values using MOEGaL AR(1) process

Input: The estimates $\widehat{\alpha}$, $\widehat{p}$ and $\widehat{\theta}$, x_{n_0}= last value of the train data set, N=size of data set.
Output: Predicted values $\{\widehat{x}_{n_0+1}, \widehat{x}_{n_0+2}, ..., \widehat{x}_{n_0+N}\}$.

Generate random sample $Y = \{y_{n_0+1}, y_{n_0+2}, ..., y_{n_0+N}\}$ from Gamma Lindley distribution with parameters $\widehat{\alpha}$ and $\widehat{\theta}$.
Generate random sample $Z = \{z_{n_0+1}, z_{n_0+2}, ..., z_{n_0+N}\}$ from Bernoulli distribution with parameter $\widehat{p}$.
$\widehat{X}$ be a new array
$\widehat{X}[n_0] \leftarrow x_{n_0}$
for $i = n_0 + 1 \rightarrow n_0 + N$ **do**
  if $Z[i] \leftarrow 1$ **then**
    $\widehat{X}[i] \leftarrow Y[i]$
  else
    $\widehat{X}[i] \leftarrow \min(\widehat{X}[i-1], Y[i])$
  end if
end for
return $\widehat{X}[n_0 + 1 \rightarrow n_0 + N]$

For the purpose of predicting $\widehat{x}_n$, $n_0 + 1 \le n \le n_0 + N$ using MOEGaL AR(1) process, we simulate K samples $\left[\widehat{x}_{n_0+1}^{(i)}, ..., \widehat{x}_{n_0+N}^{(i)}\right]_{1\le i\le K}$ based on the above algorithm, implemented in PYTHON software. Then, we compute the mean predicted value $\widehat{x}_n = \dfrac{1}{K}\sum_{i=1}^{K} \widehat{x}_n^{(i)}$. In the current application, we consider $K = 50$. In order to compare the performance of the MOEGaL AR(1) and the Standard form for the AR(1), we use the mean absolute error (*MAE*) and the mean relative error (*MRE*) as comparison criteria. The *MAE* and *MRE* are expressed respectively by $MAE = \frac{1}{20}\sum_{t=81}^{100}|\widehat{x_n} - x_n|$, $MRE = \frac{1}{20}\sum_{n=81}^{100}\frac{|\widehat{x_n} - x_n|}{x_n}$.

Model	*MAE*	MRE
MOEGaL AR(1)	0.0008	0.001
Standard form for the AR(1)	0.01	0.04

Table 6: *MAE* and *MRE* values of prediction models

From Table (6), the smallest values of *MAE* and *MRE* are obtained from MOEGaL AR(1) model. Thus, we can conclude that it provides the best predictions among the other considered model. Figure 4 illustrates the close fit between the prediction curve generated by the MOEGaL AR(1) model and the test data. However, predictions from the standard AR(1) one exhibit a single decreasing trend, unlike the varied trajectory seen in the prediction curve of MOEGaL AR(1). Hence, Figure 4 supports the results in Table 5, showing that MOEGaL AR(1) is a good predictive model due to its structure.

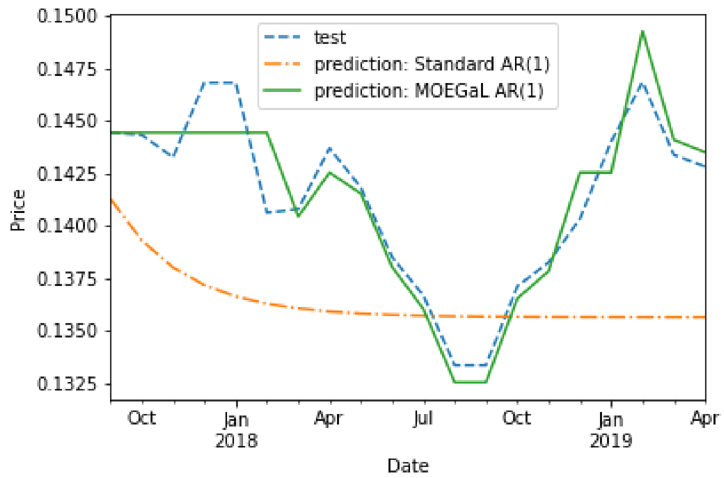


Figure 4: Predictions of time series data using MOEGaL AR(1) process and the Standard form for the AR(1)

6. Conclusion

In the current research, a new first order autoregressive process $(X_n)_{n\in\mathbb{N}}$ using Marshall-Olkin Extended Gamma Lindley distribution (MOEGaL AR(1)) was constructed. We computed the first order autocovariance function, as well as the conditional mean and the conditional cumulative distribution function of X_{n+1} given X_n. We established the joint probability distribution of (X_n, X_{n+1}). Consequently, we estimated the unknown parameters of this process using different estimation methods. We performed a simulation study in order to assess the accuracy of the obtained estimates resulting from these approaches. We developed an algorithm for predicting a gold price time series data by employing MOEGaL AR(1) model. This is suggestive that the proposed autoregressive model is a good predictive model compared to the standard AR(1) one. As a final note, we would assert that the current research work can be extend, taken further and built upon. Indeed as new perspectives we will focus in future on generalizing the MOEGaL AR(1) model to the k^{th} order autoregressive model. Another equally pertinent extension involves introducing a first autoregressive minification process with multivariate Marshall-Olkin Extended Gamma Lindley as marginal.

References

[1] H. S. Baccouch, B. V. Popović, Lindley first-order autoregressive model with applications, Comm. Statist. Theory Methods. **45**(17) (2016), 4988-5006.
[2] A. Beghriche, H. Zeghdoudi, A size biased gamma Lindley distribution, Thai. Stat. **17**(2) (2019), 179-189.
[3] J. T. Eghwerido, J. O. Ogbo, A. E. Omotoye, The Marshall-Olkin Gompertz distribution: properties and applications, Statistica. **81**(2) (2021), 183–215.
[4] M. El-Morshedy, M. S. Eliwa, The odd flexible Weibull-H family of distributions: properties and estimation with applications to complete and upper record data, Filomat. **33**(9) (2019), 2635-2652.
[5] M. E. Ghitany, E.K. Al-Hussaini, R. A. Al-Jarallah, Marshall–olkin extended weibull distribution and its application to censored data, J. Appl. Stat. **32**(10) (2005), 1025-1034.
[6] M. E. Ghitany, D. K. Al-Mutairi, S. M. Aboukhamseen, Estimation of the reliability of a stress-strength system from power lindley distributions, Comm. Statist. Simulation Comput. **44**(1) (2015), 118–136.
[7] M. E. Ghitany, B. Atieh, S. Nadarajah, Lindley distribution and its application, Math. Comput. Simulation. **78**(4) (2008), 493–506.
[8] K. Jakumar, T. Mathew, On a generalization to marshall-olkin scheme and its application to burr type xii distribution, Statist. Papers. **49**(3) (2008), 421.
[9] K. K. Jose, A. Joseph, M. M. RistiĆ: A marshall-olkin beta distribution and its applications, J. Probab. Stat.Sci. **7**(2) (2009),173–186.
[10] E. Krishna, K. K. Jose, M. M. RistiĆ, Applications of marshall–olkin fréchet distribution, Comm. Statist. Simulation Comput. **42**(1) (2013), 76–89.
[11] B. Krishna, D. George, The marshall-olkin weibull truncated negative binomial distribution and its applications, Statistica. **79**(3) (2019), 247–265.
[12] D. Laribi, A. Masmoudi, I. Boutouria, Characterization of generalized Gamma-Lindley distrubution using truncated moments of order statistics, Math. Slovaca. **71**(2) (2021), 455-474.
[13] D. Laribi, A. Masmoudi, I. Boutouria, An Extended Gamma-Lindley Model and Inference for the Prediction of Covid-19 in Tunisia, Math. Slovaca. **73**(4) (2023), 1055-1074.
[14] D. V. Lindley, Fiducial Distributions and Bayes Theorem, J. R. Stat. Soc. Ser. B. Stat. Methodol. Series B (Methodological). (1958), 102–107.
[15] P. A. Lewis, E. McKenzie, Minification processes and their transformations, J. Appl. Proba. **28**(1) (1991), 45-57.
[16] D. Mališić, On exponential autoregressive time series models, Math. Statis. Prob. Theory B (1987), 147-153.
[17] A. W. Marshall, I. Olkin, A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families, Biometrika. **84**(3) (1997), 641-652.
[18] A. B. Mello, M. C. Lima, A. D. Nascimento, A notable Gamma-Lindley first-order autoregressive process: An application to hydrological data, Environmetrics. **33**(4) (2022), e2724.
[19] H. Messaadia, H. Zeghdoudi, Around gamma Lindley distribution, J. Mod. Appl. Stat. Methods. **16**(2) (2017), 23.
[20] S. Nedjar, H. Zeghdoudi, On gamma Lindley distribution: Properties and simulations, J. Comput. Appl. Math. 298 (2016) 167-174.
[21] S. Nedjar, H. Zeghdoudi, New Compound Poisson distribution: Properties, Inflated distribution and applications, Int. J. Agricult. Stat. **16**(2) (2020).
[22] M. H. Neumann, E. Paparoditis, Goodness-of-fit tests for Markovian time series models: Central limit theory and bootstrap approximations, Bernoulli. **14**(1) (2008), 14-46.
[23] C.S. Rajitha, A. Akhilnath, Generalization of the lindley distribution with application to covid-19 data, Int. J. Data. Sci. Anal. (2022), 1–21.
[24] H. Zeghdoudi, S. Nedjar, Gamma Lindley distribution and its application, J. Appl. Probab. Stat. **11**(1) (2016), 129–138.