Computational Journal of Mathematical and Statistical Sciences
x(x), -
DOI:10.21608/cjmss.2025.391763.1200
https://cjmss.journals.ekb.eg/

Autoregressive Moving-Average Time Series Model with Errors Following a Log-Logistic Distribution

B.Reihaninia[1], **F.Hormozinejad**[1,*], **M.Khodamoradi**[2], **M.R.Ghalani**[1]

[1] Department of statistics, Ahv.C., Islamic Azad University, Ahvaz 61349, Iran; behzad.reihani@iau.ac.ir, m.ghalani@iau.ac.ir.

[2] Department of statistics, Iz.C., Islamic Azad University, Izeh 64171, Iran; Mohammad.khodamoradi@iau.ac.ir.

* **Correspondence:** Farshin.Hormozinejad@iau.ac.ir

Abstract: In the class of time series' models, the errors of the fitted models may follow a Log-Logistic distribution instead of the Normal distribution. This new model is called the Log-Logistic Autoregressive Moving-Average(L-LARMA) model. Following the introduction of this model's structure, its parameters have been estimated using numerical methods and the conditional maximum likelihood function. The structure for hypothesis testing, the information matrix, and the forecast function have also been developed and described. One application of this model is the forecasting and modeling of the overall stock market index. The overall index is one of the most critical indicators that economists and researchers in the field of economics closely monitor. The model and its forecast function have been fitted and determined using simulation and actual data of Iran stock market index from 2008 to 2022.

Keywords: Log-Logistic Distribution, Autoregressive Moving-Average Model, Conditional Maximum Likelihood Function, Information Matrix, Forecast Function, Overall Stock Market Index
Mathematics Subject Classification: 62M10, 62P20.
Received: 7 June 2025; Revised: 20 August 2025; Accepted: 30 August 2025; Online: 9 September 2025.

1. Introduction

The Log Logistic probability distribution function belongs to the family of probability functions with non-negative variable values, which has numerous applications, especially in survival analysis and parametric models.

The probability density function of the Log-Logistic distribution is shown below. [15]

$$f(y) = \frac{(\frac{b}{a})(\frac{y}{a})^{b-1}}{\left(1 + (\frac{y}{a})^{b}\right)^{2}} \quad y \geq 0, \quad a > 0, \quad b > 0, \tag{1.1}$$

Where a and b are non-negative parameters. The mean, median, and variance of Y are, respectively, as follows:

$$E(Y) = \frac{ab/\pi}{\sin(b/\pi)} \tag{1.2}$$

$$median(Y) = a \tag{1.3}$$

$$Var(Y) = a^2 \left(\frac{2b/\pi}{\sin(2b/\pi)} - \left(\frac{b/\pi}{\sin(b/\pi)} \right)^2 \right) \tag{1.4}$$

A time series model with a dependent variable following a Log-Logistic distribution leads to a new model within the time series class, which is briefly referred to as Log Logistic Auto-Regressive Moving Average (L-LARMA), similar to the model previously developed by Ferrari, S.L.P, and Cribari-Neto, for the Beta distribution [7], [13]. Prior to this paper, many others paved the way for how we have walked and came to the idea. To name a few, we can mention the papers [3], [8], and [4]. Regarding this type of time series, one can refer to the framework definition for non-Gaussian time series models. Generalized dynamic time series models with an exponential family distribution have been fitted [10], [8].

The Kumaraswamy Autoregressive Moving Average (KARMA) models within the class of truncated time series and possessing a Kumaraswamy distribution (a generalized Beta distribution), have also been fitted [2]. The Gamma autoregressive model has been introduced with a specific type of stationary Gamma process [16]. The modified quasi-Newton mehod was propoed for solving the nonlinear equation F(x)=0, which is based on a new quasi-Newton aooroach [18]. A natural generalization of the ARCH (Autoregressive Conditional Heteroskedastic) prossess to allow for past conditional variances was proposed [5].

Discussed a quasi-likelihood (QL) approach to regression analysis with time series data [19]. Furthermore, a closed-form for the moments of the non-central Gamma distribution with long-term memory retention properties has been obtained [9].

The framework for our discussion is the LLARMA model described in Section 2. Estimation of model parameters is discussed using the conditional likelihood method and building confidence intervals and hypothesis testing are considered and deals with the description of the diagnostic analysis and forecasting in Section 3. The proposed model is evaluated using the simulation and real data of the stock index in Section 4. Conclusions for the introduced model are presented, as well as a simulation and case study of stock market data in Section 5.

2. Autoregressive Moving-Average Model with Log-Logistic Distribution (L-LARMA)

2.1. The model

Assuming that $\{y_t\}$ for $t = 1, 2, ..., n$ is a non-negative stochastic process $y_t \geq 0$ with probability 1 and the set $\mathcal{F}_{t-1} = \sigma\{y_{t-1}, y_{t-2}, \ldots\}$, be a sigma-field that is $y_1, y_2, \ldots, y_{t-1}$ are measurable and the

conditional distribution of each y_t given the prior information set $\mathcal{F}_{t-1} = \sigma\{y_{t-1}, y_{t-2}, \dots\}$ follows a Log-Logistic distribution with parameters a and b.
According to the definition of the sigma field and distribution of the process $\{y_t\}$ and introducing new parameters into the conditional distribution as $\mathbf{Y}_t \mid \mathcal{F}_{t-1} \sim$ Log-Logistic$(\tilde{\mu}_t, \Phi)$ it has a Log-Logistic distribution with parameters $\tilde{\mu}_t = a$ and $\Phi = b$.
The conditional density function is defined as follows:

$$f_{\tilde{\mu}_t}(\mathbf{y}_t \mid \mathcal{F}_{t-1}) = \frac{\frac{\Phi}{\tilde{\mu}_t}(\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi-1}}{\left(1 + (\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}\right)^2} \qquad \tilde{\mu}_t \geq 0, \quad \Phi \geq 0. \tag{2.1}$$

The function $\mathbf{g}(\cdot)$ is defined as a link function that maps from the interval (0,1) to $\mathbb{R}$ ($\mathbf{g} : (0, 1) \to \mathbb{R}$ or $\mathbf{g}^{-1} : \mathbb{R} \to (0, 1)$), with functions such as logit or probit and log-log. In general, the Log-Logistic Autoregressive Moving-Average (L-LARMA) model is defined as below:

$$\eta_t = \mathbf{g}(\tilde{\mu}_t) = \mathbf{x}_t'\beta + \tau_t \tag{2.2}$$

where $\beta = (\beta_1, \beta_2 \dots, \beta_k)'$ is a set of unknown linear parameters and τ_t is the autoregressive moving-average component of the model.
In general, an ARMA process of order p and q is represented by τ_t

$$\tau_t = \alpha + \sum_{i=1}^{p} \phi_i \xi_{t-i} + \sum_{j=1}^{q} \theta_j r_{t-j} + r_t \tag{2.3}$$

where ϕ_i and θ_j are the unknown Autoregressive and moving-average parameters, respectively $\alpha \in \mathbb{R}$ is constant value, and r_t is the random error component.
The ARIMA models can be used to produce forecasts for time series data and have three parts. Not all parts are always necessary, but it depends on the type of time series data at hand. The three parts are the autoregressive (AR), the integrated (I) and lastly, the moving average (MA). Assumption for the AR part of a time series data is that the observed value depends on some linear combinations of previous observed values up to some maximum lags, plus an error term. Assumption for the MA part of time series data is that the observed value is a random error term plus some linear combinations of previous random error terms up to some maximum lags [6]. Using the σ-algebra $\mathcal{F}_{t-1}$ and not defining r_t in it, then $E(r_t \mid \mathcal{F}_{t-1}) = 0$ holds, and based on this conditional expectation, the ARMA part of the model in (2.3) is introduced as (2.4) :

$$\tau_t = \alpha + \sum_{i=1}^{p} \phi_i \xi_{t-i} + \sum_{j=1}^{q} \theta_j r_{t-j} \tag{2.4}$$

where ξ_{t-i} for $i > 0$ is measurable in $\mathcal{F}_{t-1}$ and $E(\xi_t \mid \mathcal{F}_{t-1}) \approx \tau_t$.
Based on the relation (2.2), $\xi_{t-i} = \mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}_{t-i}'\beta$, and by subsittuting in (2.4), the result is:

$$\tau_t = \alpha + \sum_{i=1}^{p} \phi_i(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}_{t-i}'\beta) + \sum_{j=1}^{q} \theta_j r_{t-j} \tag{2.5}$$

Considering (2.2) and (2.5), the Log-Logistic Autoregressive Moving-Average (L-LARMA(p, q)) model is introduced by (2.1) and (2.6)

$$\eta_t = \mathbf{g}(\tilde{\mu}_t) = \alpha + \mathbf{x}_t'\beta + \sum_{i=1}^{p} \phi_i(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}_{t-i}'\beta) + \sum_{j=1}^{q} \theta_j r_{t-j} \tag{2.6}$$

2.2. *The Consistency and Error*

In this model, the error is scaled based on the moving-average part and in the sigma field $\mathcal{F}_t$, Two methods are introduced for determining model errors: the first method uses the original values of the variable $r_t = \mathbf{y}_t - \mu_t$ and the second method uses the forecast function, i.e., $r_t = \mathbf{g}(\mathbf{y}_t) - \mathbf{g}(\mu_t)$

In the first case, for errors $E(\mathbf{y}_t - \mu_t \mid \mathcal{F}_{t-1}) = 0$ and $Var(\mathbf{y}_t - \mu_t) = \delta\tilde{\mu}_t^{\,2}$, and in the special case $E(\mathbf{y}_t - \mu_t) = 0$ results, and also for every $i < j$, $E(\mathbf{y}_i - \mu_i)(\mathbf{y}_j - \mu_j) = E(\mathbf{y}_i - \mu_i)E((\mathbf{y}_j - \mu_j) \mid \mathcal{F}_{t-1})) = 0$. It is concluded that the errors are orthogonal.

2.3. *The Sensitivity to Outliers*

In statistical analysis as well as in time series, the issue of outlier data is of particular importance. Although the outlier data is a sign of the reality of the data and the event, unfortunately, in statistical analysis, they cause errors and biased estimates of parameters. Various ways to make this data less effective or ineffective have been proposed, including deleting or replacing it with other data. But another way to solve this problem is to use the median index instead of the mean. Using the median has the advantage that even if there are outlier data, it as does not affectthe estimation of model parameters.

3. Estimation of L-LARMA Model Parameters

The model parameters are estimated based on the conditional maximum likelihood method. Considering the model structure, the parameter vector of the model is introduced as $\gamma = (\alpha, \beta', \phi', \theta', \Phi)'$ where $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$, $\phi = (\phi_1, \phi_2, \ldots, \phi_p)'$ are the parameter vectors of the autoregressive part of the model, and $\theta = (\theta_1, \theta_2, \ldots, \theta_q)'$ is the parameter vector of the moving-average part of the model.

Note: The use of the conditional maximum likelihood method for the first m observations must be defined as $m = \max\{p, q\}$, and for the first m observations, the errors are assumed to be nearly zero. Also, in the autoregressive and moving-average parts, since $m = \max\{p, q\}$ and $q \leq m$, the first q errors are equal to zero.

3.1. *Estimation of Parameter Vector*

The logarithm of the likelihood function is defined as follows:

$$\begin{aligned} \log(f_{\tilde{\mu}_t}(\mathbf{y}_t \mid \mathcal{F}_t)) &= l_t(\tilde{\mu}_t, \Phi) \\ &= \log\Phi - \log\tilde{\mu}_t + (\Phi - 1)(\log\mathbf{y}_t - \log\tilde{\mu}) - 2\log\left(1 + (\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}\right) \end{aligned} \tag{3.1}$$

Where $l = \sum_{t=m+1}^{n} l_t(\tilde{\mu}_t, \Phi)$ is the conditional likelihood function.
The parameters vector is of the form $\Lambda = (\alpha, \beta_l, \phi_i, \theta_j, \Phi)$, and the derivative concerning each of the parameters $\lambda_k \in \Lambda$ is shown as:

$$\frac{\partial l}{\partial \lambda_k} = \sum_{t=m+1}^{n} \frac{\partial l_t(\tilde{\mu}_t, \phi)}{\partial \tilde{\mu}_t} \frac{\partial \tilde{\mu}_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \lambda_k} \tag{3.2}$$

Given that $\eta_t = \mathbf{g}(\mu_t)$, it can be shown that:

$$\frac{\partial \mu_t}{\partial \eta_t} = 1/\frac{\partial \eta_t}{\partial \mu_t} = \frac{1}{\mathbf{g}'(\mu_t)} \tag{3.3}$$

Theorem 1. *The partial derivative of the likelihood function is equal to* $\frac{\partial l_t(\tilde{\mu}_t,\Phi)}{\partial \tilde{\mu}_t} = \frac{\Phi}{\tilde{\mu}_t}\frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t{}^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t{}^{\Phi}}$

Proof. Using (3.1):

$$\log(f(\mathbf{y}_t \mid \mathcal{F}_{t-1})) = \log\Phi - \log\tilde{\mu}_t + (\Phi-1)(\log\mathbf{y}_t - \log\tilde{\mu}_t) - 2\log(1+\left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi}).$$

$$\begin{aligned}\frac{\partial l_t(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t} &= -\frac{1}{\tilde{\mu}_t} - \frac{\Phi}{\tilde{\mu}_t} + \frac{1}{\tilde{\mu}_t} + \frac{2\Phi(\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}}{\tilde{\mu}_t(1+\left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi})}\\ &= \frac{-\Phi\left(\left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi}+1\right)+2\Phi\left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi}}{\tilde{\mu}_t(1+\left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi})} = \frac{\Phi(\left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi}-1)}{\tilde{\mu}_t(1+\left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi})}\\ &= \frac{\Phi(\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi})}{\tilde{\mu}_t(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})} = \frac{\Phi}{\tilde{\mu}_t}\frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}.\end{aligned}$$

□

Considering (3.2), (3.3), and using Theorem1, we have:

$$\frac{\partial l}{\partial\lambda} = \sum_{t=m+1}^{n}\frac{\Phi}{\tilde{\mu}_t}\frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t{}^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t{}^{\Phi}}\frac{1}{\mathbf{g}_t'(\mu_t)}\frac{\partial\eta_t}{\partial\lambda_k}. \tag{3.4}$$

We define simple original scale error using $r_t = y_t - \mu_t$, so that we can put r_t in (2.6) for $\frac{\partial\eta_t}{\partial\lambda_k}$ leads to (3.4).

Since the values are non-negative in the log-Logistic distribution, the absolute values of the residuals are used for parameter estimation. This change in the least squares error estimation does not cause any issues because in the error squared, there is no difference in the sign of the residuals, and the goal is to determine the least squares error. Accordingly, in this article, the fitted models use the absolute values of the residuals

Theorem 2. *In the Log-Logistic ARMA process, the estimate for each of the parameters are as follows:*

$$\frac{\partial l}{\partial\alpha} = U(\alpha) = \mathbf{1}'TC, \qquad \frac{\partial l}{\partial\beta_l} = U(\beta) = \mathbf{M}'TC, \qquad \frac{\partial l}{\partial\phi_i} = U(\phi) = \mathbf{N}'TC,$$

$$\frac{\partial l}{\partial\theta_j} = U(\theta) = \mathbf{R}'TC, \qquad \frac{\partial l}{\partial\Phi} = U(\Phi) = \sum_{t=m+1}^{n}\frac{1}{\Phi} - \frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t{}^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t{}^{\Phi}}\log(\frac{\mathbf{y}_t}{\tilde{\mu}_t}).$$

where

$$\mathbf{g}(\tilde{\mu}_t) = a + \mathbf{x}_t'\beta + \sum_{i=1}^{p}\phi_i(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}_{t-i}'\beta) + \sum_{j=1}^{q}\theta_j r_{t-j}$$

$$c_t = \frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}, \qquad C = (\Phi c_{m+1}, \Phi c_{m+2}, \ldots, \Phi c_n)'_{(n-m)\times 1}$$

$$T = diag\left\{ \frac{1}{\mathbf{g}'(\mu_{m+1})}, \frac{1}{\mathbf{g}'(\mu_{m+2})}, \ldots, \frac{1}{\mathbf{g}'(\mu_n)} \right\}_{(n-m)\times(n-m)}$$

$$\mathbf{M} = \left\{ \mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l} \right\}_{(n-m)\times(k)} \qquad \mathbf{N} = \{\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}'_{t-i}\beta\}_{(n-m)\times(p)}$$

$$\mathbf{R} = \left\{ r_{t-j} \right\} \textit{is matrix with dimension } (n-m)\times q$$

Proof.

$$\eta_t = \mathbf{g}(\tilde{\mu}_t) = a + \mathbf{x}'_t\beta + \sum_{i=1}^{p} \phi_i(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}'_{t-i}\beta) + \sum_{j=1}^{q} \theta_j r_{t-j},$$

$$\begin{aligned} \log(f_{\tilde{\mu}_t}(\mathbf{y}_t \mid \mathcal{F}_t)) &= l_t(\tilde{\mu}_t, \Phi) \\ &= \log\Phi - \log\tilde{\mu}_t + (\Phi-1)(\log\mathbf{y}_t - \log\tilde{\mu}_t) - 2\log\left(1 + (\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}\right), \end{aligned}$$

$$l = \sum_{t=m+1}^{n} l_t(\tilde{\mu}_t, \Phi),$$

$$\frac{\partial l}{\partial \lambda_k} = \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}^{\Phi_t} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} \frac{\partial \eta_t}{\partial \lambda_k}, \qquad \lambda_k \in \{\alpha, \beta_l, \Phi_i, \theta_j, \Phi\}.$$

For parameter a :

$$\frac{\partial \eta_t}{\partial a} = 1 + \sum_{j=1}^{q} \theta_j \frac{\partial}{\partial a}(\mathbf{y}_{t-j} - \tilde{\mu}_{t-j}),$$

$$\frac{\partial l}{\partial a} = \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} \frac{\partial \eta_t}{\partial a},$$

$$c_t = \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}}, \qquad C = (\Phi c_{m+1}, \Phi c_{m+2}, \ldots, \Phi c_n)'_{(n-m)\times 1},$$

$$T = \text{diag}\left\{ \frac{1}{\mathbf{g}'(\mu_{m+1})}, \frac{1}{\mathbf{g}'(\mu_{m+2})}, \ldots, \frac{1}{\mathbf{g}'(\mu_n)} \right\}_{(n-m)\times(n-m)},$$

$$U(a) = \frac{\partial l}{\partial a} = \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_\mathbf{t}} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} \frac{\partial \eta_t}{\partial a} = \mathbf{1}'TC.$$

For parameter β_l :

$$\frac{\partial \eta_t}{\partial \beta_l} = \mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l},$$

$$U(\beta) = \frac{\partial l}{\partial \beta_l} = \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_l}$$

$$= \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} (\mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l}) = \mathbf{M}'TC,$$

$$M = \left\{ \mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l} \right\} \text{is matrix with dimension } (n-m) \times k.$$

For parameter ϕ_i :

$$\frac{\partial \eta_t}{\partial \phi_i} = \mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}'_{t-i}\beta,$$

$$U(\phi) = \frac{\partial l}{\partial \phi_i} = \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} \frac{\partial \eta_t}{\partial \phi_i}$$

$$= \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} (\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}'_{t-i}\beta) = \mathbf{N}'TC,$$

$$\mathbf{N} = \{\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}'_{t-i}\beta\} \text{ is a matrix with dimension } (n-m) \times p.$$

For parameter θ_j :

$$\frac{\partial \eta_t}{\partial \theta_j} = r_{t-j},$$

$$U(\theta) = \frac{\partial l}{\partial \phi_i} = \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} \frac{\partial \eta_t}{\partial \theta_j}$$

$$= \sum_{t=m+1}^{n} \frac{\Phi}{\tilde{\mu}_t} \frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \frac{1}{\mathbf{g}'(\mu_t)} (r_{t-j}) = \mathbf{R}'TC,$$

$$\mathbf{R} = \left\{ r_{t-j} \right\} \text{ is matrix with dimension } (n-m) \times q.$$

For parameter $\mathbf{\Phi}$:

$$U(\Phi) = \frac{\partial l}{\partial \Phi} = \sum_{t=m+1}^{n} \left(\frac{1}{\Phi} + (\log \mathbf{y}_t - \log \tilde{\mu}_t) - \left((2 \log \frac{\mathbf{y_t}}{\tilde{\mu}_t}) - \frac{2 \log \frac{\mathbf{y}_t}{\tilde{\mu}_t}}{1 + \left(\frac{\mathbf{y}_t}{\tilde{\mu}_t}\right)^{\Phi}} \right) \right)$$

$$= \sum_{t=m+1}^{n} (\frac{1}{\Phi} + \frac{-(\log \mathbf{y}_t - \log \tilde{\mu}_t)(1 + (\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}) - \log \mathbf{y}_t + \log \tilde{\mu}_t + 2 \log \mathbf{y}_t - 2 \log \tilde{\mu}_t}{1 + (\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}})$$

$$= \sum_{t=m+1}^{n} \left(\frac{1}{\Phi} + (\log \mathbf{y}_t - \log \tilde{\mu}_t) \left(\frac{1 - (\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}}{1 + (\frac{\mathbf{y}_t}{\tilde{\mu}_t})^{\Phi}} \right) \right)$$

$$= \sum_{t=m+1}^{n} (\frac{1}{\Phi} + (\log \mathbf{y}_t - \log \tilde{\mu}_t)(\frac{\tilde{\mu}_t^{\Phi} - \mathbf{y}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}})) = \sum_{t=m+1}^{n} \left(\frac{1}{\mathbf{\Phi}} - \left(\frac{\mathbf{y}_t^{\Phi} - \tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi} + \tilde{\mu}_t^{\Phi}} \right) \log(\frac{\mathbf{y}_t}{\tilde{\mu}_t}) \right).$$

□

The matrix form of the parameter vector is $U(\gamma) = (U(\alpha), U'(\beta), U'(\phi), U'(\theta), U(\Phi))'$. The parameters are estimated using the conditional maximum likelihood method and solving the equation $U(\gamma) = 0$, which can be solved using numerical methods such as Newton's method and the EM algorithm.
Note that $U(\gamma) = 0$ is a matrix equation, a column vector with dimension $k + p + q + 2$. Hence, conditional maximum likelihood estimates should be obtained by using numerical methods such as Newton or Quasi-Newton nonlinear optimization or the EM algorithm [18].

3.2. Fisher information matrix

In this section, the Fisher information matrix is calculated. According to Theorem 3, the condition of the zero derivative of the Fisher information function is established. Therefore, to calculate the Fisher information matrix, the second derivative of each of the model parameters defined in (3.4) is used and is shown as follows:

$$\begin{aligned}\frac{\partial^2 l_t(\tilde{\mu}_t,\Phi)}{\partial\lambda_i\partial\lambda_j} &= \sum_{t=m+1}^{n} \frac{\partial}{\tilde{\mu}_t}\left(\frac{\partial l_t(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t}\frac{\partial\tilde{\mu}_t}{\partial\eta_t}\frac{\partial\eta_t}{\partial\lambda_i}\right)\frac{\partial\tilde{\mu}_t}{\partial\eta_t}\frac{\partial\eta_t}{\partial\lambda_j} = \\ &= \sum_{t=m+1}^{n}\left(\frac{\partial^2 l_t(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t^2}\frac{\partial\tilde{\mu}_t}{\partial\eta_t}\frac{\partial\eta_t}{\partial\lambda_j} + \frac{\partial l_t(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t}\left(\frac{\partial}{\tilde{\mu}_t}\left(\frac{\partial\tilde{\mu}_t}{\partial\eta_t}\frac{\partial\eta_t}{\partial\lambda_j}\right)\right)\right)\frac{\partial\tilde{\mu}_t}{\partial\eta_t}\frac{\partial\eta_t}{\partial\lambda_j}.\end{aligned} \tag{3.5}$$

In general case, the expected value of (3.5) can be shown as:

$$\sum_{t=m+1}^{n} E\left(\frac{\partial^2 l_t(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t^2} \mid \mathcal{F}_{t-1}\right)\left(\frac{\partial\tilde{\mu}_t}{\partial\eta_t}\right)^2\frac{\partial\eta_t}{\partial\lambda_i}\frac{\partial\eta_t}{\partial\lambda_j}. \tag{3.6}$$

Theorem 3. *The condition that the expectation is zero for the first derivative of the Fisher information function also holds for the probability density.*

$$E\left(\frac{\partial f(y_t,\Phi,\tilde{\mu}_t)}{\partial\tilde{\mu}_t}\right) = 0.$$

Proof.

$$\frac{\partial f(\mathbf{y}_t,\Phi,\tilde{\mu}_t)}{\partial\tilde{\mu}_t} = \frac{\Phi}{\tilde{\mu}_t}\left(\frac{\mathbf{y}_t^\Phi - \tilde{\mu}_t^\Phi}{\mathbf{y}_t^\Phi + \tilde{\mu}_t^\Phi}\right) = C_t.$$

let $W = \frac{\mathbf{y}_t^\Phi - \tilde{\mu}_t^\Phi}{\mathbf{y}_t^\Phi + \tilde{\mu}_t^\Phi}$ then $F_W(w) = P(W \le w) = P(\frac{\mathbf{y}_t^\Phi - \tilde{\mu}_t^\Phi}{\mathbf{y}_t^\Phi + \tilde{\mu}_t^\Phi} \le w)$,

$$\begin{aligned}P\left(\mathbf{y}_t^\Phi - \tilde{\mu}_t^\Phi \le w\mathbf{y}_t^\Phi + w\tilde{\mu}_t^\Phi\right) &= P\left((1-w)\,\mathbf{y}_t^\Phi \le (1+w)\,\tilde{\mu}_t^\Phi\right) = P\left(\mathbf{y}_t^\Phi \le \frac{1+w}{1-w}\tilde{\mu}_t^\Phi\right) \\ &= P(\mathbf{y}_t \le \sqrt[\Phi]{\frac{1+w}{1-w}}|\tilde{\mu}_t|)\end{aligned}$$

$f_W(w)$ is probability density function of W.

$$f_W(w) = \frac{2\tilde{\mu}_t\left(\frac{1+w}{1-w}\right)^{\frac{1}{\Phi}}}{\Phi(1-w^2)} \frac{\frac{\Phi}{\tilde{\mu}_t}\left(\frac{\tilde{\mu}_t\left(\frac{1+w}{1-w}\right)^{\frac{1}{\Phi}}}{\tilde{\mu}_t}\right)^{\Phi-1}}{(1+\left(\frac{\tilde{\mu}_t\left(\frac{1+w}{1-w}\right)^{\frac{1}{\Phi}}}{\tilde{\mu}_t}\right)^{\Phi})^2} = \frac{2\left(\frac{1+w}{1-w}\right)^{\frac{1}{\Phi}}\left(\frac{1+w}{1-w}\right)^{1-\frac{1}{\Phi}}}{(1-w^2)(1+\frac{1+w}{1-w})^2}$$

$$= \frac{2\frac{1+w}{1-w}}{(1-w^2)\left(\frac{2}{1-w}\right)^2} = \frac{1}{2}.$$

$f_W(w) = \frac{1}{2}$ for $-1 < w < 1$ so that $E(W) = 0$ or $E\left(\frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}\right) = 0$ or $E(C_t) = 0$

$$E\left(\frac{\partial f(y_t,\Phi,\tilde{\mu}_t)}{\partial\tilde{\mu}_t}\right) = 0.$$

□

Using Theorem 3, relation (3.6) can be expressed as follows:(refer to Theorem 4)

$$E\left(\frac{\partial^2 l_t(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t^2} \mid \mathcal{F}_{t-1}\right) = \frac{\Phi^2}{2\tilde{\mu}_t^{2\Phi}} - \frac{3\tilde{\mu}_t^{\Phi}-2}{\tilde{\mu}_t^{\Phi/2}} = w_t \tag{3.7}$$

Theorem 4. *In the Log-Logistic Autoregressive Moving-Average model, the relation (3.7) is established.*

Proof.

$$E\left(\frac{\partial^2 l_t(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t^2} \mid \mathcal{F}_{t-1}\right) = E\left(\frac{\Phi^2\tilde{\mu}_t^{\Phi-2}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}\right) - E\left(\frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}\right) + E\left(\frac{\Phi^2\tilde{\mu}_t^{\Phi-2}(\mathbf{y}^{\Phi}{}_t-\tilde{\mu}_t^{\Phi})}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2}\right)$$

Let $W = \frac{1}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}$. So that, $P(W \le w) = P(\frac{1}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}} \le w)$ and $P(\mathbf{y}_t^{\Phi} \ge \frac{1-w\tilde{\mu}_t^{\Phi}}{w}) = 1 - P(\mathbf{y}_t \le \sqrt[\Phi]{\frac{1-w\tilde{\mu}_t^{\Phi}}{w}})$

$$f_W(w) = \frac{-\left(\frac{w\tilde{\mu}_t^{\Phi}-1}{w}\right)^{\frac{1}{\Phi}-1}}{\Phi w^2} \frac{\frac{\Phi}{\tilde{\mu}_t}\left(\frac{\left(\frac{1-w\mu_t^{\Phi}}{w}\right)^{\frac{1}{\Phi}}}{\tilde{\mu}_t}\right)^{\Phi-1}}{(1+\frac{\left(\frac{1-w\tilde{\mu}_t^{\Phi}}{w}\right)^{\frac{1}{\Phi}}}{\tilde{\mu}_t})^2} = -\tilde{\mu}_t{}^{\Phi}$$

$$f_W(w) = \tilde{\mu}_t^{\Phi}, \quad 0 \le w \le \frac{1}{\tilde{\mu}_t^{\Phi}}, \quad E(W) = \frac{1}{2\tilde{\mu}_t{}^{2\Phi}}$$

To prove that $E\left(\frac{\Phi^2\tilde{\mu}_t^{\Phi-2}(\mathbf{y}^{\Phi}{}_t-\tilde{\mu}_t^{\Phi})}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2}\right) = \frac{3\tilde{\mu}_t^{\Phi}-2}{3\tilde{\mu}_t^{\frac{\Phi}{2}}}$, let $W = \frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi}}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2} = 1 - \frac{2\tilde{\mu}_t^{\Phi}}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2}$, so

$$P(W \le w) = P\left(1 - \frac{2\tilde{\mu}_t^{\Phi}}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2} \le w\right) = 1 - P\left(\frac{2\tilde{\mu}_t^{\Phi}}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2} \le 1-w\right)$$

$$P\left(\frac{2\tilde{\mu}_t^{\Phi}}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2}\leq 1-w\right)=P\left(\frac{\sqrt{2}\tilde{\mu}_t^{\frac{\Phi}{2}}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}\leq\sqrt{1-w}\right)$$
$$=P\left(\sqrt{2}\tilde{\mu}_t^{\frac{\Phi}{2}}\leq\mathbf{y}_t^{\Phi}\sqrt{1-w}+\tilde{\mu}_t^{\Phi}\sqrt{1-w}\right)$$
$$=P\left(\mathbf{y}_t^{\Phi}\leq\frac{\sqrt{2}\tilde{\mu}_t^{\frac{\Phi}{2}}}{\sqrt{1-w}}-\tilde{\mu}_t^{\Phi}\right)=P\left(\mathbf{y}_t\leq\left(\frac{\sqrt{2}\tilde{\mu}_t^{\frac{\Phi}{2}}}{\sqrt{1-w}}-\tilde{\mu}_t^{\Phi}\right)^{\frac{1}{\Phi}}\right).$$

$$f_W(w)=\frac{\tilde{\mu}_t^{\frac{\Phi}{2}}}{2\sqrt{2(1-w)}},\quad 1-\frac{2}{\tilde{\mu}_t^{\Phi}}\leq w\leq 1,\quad E\,(W)=2\sqrt{\frac{2}{\tilde{\mu}_t^{\Phi}}}-\frac{2}{3}\sqrt{(\frac{2}{\tilde{\mu}_t^{\Phi}})^3}$$
$$E\left(\frac{\Phi^2\tilde{\mu}_t^{\Phi-2}(\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi})}{(\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi})^2}\right)=E\left(\Phi^2\tilde{\mu}_t^{\Phi-2}W\right)=\frac{3\tilde{\mu}_t^{\Phi}-2}{3\tilde{\mu}_t^{\frac{\Phi}{2}}}$$
$$E\left(\frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}\right)=0$$
$$E\left(\frac{\partial^2 l_t\,(\tilde{\mu}_t,\Phi)}{\partial\tilde{\mu}_t^2}\mid\mathcal{F}_{t-1}\right)=\frac{\Phi^2}{2\tilde{\mu}_t^{2\Phi}}-\frac{3\tilde{\mu}_t^{\Phi}-2}{\tilde{\mu}_t^{\Phi/2}}=w_t.$$

□

Based on (3.6) and (3.7), the conditional Fisher information matrix is defined as follows:

$$K(\gamma)=E\left(\frac{\partial^2 l_t(\tilde{\mu}_t,\Phi)}{\partial\lambda_i\partial\lambda_j}\mid\mathcal{F}_{t-1}\right)=-\sum_{t=m+1}^{n}\left(\frac{\Phi^2}{2\tilde{\mu}_t^2}-\frac{3\tilde{\mu}_t^{\Phi}-2}{\tilde{\mu}_t^{\Phi/2}}\right)\frac{1}{\mathbf{g}'(\tilde{\mu}_t)}\frac{\partial\eta_t}{\partial\lambda_i}\frac{\partial\eta_t}{\partial\lambda_j}$$
$$=\sum_{t=m+1}^{n}w_t\frac{\partial\eta_t}{\partial\lambda_i}\frac{\partial\eta_t}{\partial\lambda_j}. \tag{3.8}$$

Theorem 5. *In the L-LARMA model, the components of the conditional Fisher's information matrix are calculated as follows:*

$$\mathbf{K}_{\alpha\alpha}=-tr(\mathbf{W}),\quad \mathbf{K}_{\alpha\beta}=\mathbf{K}'_{\beta\alpha}=-\mathbf{M}'\mathbf{W}\mathbf{1},\quad \mathbf{K}_{\alpha\phi}=\mathbf{K}'_{\phi\alpha}=-\mathbf{N}'\mathbf{W}\mathbf{1}$$
$$\mathbf{K}_{\alpha\theta}=\mathbf{K}'_{\theta\alpha}=-\mathbf{R}'\mathbf{W}\mathbf{1},\quad \mathbf{K}_{\alpha\Phi}=\mathbf{K}'_{\Phi\alpha}=-tr(\mathbf{D}\mathbf{W}),\quad \mathbf{K}_{\beta\beta}=-\mathbf{M}\mathbf{W}\mathbf{M}$$
$$\mathbf{K}_{\beta\phi}=\mathbf{K}'_{\phi\beta}=-\mathbf{M}'\mathbf{W}\mathbf{N},\quad \mathbf{K}_{\beta\theta}=\mathbf{K}'_{\theta\beta}=-\mathbf{M}'\mathbf{W}\mathbf{R},\quad \mathbf{K}_{\beta\Phi}=\mathbf{K}'_{\Phi\beta}=-\mathbf{M}'\mathbf{D}\mathbf{W}\mathbf{1}$$
$$\mathbf{K}_{\phi\phi}=-\mathbf{N}'\mathbf{W}\mathbf{N},\quad \mathbf{K}_{\phi\theta}=\mathbf{K}'_{\theta\phi}=-\mathbf{N}'\mathbf{W}\mathbf{R},\quad \mathbf{K}_{\phi\Phi}=\mathbf{K}'_{\Phi\phi}=-\mathbf{N}'\mathbf{D}\mathbf{W}\mathbf{1}$$
$$\mathbf{K}_{\theta\theta}=-\mathbf{R}'\mathbf{W}\mathbf{R},\quad \mathbf{K}_{\theta\Phi}=\mathbf{K}'_{\Phi\theta}=-\mathbf{R}\mathbf{D}\mathbf{W}\mathbf{1},\quad \mathbf{K}_{\Phi\Phi}=-tr(\mathbf{D}^2\mathbf{W}).$$

Proof. Note that the matrices, $\mathbf{M},\mathbf{N}$, and $\mathbf{R}$ were defined in the Theorem 3, to define matrices $\mathbf{D},\mathbf{W}$ and proving this Theorem ,

$$\frac{\partial l}{\partial\lambda_k}=\sum_{t=m+1}^{n}\frac{\Phi}{\tilde{\mu}_t}\frac{\mathbf{y}_t^{\Phi}-\tilde{\mu}_t^{\Phi}}{\mathbf{y}_t^{\Phi}+\tilde{\mu}_t^{\Phi}}\frac{1}{\mathbf{g}'(\mu_t)}\frac{\partial\eta_t}{\partial\lambda_k}=\sum_{t=m+1}^{n}w_t\frac{\partial\eta_t}{\partial\lambda_k}$$

$$\eta_t = g(\tilde{\mu}_t) = a + \mathbf{x}_t'\beta + \sum_{i=1}^{p} \phi_i(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}_{t-i}'\beta) + \sum_{j=1}^{q} \theta_j r_{t-j}$$

$$\mathbf{K}_{\alpha\alpha} = E(\frac{\partial^2 l}{\partial \alpha^2}) = -\sum_{t=m+1}^{n} w_t \frac{\partial \eta_t}{\partial \alpha}\frac{\partial \eta_t}{\partial \alpha} = -\sum_{t=m+1}^{n} w_t = -tr(W)$$

$$\begin{aligned}
\mathbf{K}_{\alpha\beta} = \mathbf{K}_{\beta\alpha}' &= E(\frac{\partial^2 l}{\partial \alpha \partial \beta}) = -\sum_{t=m+1}^{n} w_t \frac{\partial \eta_t}{\partial \alpha}\frac{\partial \eta_t}{\partial \beta} \\
&= -\sum_{t=m+1}^{n} w_t(\mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l}) = -\mathbf{M}'W\mathbf{1}. \\
\mathbf{K}_{\alpha\phi} = \mathbf{K}_{\phi\alpha}' &= E(\frac{\partial^2 l}{\partial \alpha \partial \phi}) = -\sum_{t=m+1}^{n} w_t \frac{\partial \eta_t}{\partial \alpha}\frac{\partial \eta_t}{\partial \phi} = -\sum_{t=m+1}^{n} w_t(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}'_{t-i}\beta) = -\mathbf{N}'W\mathbf{1}. \\
\mathbf{K}_{\alpha\theta} = \mathbf{K}_{\theta\alpha}' &= E(\frac{\partial^2 l}{\partial \alpha \partial \phi}) = -\sum_{t=m+1}^{n} w_t r_{t-j} = -\mathbf{R}'W\mathbf{1}. \\
\mathbf{K}_{\alpha\Phi} = \mathbf{K}_{\Phi\alpha}' &= E(\frac{\partial^2 l}{\partial \alpha \partial \Phi}) = -\sum_{t=m+1}^{n} w_t \frac{-\Phi\mu_t^{\Phi-1}(\log \mu_t)y_t^{-\Phi} + \mu_t^{\Phi-1}(\Phi \log(y_t) - 1) + \frac{1}{\mu_t}}{2\mathbf{g}'(\mu_t)} \\
&= -\sum_{t=m+1}^{n} w_t d_t = -tr(DW).
\end{aligned}$$

$$\begin{aligned}
\mathbf{K}_{\beta\beta} &= E(\frac{\partial^2 l}{\partial \beta^2}) = -\sum_{t=m+1}^{n} w_t \frac{\partial \eta_t}{\partial \beta}\frac{\partial \eta_t}{\partial \beta} \\
&= -\sum_{t=m+1}^{n} w_t \left(\mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l}\right)\left(\mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l}\right) = -\mathbf{M}'W\mathbf{M}. \\
\mathbf{K}_{\beta\phi} = \mathbf{K}_{\phi\beta}' &= E(\frac{\partial^2 l}{\partial \beta \partial \phi}) = -\sum_{t=m+1}^{n} w_t \frac{\partial \eta_t}{\partial \beta}\frac{\partial \eta_t}{\partial \phi} \\
&= -\sum_{t=m+1}^{n} w_t \left(\mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l}\right)(\mathbf{g}(y_{t-i}) - \mathbf{x}'_{t-i}\beta) = -\mathbf{M}'W\mathbf{N}.
\end{aligned}$$

$$\begin{aligned}
\mathbf{K}_{\beta\theta} = \mathbf{K}_{\theta\beta}' &= E(\frac{\partial^2 l}{\partial \beta \partial \theta}) = -\sum_{t=m+1}^{n} w_t \frac{\partial \eta_t}{\partial \beta}\frac{\partial \eta_t}{\partial \theta} \\
&= -\sum_{t=m+1}^{n} w_t \left(\mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l}\right) r_{t-j} = -\mathbf{M}'W\mathbf{R}. \\
\mathbf{K}_{\beta\Phi} = \mathbf{K}_{\Phi\beta}' &= E(\frac{\partial^2 l}{\partial \beta \partial \Phi}) = -\sum_{t=m+1}^{n} w_t \frac{\partial \eta_t}{\partial \beta}\frac{\partial \eta_t}{\partial \Phi}
\end{aligned}$$

$$= -\sum_{t=m+1}^{n} w_t \left(\mathbf{x}_{tl} - \sum_{i=1}^{p} \phi_i \mathbf{x}_{(t-i)l} \right) d_t = -\mathbf{M}' D W \mathbf{1}.$$

$$\mathbf{K}_{\phi\phi} = E(\frac{\partial^2 l}{\partial\phi\partial\phi}) = -\sum_{t=m+1}^{n} w_t \frac{\partial\eta_t}{\partial\phi}\frac{\partial\eta_t}{\partial\phi}$$

$$= -\sum_{t=m+1}^{n} w_t(\mathbf{g}(y_{t-i}) - \mathbf{K}_{t-i}\beta)(\mathbf{g}(y_{t-i}) - \mathbf{x}'_{t-i}\beta) = -\mathbf{N}' W \mathbf{N}.$$

$$\mathbf{K}_{\phi\theta} = \mathbf{K}'_{\theta}\phi = E(\frac{\partial^2 l}{\partial\phi\partial\theta}) = -\sum_{t=m+1}^{n} w_t \frac{\partial\eta_t}{\partial\phi}\frac{\partial\eta_t}{\partial\theta} = -\sum_{t=m+1}^{n} w_t(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{K}'_{t-i}\beta) r_{t-j} = -\mathbf{N}' W \mathbf{R}.$$

$$\mathbf{K}_{\phi\Phi} = \mathbf{K}'_{\Phi\phi} = E(\frac{\partial^2 l}{\partial\phi\partial\Phi}) = -\sum_{t=m+1}^{n} w_t \frac{\partial\eta_t}{\partial\phi}\frac{\partial\eta_t}{\partial\Phi} = -\sum_{t=m+1}^{n} w_t(\mathbf{g}(\mathbf{y}_{t-i}) - \mathbf{x}'_{t-i}\beta) r_{t-j} d_t = -\mathbf{N}' D W 1.$$

$$\mathbf{K}_{\theta\theta} = E(\frac{\partial^2 l}{\partial\theta^2}) = -\sum_{t=m+1}^{n} w_t r_{t-j}^2 = -\mathbf{R}' W \mathbf{R}.$$

$$\mathbf{K}_{\theta\Phi} = \mathbf{K}'_{\Phi\theta} = E(\frac{\partial^2 l}{\partial\theta\partial\Phi}) = -\sum_{t=m+1}^{n} w_t d_t r_{t-j} = -\mathbf{R} D W \mathbf{1}.$$

$$\mathbf{K}_{\Phi\Phi} = E(\frac{\partial^2 l}{\partial\Phi^2}) = -\sum_{t=m+1}^{n} w_t d_t^2 = -tr(D^2 W).$$

□

All model parameters, it is calculated, and the conditional Fisher information matrix is formed as follows (see Theorem 5).

$$\mathbf{K}(\gamma) = \begin{bmatrix} \mathbf{K}_{\alpha\alpha} & \mathbf{K}'_{\beta\alpha} & \mathbf{K}'_{\phi\alpha} & \mathbf{K}'_{\theta\alpha} & \mathbf{K}'_{\Phi\alpha} \\ \mathbf{K}_{\alpha\beta} & \mathbf{K}_{\beta\beta} & \mathbf{K}'_{\phi\beta} & \mathbf{K}'_{\theta\beta} & \mathbf{K}_{\Phi\beta} \\ \mathbf{K}_{\alpha\phi} & \mathbf{K}_{\beta\phi} & \mathbf{K}_{\phi\phi} & \mathbf{K}'_{\theta\phi} & \mathbf{K}'_{\Phi\phi} \\ \mathbf{K}_{\alpha\theta} & \mathbf{K}_{\beta\theta} & \mathbf{K}_{\phi\theta} & \mathbf{K}_{\theta\theta} & \mathbf{K}'_{\Phi\theta} \\ \mathbf{K}_{\alpha\Phi} & \mathbf{K}_{\beta\Phi} & \mathbf{K}_{\phi\Phi} & \mathbf{K}_{\theta\Phi} & \mathbf{K}_{\Phi\Phi} \end{bmatrix}_{(k+p+q+2)\times(k+p+q+2)} \tag{3.9}$$

Based on the common conditions in the maximum likelihood method, for large samples, it follows a Normal distribution [2] therefore , we have:

$$\hat{\gamma} \sim \mathcal{N}_{k+p+q+2}(\gamma, \mathbf{K}^{-1}). \tag{3.10}$$

Where $\hat{\gamma} = (\hat{\alpha}, \hat{\beta}, \hat{\phi}, \hat{\theta}, \hat{\Phi})'$ is the vector of estimated model parameters by the conditional maximum likelihood estimation (CMLE) method, respectively for $\alpha, \beta, \phi, \theta, \Phi$.

3.3. The confidence interval and hypothesis testing

In this model, assuming that and based on what was said in the previous section if i-the element from the parameter vector γ and $\mathbf{K}^{ij}(\gamma)$ is equal to the element of row i and column j of the matrix $\mathbf{K}^{-1}$ we have:

$$\frac{\hat{\gamma}_i - \gamma_i}{\sqrt{K^{ij}(\hat{\gamma})}} \xrightarrow{D} \mathcal{N}(0, 1) \tag{3.11}$$

Moreover, for large samples, a confidence interval for γ_i is follows:

$$\hat{\gamma}_i - z_{1-\alpha/2}\sqrt{K^{ii}(\hat{\gamma})} < \gamma_i < \hat{\gamma}_i + z_{1-\alpha/2}\sqrt{K^{ii}(\hat{\gamma})} \tag{3.12}$$

In the limit and approximately for large samples, the test statistic with known value γ_i^0 for the parameter γ_i for testing $\mathcal{H}_0 : \gamma_i = \gamma_i^0$ against $\mathcal{H}_1 : \gamma_i \neq \gamma_i^0$ is the same as the Wald test.Under $\mathcal{H}_0$, the limit distribution of Z is standard Normal [17].

$$Z = \frac{\hat{\gamma}_i - \gamma_i^0}{\sqrt{K^{ij}(\hat{\gamma})}} \tag{3.13}$$

3.4. Model fit and forecast

This section introduces model fit criteria and forecast models. Important indicators of model fit criteria include Akaike's Information Criterion (AIC) [1], Bayesian Information Criterion [12], and Schwarz's Criterion (SIC) [14]. Residuals are good indicators for detecting model fit. Various types of residuals are available for different classes of fitted models [11]. Standardized residuals have been used for the proposed model [8].

Based on the Conditional Maximum Likelihood Estimation(CMLE) method, estimates $\hat{\hat{\mu}}_{m+1}, \hat{\hat{\mu}}_{m+2}, \ldots, \hat{\hat{\mu}}_n$,can be obtained as forecast steps .

$$\hat{\hat{\mu}}_t = \mathbf{g}^{-1}(\hat{\alpha} + \mathbf{x}'_t\hat{\beta} + \sum_{i=1}^{p}\hat{\phi}_i(\mathbf{g}(y_{t-i}) - \mathbf{x}_{t-i}\hat{\beta}) + \sum_{j=1}^{q}\hat{\theta}_j r_{t-j}. \tag{3.14}$$

And $r_t = \mathbf{g}(\mathbf{y_t}) - \mathbf{g}(\hat{\hat{\mu}}_\mathbf{t})$ for $t \in \{m+1,\ m+2, \ldots, n\}$ and $h = 1, 2, \ldots, h_0$.
In general, for h steps ahead, the forecast model is obtained as follows:

$$\hat{\hat{\mu}}_{n+h} = \mathbf{g}^{-1}(\hat{\alpha} + \mathbf{x}'_{n+h}\hat{\beta} + \sum_{i=1}^{p}\hat{\phi}_i(\mathbf{g}(\mathbf{y}_{n+h-i}) - \mathbf{x}'_{n+h-i}\hat{\beta}). \tag{3.15}$$

Where $r_t = 0$ for $t > n$ and we have:

$$\mathbf{g}(\mathbf{y_t}) = \begin{cases} \mathbf{g}(\hat{\hat{\mu}}_t), & t > n, \\ \mathbf{g}(y_t), & t \leq n. \end{cases}$$

4. Simulation and Modeling of Real Data

4.1. Simulation

Simulated data were created using the method of transforming the Log-Logistic probability distribution function to a Uniform distribution function. The execution and generation of simulated data were performed using the R software and data modeling in C# software. Considering the introduced model, in the specific case the LLARMA(0.4,0.3) model with 1000 and 10000 observations, it was simulated (Figure 1b), and for the known parameters of the model, the parameter estimates were obtained. The results obtained confirm that the estimates are consistent with the parameters, especially in larger sample sizes; this consistency is more evident (Table 1).

Table 1. The estimate of parameters in LLARMA(0.4,0.3) for different size samples

Known parameter		$\Phi = 0.3$	$\alpha = 0.8$	$\beta_1 = 2$	$\phi_1 = 0.4$	$\theta_1 = 0.3$
n	Est.parameter	$\hat{\Phi}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\phi}_1$	$\hat{\theta}_1$
100	Estimate	0.6032	0.4547	4.9804	0.3024	0.3974
	Standard error	0.2154	0.1523	1.2851	0.3722	0.2547
500	Estimate	0.5178	0.8425	3.3564	0.4103	0.3703
	Standard error	0.2813	0.3587	1.5241	0.0590	0.0592
1000	Estimate	0.4258	0.6537	3.6987	0.3724	0.3569
	Standard error	0.1942	0.2571	1.0024	0.0458	0.0464
5000	Estimate	0.3690	0.7021	3.0870	0.4138	0.2926
	Standard error	0.1764	0.3458	1.2175	0.0206	0.0218
10000	Estimate	0.2894	0.7908	2.2501	0.3974	0.3094
	Standard error	0.1005	0.3054	0.9857	0.0144	0.0217
	Box-Ljung test	X squared = 0.02726		p value = .868		

Additionally, in addition to conjunction with tests, estimates, and goodness-of-fit measures, the structure and cumulative frequency plot can also confirm the fitted model. The Ljung-Box test is used to check if residuals from a time series model exhibit serial correlation (autocorrelation). It assesses whether the model has adequately captured the data's structure or if there's remaining autocorrelation in the residuals, indicating a potential lack of fit. Essentially, it tests if the residuals behave like white noise. It is observed that in this model, the simulation results from the LLARMA(1,1) series, where the dependent variable follows the log-logistic distribution, are observed (Figure 1).

In large sample sizes, the accuracy of the estimate is higher, indicating the consistency of the model parameter estimates. Especially in the case of the 10000 sample size, there is no significant difference between the parameter estimates and the exact parameter values (Table 1).

One of the preliminary criteria for model identification is the observation and examination of the autocorrelation coefficient and the partial autocorrelation coefficient, which this model, considering the structure and simulation from the LLARMA(1,1) series these two indices adequately indicate the accuracy of the simulation and the desirable estimation of the model parameters for a sample with a volume of 10,000 (Figure 2).

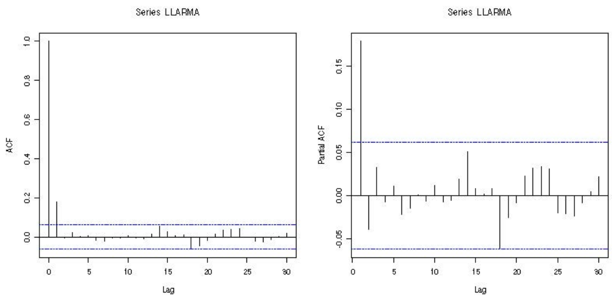
The aim of efficiency and superiority of the proposed LLARMA model, and comparison to ARMA model with Normal distribution two simulations. Although the loglikelihoods index in both models are nearly same, but according to the BIC index, the LLARMA model has been optimized (Table 2).

Table 2. The estimation of parameters in LLARMA(1,1)(Log-Logistic ARMA) and ARMA(1,1)(Normal) for n=10000

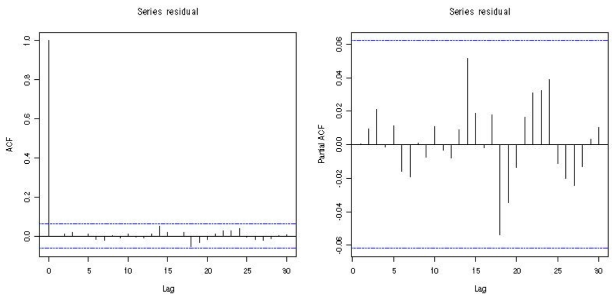
Known parameter		$\Phi = 0.3$	$\alpha = 0.8$	$\beta_1 = 2$	$\phi_1 = 0.4$	$\theta_1 = 0.3$	log *likelihood*
Distribution	Est. parameter	$\hat{\Phi}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\phi}_1$	$\hat{\theta}_1$	*BIC*
LLARMA(1,1)	Estimate	0.2894	0.7908	2.2501	0.3974	0.3094	-14228.06
	Standard error	0.1005	0.3054	0.9857	0.0144	0.0217	24357.37
ARMA(1,1)	Estimate				0.4130	0.2829	-14092.44
	Standard error				0.0146	0.0155	28192.87

(a) n = 1000 **(b)** n = 10000

Figure 1. Cumulative frequency chart of the simulated LLARMA(1,1) series for samples of 1000 and 10000.



(a) ACF of Data **(b)** PACF of Data

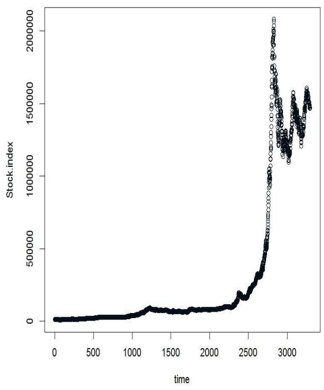


(c) ACF of Residuals **(d)** PACF of Residulas

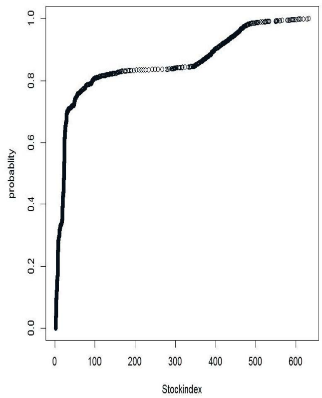
Figure 2. Autocorrelation and partial autocorrelation chart of the simulated LLARMA(1,1) series for data and residuals in a sample with a volume of 10,000

4.2. Application in Modeling and Forecasting of the Overall Stock Market Index

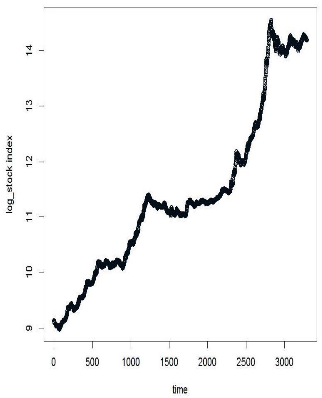
One of the important indicators of the stock market is the overall stock index. This index, also known as the price and cash return index, actually shows whether, on average, the price of stocks and cash dividends in the securities exchange has decreased or increased. Modeling and forecasting this index are of special importance in determining the condition of the stock market of any country. Considering the growth trend of this index in recent years and referring to the charts related to the growth or changes of this index, and drawing the cumulative frequency chart and its probability distribution, the assumption of its normality is weak, and it may deviate from the Normal distribution and follow other distributions. One of them, which is used in economic issues, is the Log-Logistic distribution, from which, for example, the Gini coefficient index is calculated.



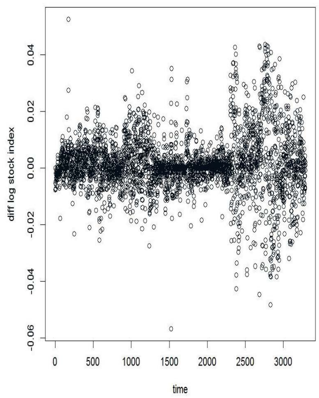
(a) Total Stock Exchange Index



(b) Cumulative Frequency



(c) Log of the Total Stock Index



(d) Diff Log of the Total Stock Index

Figure 3. Functions applied to the total stock market index

The data used in this research are the overall index of the Tehran Stock Exchange in Iran from April 2008 to July 2022, which has been collected and extracted weekly from the Iran Stock Exchange information website.The trend of the data shows a nonlinear pattern of the stock index during the research period. (Figure 3a); the cumulative probability distribution function also shows a structure other than the Normal distribution, and the structure and form of the chart graphically resemble the cumulative distribution of the Log-Logistic distribution to a great extent. (Figure 3b).

Considering the time series chart of the overall stock index, the non-stationarity in the mean is apparent; therefore, by using logarithm operators and differencing once, the non-stationary series in the mean has been transformed into a stationary series (Figures 3c and 3d). Autocorrelation and partial autocorrelation indices of the data (Figure 4a and 4b) and transformed by using logarithm operators and differencing once of the overall stock market index, were determined.(Figure 4c and 4d). Also, according to ACF and PACF in Figure 4c and 4d, model LLARIMA(2,2) can be considered suitable.

Residual's plot, ACF and PACF of residuals of model (LLARIMA(2,2) indicate that the fitted model has maintained the goodness of fit criteria (Table 3, Table 4 and Figure 5).

Noteworthy, since the values of the Log-Logistic distribution are only non-negative data, negative values are createdin logarithmic transformations and differencing. To resolve this issue, the absolute function is applied to negative values.

The model fitting on the observations with logarithmic transformation and the estimation of model parameters using numerical methods based on the conditional likelihood method, particularly the Newton-Raphson method, were obtained. The resulting model is optimized and unique. Given the structure of the Log-Logistic Autoregressive Moving-Average model, besides the overall stock index, which serves as an auxiliary dependent variable in the model, the overall industry index is included as another important index as an independent variable in the model (Table 3). The criterion of stationary was confirmed, and the presence of a unit root was rejected, which favors the appropriateness of the model (Table 4). Related to the autocorrelation and partial autocorrelation coefficients of the data and residuals, and also demonstrate the accuracy of the model fit.

After determining and fitting the model to the observations, forecasting is one of the matters of interest in data analysis and models. In this model, forecasts have also been made for another 100 and 500 times using the forecast function. It should be mentioned that the forecast was calculated after applying changes and transformations to the original function, and the resulting charts in this regard show the actual data on the stock index without the need for retransformation (Figure 6).

Table 3. The model fit criteria

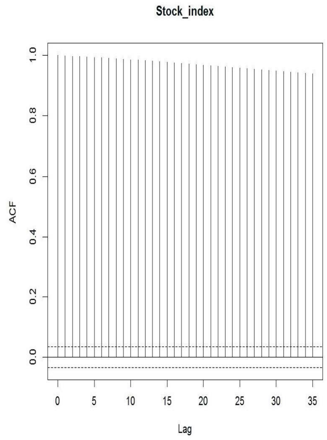
Test name	Statistics	Null hypothesis	P-value
Dickey-Fuller	-11.223	non-stationary	0.01
Dickey-Fuller-unit root test	-37.1407	presence of a unit root	0.000

Table 4. The estimation of parameters of LLARMA(2,2) for the Overall Stock Market Index

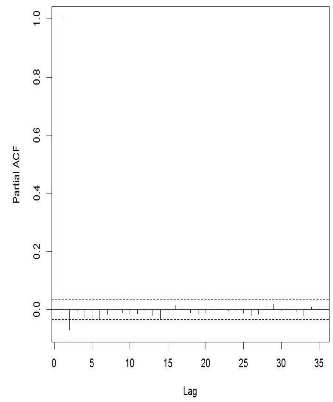
MODEL	Es/Par	$\hat{\Phi}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\Phi = (\phi_1, \phi_2, ..., \phi_p)$ $\hat{\phi}_1$	$\hat{\phi}_2$	$\theta = (\theta_1, \theta_2, ..., \theta_q)$ $\hat{\theta}_1$	$\hat{\theta}_2$	*BIC*	*AIC*	*AICc*
LLARiMA(2,1,2)	Estimate	1.084	11.02	0.0009	0.6672	0.2367	-0.2510	-0.4901	-21383.45	-21420.04	-21420.01
	Standard error	0.034	0.095	0.05	0.069	0.0494	0.0666	0.0335			

5. Conclusions

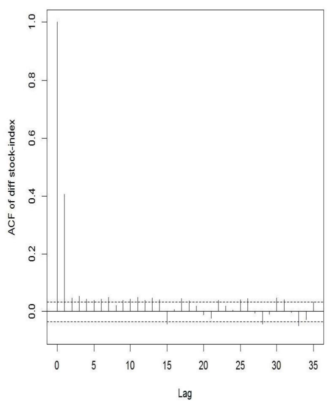
In time series analysis, the assumption of normality of residuals does not always hold; they may follow other distributions. This issue makes the analysis of non-Gaussian series of interest to researchers. Another important matter is the nature of the data and the pattern and probability distribution of data in different subjects, which may not follow conventional patterns or models in time series analysis. In analyzing stock market data, especially the overall index, it is clear from the cumulative probability distribution chart that it does not follow a Normal distribution and shows a considerable match with the Log-Logistic distribution. Therefore, with this resemblance, the time series model, assuming that the



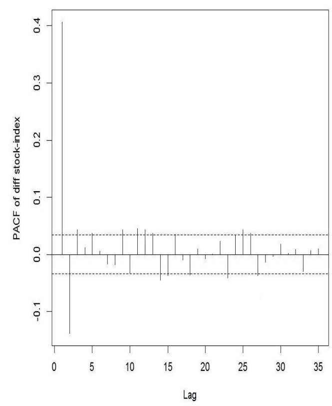
(a) ACF of Total Stock Index



(b) PACF of Total Stock Index



(c) ACF of Diff Log of the Total Stock Index



(d) PACF of Diff Log of the Total Stock Index

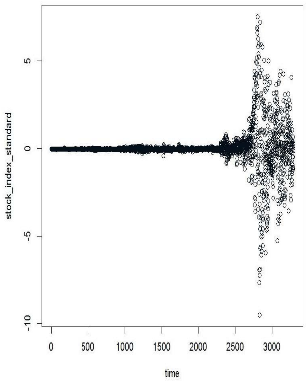
Figure 4. Chart of autocorrelation (ACF) and partial autocorrelation (PACF) for the raw total stock index and for the first-difference of its logarithm.

dependent values of the time series follow the Log-Logistic distribution, the new model introduced in this research was fitted to the data of the overall stock market index. The forecasts made by this model for the overall stock index are accurate and closely match the actual value of the index.

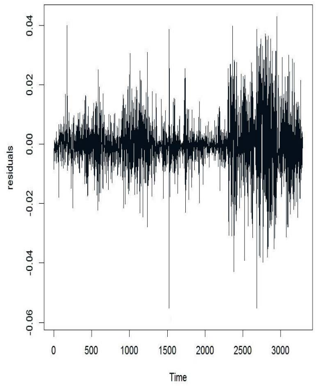
Acknowledgement

I am expressing our deepest gratitude and appreciation for the assistance and support received throughout the completion of this research paper. I want to take this opportunity to acknowledge the contributions of the individuals and institutions who have played a significant role in successfully completing this research project.

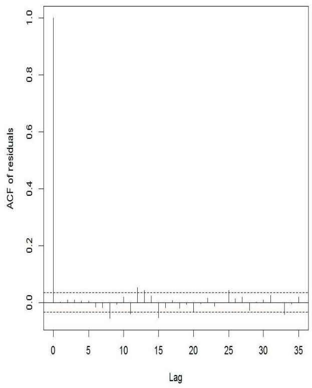
First of all, I would like to thank my supervisor for his invaluable guidance, encouragement, and support throughout this research endeavor. Dr Farshin Hormozi nezhad's insightful feedback was in-



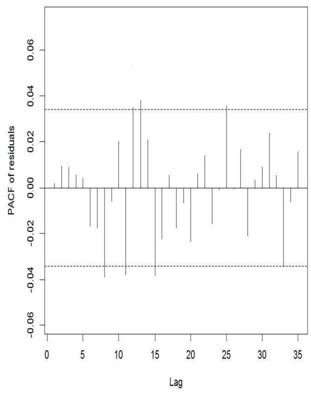
(a) Standardized Diff Log of the Total Stock Index



(b) Residual's plot



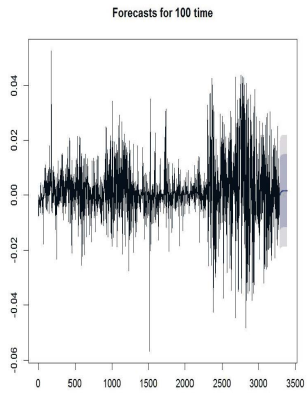
(c) Residual's ACF



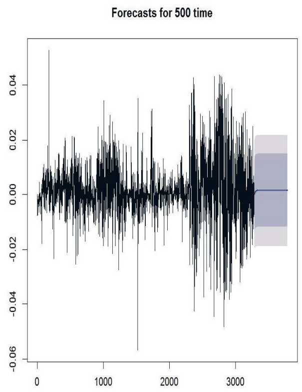
(d) Residual's PACF

Figure 5. ACF and PACF of Residuals plot

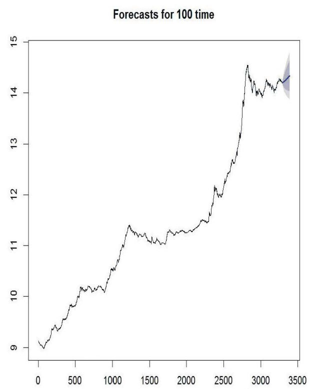
strumental in shaping the directions of this research paper.
I acknowledge the editorial teams and reviewers' contributions, which strengthened the clarity and consistency of this research paper.
Once again, I am grateful to all those who have contributed immensely to my academic journey.



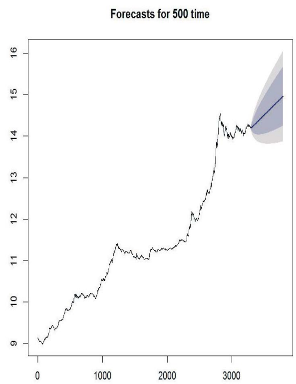
(a) Forecast for 100 future periods by stationary model



(b) Forecast for 500 future periods by stationary model



(c) Forecast for 100 future periods by original data



(d) Forecast for 100 future periods by original data

Figure 6. Forecast of the overall stock market index for 100 and 500 future periods (days).

References

1. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716—723.
2. Bayer, F.M., Bayer, D.M., Pumi, G. (2017). Kumaraswamy autoregressive moving average models for double bounded environmental data. *Journal of Hydrology*, 55(5), 385—396.
3. Benjamin, M.A., Rigby, R.A., Stasinopoulos, D.M. (1998). Fitting Non-Gaussian Time Series Models in COMPSTAT Proceedings in Computational Statistics, eds. R. Payne and P. Green, Heidelberg. *Physica-Verlag*, 191–196.
4. Benjamin, M.A., Rigby, R.A., Stasinopoilos, D.M. (2014). Generalized autoregressive moving average models. *Journal of the American Statistical Association*, 98, 214–223.

5. Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 304–327.
6. Box, G.E. Jenkins, G.M. (1994). *Time Series Analysis: Forecasting and Control*. Prentice Hall, Englewood Cliffs.
7. Ferrari, S.L.P., Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Applied Statistics*, 31, 799–815.
8. Fokianos, K., Kedem, B., Benjamin, M.A. (2003). Regression Theory for Categorical Time Series. *Statistical Science*, 18(3), 357–376.
9. Gourieroux, C.J., Jasiak, J. (2006). Autoregressive Gamma Processes. *Journal of Forecasting*, 25, 129–152.
10. Li, W.K. (1994). Time series models based on generalized linear models: some further results. *Biometrics*, 50, 506–511.
11. Mauricio, J.A. (2008), Computing and using residuals in time series models. *Computational Statistics and Data Analysis*, 52(3), 1746–1763.
12. Raftery, A.E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–196.
13. Rocha, A.V., Cribari-Nemo, F. (2009). Beta autoregressive moving average models. *Test*, 18, 529–545.
14. Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461–464
15. Shoukri, M.M., Mian, I.U.M., Tracy, D.S. (1988). Sampling Properties of Estimators of the Log-Logistic Distribution with Application to Canadian Precipitation Data. *The Canadian Journal of Statistics*, 16(3), 223-–236.
16. Sim, C.H. (1990). First-order autoregressive models for gamma and exponential processes. *J. Prob*, 27, 325–332.
17. Wald, A. (1943). On the Efficient Design of Statistical Investigations. *Ann. Math. Statist*, 14(2), 134-140.
18. Xiaowei, F., Qin, N., Meilan, Z. (2018). A modified quasi-Newton method for nonlinear equations. *Journal of Computational and Applied Mathematics*, 328, 44-–58.
19. Zeger, S.L., Qaqish, B. (1988). Markov Regression Model for Time Serise: A Quasi-Likelihood Approach. *Biometrics*, 44, 1019–1032.

SASAR
Open Access Journal