JIRSS (2020)
Vol. 19, No. 02, pp 33-66
DOI:10.52547/jirss.19.2.33

Modified Maximum Likelihood Estimation in First-Order Autoregressive Moving Average Models with some Non-Normal Residuals

Maryam Kasraie [1] and Abdolreza Sayyareh [2]

[1] Department of Statistics, Science and Research branch, Islamic Azad University, Tehran, Iran.
[2] Department of Computer Science and Statistics, Faculty of Mathematics, K. N. Toosi University of Technology, Tehran, Iran.

Received: 13/02/2020, Revision received: 16/07/2020, Published online: 28/08/2020

Abstract. When modeling time series data using autoregressive-moving average processes, it is a common practice to presume that the residuals are normally distributed. However, sometimes we encounter non-normal residuals and asymmetry of data marginal distribution. Despite widespread use of pure autoregressive processes for modeling non-normal time series, the autoregressive-moving average models have less been used. The main reason is the difficulty in estimating the autoregressive-moving average model parameters. The purpose of this study is to address this intricacy by approximating maximum likelihood estimators, which is particularly important from model selection perspective. Accordingly, the coefficients and residual distribution parameters of the first-order stationary autoregressive-moving average model with residuals that follow exponential and Weibull families, were estimated. Then based on the simulation study, the obtained theoretical results were investigated and it was shown that the modified maximum likelihood estimators were suitable estimators to estimate the first-order autoregressive-moving average model parameters in non-normal mode. In a numerical example positive skewness of obtained residuals from fitting the first-order autoregressive-moving average model was shown. Following that, the parameters of candidate residual distributions estimated by modified maximum likelihood estimators and one of the estimated models for modeling the data was selected.

Keywords. Autoregressive-Moving Average Model, Exponential Family, Modified

M. Kasraie (maryamkasraie@yahoo.com)
Corresponding Author: A. Sayyareh (asayyareh@kntu.ac.ir)

Maximum Likelihood Estimator, Non-Normal Residuals, Weibull Family.

MSC: 62F10, 62M10.

1 Introduction

One of the most popular and widely used stochastic models for describing time series is autoregressive-moving average model (ARMA). A common assumption made to implement this model is the normality of the residuals distribution which generate the process. Under this assumption, the marginal distribution of the observed values is normal and the other issues such as model order selection are considered. Box and Jenkins (1976) provided All the theory behind the ARMA models analysis in the normal mode.

However, in the real world the normality assumption of the residual process is often violated. This can cause asymmetry in the marginal distribution of the observations. This is particularly true about hydrologic datasets such as daily river flow time series (Duca et al. (2019); Sarlak and Sorman (2007); Weiss (1977); Yakowitz (1973)). The river flow time series tend to exhibit saw-tooth and skewed behavior that is not consistent with a normal distribution (Fernandez and Salas (1986); Li and McLeod (1988)). Some other examples are geophysical, financial and economic datasets. Especially, for stock price data the distribution of the residuals appears to be heavy tails and leptokurtic (Fama (1965); Kendall and Hill (1953); Mandelbrot (1967); Mandelbrot (1997); Mandelbrot and Taylor (1967); Spierdijk (2016)).

Based on this argument, it seems that some form of data transformation is often required before classical ARMA modeling (Box and Cox (1964)). However, sometimes simple monotonic transformations do not correct asymmetry (Weiss (1975)). Thus, new approaches have been developed to model non-normal time series with ARMA-type dependency structure. ARMA time series model with exponential distribution was proposed a few years ago by Gaver and Lewis (1980), Li and McLeod (1988) and several other statisticians. Lognormal and gamma distributions received a great deal of attention in the ARMA modeling of hydrological datasets and realized volatility datasets of stocks (Braga and Calmon (2017); Zhang and Li (2019)). In comparison, the Laplace distribution is more applicable to the modeling of certain types of financial and engineering datasets (Trindade et al. (2010)), Bayer et al. (2018) proposed a beta seasonal ARMA model for modeling air relative humidity datasets.

Most researches on non-normal time series modeling and parameter estimation have been conducted by pure autoregressive processes. Nevertheless, modeling based on pure autoregressive (AR) or moving average (MA) processes can lead to a high order model. Generally, in choosing a model, we should try to involve the smallest number of parameters that will represent the time series in an appropriate manner (Cryer and Chan (2008)). It is useful to consider autoregressive and moving average sentences in a single pattern when building a non-normal template. However, the essential problem with the implementation of mixed models is estimation of the parameters.

By considering the desirable properties of maximum likelihood estimators (MLE) in different domains, particularly model selection, they are suggested for estimation procedures. In some cases, however, MLE cannot be determined because the likelihood equations consist of complex functions with no explicit solution. Numerous studies have been carried out as modified methods to estimate AR model parameters by maximum likelihood (ML) method. Tiku et al. (1999) provided modified MLEs for AR models based on linearization of the log-likelihood function in complex phrases using first-order Taylor series expansion. Zamani Mehreyan and Sayyareh (2017) developed this approach to approximate parameters under non-normal residuals of the first-order AR model.

The likelihood function for the autoregressive-moving average processes, even in the normal mode, is more complex than the pure processes. The objective of this research is to address the likelihood function complexity and approximating MLEs to estimate the first-order ARMA model coefficients and residual distribution parameters while the residuals have a non-normal distribution.

2 Modified Maximum Likelihood Estimation

In estimating the parameters of the first-order ARMA models by ML method, we encounter with complex equations which can then be transformed into simpler ones using the modified estimation method. Consider the first-order ARMA model

$$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad t = 2, 3, ..., n, \tag{2.1}$$

in which φ and θ are autoregressive and moving average coefficients, respectively. For stationarity and inversion of the process, it is assumed that $|\varphi| < 1$, $|\theta| < 1$ and $\varphi \neq \theta$. The ϵ_t's are independent and identically distributed random variables (i.i.d) that pursue exponential or Weibull family. Let ω be the vector of considered distribution parameters. Thus, the vector of population parameters to be estimated is, $\gamma = (\omega, \varphi, \theta)$.

Generally, the most important step to study MLEs is to evaluate the likelihood function. The conditional likelihood function for an AR process is conditioned on the initial values of z_t's and for a MA process is conditioned on the initial values of ϵ_t's. Therefore, a common approximation to the likelihood function for an ARMA process becomes conditional on both, z_t's and ϵ_t's. By conditionalization on z_1 and ϵ_1 for the first-order ARMA model, the join distribution of the $(z_2, z_3, ..., z_n)$ can be calculated based on the joint distribution of $(\epsilon_2, \epsilon_3, ..., \epsilon_n)$ (Wei (2006)). On the basis of what is commonly obtained for ARMA models in normal case, we can calculate the conditional likelihood function.

Let us first assumed that the residuals follow a distribution of exponential family. Canonical form of exponential family is

$$f^{\eta}(\epsilon_t) = \exp\{\sum_{j=1}^{k} \eta_j T_j(\epsilon_t) - A(\eta)\}h(\epsilon_t) = \exp\{\eta^T T(\epsilon_t) - A(\eta)\}h(\epsilon_t), \quad t = 2, 3, ..., n.$$

The vector $\eta = (\eta_1, \eta_2, ..., \eta_k)$ is called the natural parameter vector, $\eta_j = c_j(\omega), j = 1, 2, ..., k,$ and $T(\epsilon_t) = (T_1(\epsilon_t), T_2(\epsilon_t), ..., T_k(\epsilon_t))$ represents the sufficient statistic. The function $A(\eta)$ is referred to as the cumulant function and $h(\epsilon_t)$ is a function of ϵ_t. The joint probability density function for $(\epsilon_2, \epsilon_3, ..., \epsilon_n)$ is $f^{\eta}(\epsilon_2, \epsilon_3, ..., \epsilon_n) = \prod_{t=2}^{n} f^{\eta}(\epsilon_t)$. By rewriting model (2.1), we have $\epsilon_t = z_t - \varphi z_{t-1} + \theta\epsilon_{t-1}$. Also, using an iterative search technique ϵ_t is:

$$\epsilon_t = \sum_{j=0}^{t-2} \theta^j (z_{t-j} - \varphi z_{t-j-1}) + \theta^{t-1}\epsilon_1, \quad t = 2, 3, ..., n. \tag{2.2}$$

Although this equation is clearly nonlinear with respect to the parameters, there is a transformation between $(\epsilon_2, \epsilon_3, ..., \epsilon_n)$ and $(z_2, z_3, ..., z_n)$, with Jacobian equal to 1, when we have conditioned on $Z_1 = z_1$ and $\epsilon_1 = E(\epsilon_t)$. Thus, the joint probability density function of $(z_2, z_3, ..., z_n)$ given $Z_1 = z_1$ and $\epsilon_1 = E(\epsilon_t)$ is:

$$\begin{aligned} f^{\eta}(z_2, z_3, ..., z_n | Z_1 = z_1, \epsilon_1 = E(\epsilon_t)) &= \prod_{t=2}^{n} \exp\{\sum_{j=1}^{k} \eta_j T_j(\epsilon_t) - A(\eta)\} h(\epsilon_t) \\ &= \exp\{\sum_{t=2}^{n}\sum_{j=1}^{k} \eta_j T_j(\epsilon_t) - (n-1)A(\eta)\} \prod_{t=2}^{n} h(\epsilon_t). \end{aligned}$$

The logarithm of the likelihood function is obtained as:

$$\begin{aligned} \ell(\eta, \varphi, \theta) &= \sum_{t=2}^{n}\sum_{j=1}^{k} \eta_j T_j(\epsilon_t) - (n-1)A(\eta) + \sum_{t=2}^{n} \log h(\epsilon_t) \\ &= \sum_{t=2}^{n} \eta^T T(\epsilon_t) - (n-1)A(\eta) + \sum_{t=2}^{n} \log h(\epsilon_t. \end{aligned}$$

Then, MLEs are the answers to the following equations:

$$\begin{aligned} \frac{\partial}{\partial\varphi}\ell(\eta, \varphi, \theta) &= \sum_{t=2}^{n} \eta^T \frac{\partial}{\partial\varphi} T(\epsilon_t) + \sum_{t=2}^{n} \frac{\partial}{\partial\varphi} \log h(\epsilon_t) \\ &= \sum_{t=2}^{n} \eta^T \frac{\partial}{\partial\varphi} T(\epsilon_t^* + \mu_\epsilon) + \sum_{t=2}^{n} \frac{\partial}{\partial\varphi} \log\ h(\epsilon_t^* + \mu_\epsilon) = 0, \end{aligned} \tag{2.3}$$

$$\begin{aligned} \frac{\partial}{\partial\theta}\ell(\eta, \varphi, \theta) &= \sum_{t=2}^{n} \eta^T \frac{\partial}{\partial\theta} T(\epsilon_t) + \sum_{t=2}^{n} \frac{\partial}{\partial\theta} \log h(\epsilon_t) \\ &= \sum_{t=2}^{n} \eta^T \frac{\partial}{\partial\theta} T(\epsilon_t^* + \mu_\epsilon) + \sum_{t=2}^{n} \frac{\partial}{\partial\theta} \log\ h(\epsilon_t^* + \mu_\epsilon) = 0, \end{aligned} \tag{2.4}$$

$$\frac{\partial}{\partial\eta_j}\ell(\eta, \varphi, \theta) = \sum_{t=2}^{n} T_j(\epsilon_t) - (n-1)\frac{\partial}{\partial\eta_j} A(\eta) = 0, \quad j = 1, 2, ..., k, \tag{2.5}$$

where $\epsilon_t^* = \epsilon_t - \mu_\epsilon$ and μ_ϵ is mean of ϵ_t. Applying this concept is useful for writing recursive relations as well as simplifying Taylor's expansion relationships. In the case of non-normal residuals, for estimating the model coefficients, the listed log-likelihood equations involve complex functions that there is no explicit solution for them.

Therefore, to make the equations easier, we apply some approximations. For Equations (2.3) and (2.4) Equations (2.6) and (2.7) are defined, respectively, as:

$$M_1(\epsilon_t^*) = \eta^T \frac{\partial}{\partial\varphi} T(\epsilon_t^* + \mu_\epsilon) + \frac{\partial}{\partial\varphi} \log\ h(\epsilon_t^* + \mu_\epsilon), \tag{2.6}$$

$$M_2(\epsilon_t^*) = \eta^T \frac{\partial}{\partial\theta} T(\epsilon_t^* + \mu_\epsilon) + \frac{\partial}{\partial\theta} \log\ h(\epsilon_t^* + \mu_\epsilon). \tag{2.7}$$

By linearizing the functions $M_1(\epsilon_t^*),\ M_2(\epsilon_t^*)$ based on the Taylor's expansion, we have:

$$M_1(\epsilon_t^*) \simeq M_1(\bar{\epsilon}_c) + M_1^{'}(\bar{\epsilon}_c)(\epsilon_t^* - \bar{\epsilon}_c), \tag{2.8}$$

$$M_2(\epsilon_t^*) \simeq M_2(\bar{\epsilon}_c) + M_2^{'}(\bar{\epsilon}_c)(\epsilon_t^* - \bar{\epsilon}_c), \tag{2.9}$$

where $\bar{\epsilon}_c = \bar{\epsilon}_t - \mu_\epsilon$ and ϵ_t^* is defined as:

$$\begin{aligned} \epsilon_t^* &= (z_t - \mu_z) - \varphi(z_{t-1} - \mu_z) + \theta(\epsilon_{t-1} - \mu_\epsilon), \\ \epsilon_t^* &= z_t^* - \varphi z_{t-1}^* + \theta\epsilon_{t-1}^*, \quad t = 2, 3, \ldots, n, \end{aligned} \tag{2.10}$$

where μ_z is the expected value of Z_t. Also, using an iterative search technique and assigning zero to ϵ_1^*, an equal definition of ϵ_t^* is:

$$\epsilon_t^* = \sum_{j=0}^{t-2} \theta^j [z_{t-j}^* - \varphi z_{t-j-1}^*], \quad t = 2, 3, \ldots, n. \tag{2.11}$$

After replacing (2.8) in (2.3) and (2.9) in (2.4), the modified MLEs can be derived from Equations (2.12) and (2.13). However, these equations are based on non-recursive definition of ϵ_t^* that was given in Equation (2.10),

$$\sum_{t=2}^{n} M_1(\epsilon_t^*) = \sum_{t=2}^{n} M_1(\bar{\epsilon}_c) + M_1^{'}(\bar{\epsilon}_c)[z_t^* - \varphi z_{t-1}^* + \theta\epsilon_{t-1}^* - \bar{\epsilon}_c] = 0, \tag{2.12}$$

$$\sum_{t=2}^{n} M_2(\epsilon_t^*) = \sum_{t=2}^{n} M_2(\bar{\epsilon}_c) + M_2^{'}(\bar{\epsilon}_c)[z_t^* - \varphi z_{t-1}^* + \theta\epsilon_{t-1}^* - \bar{\epsilon}_c] = 0. \tag{2.13}$$

On the other hand, by replacing ϵ_t^* recursive relation given in (2.11) in Equations (2.8) and (2.9), we get two equivalent Equations (2.14) and (2.15),

$$\sum_{t=2}^{n} M_1(\epsilon_t^*) = \sum_{t=2}^{n} [M_1(\bar{\epsilon}_c) + M_1^{'}(\bar{\epsilon}_c)(\sum_{j=0}^{t-2} \theta^j [z_{t-j}^* - \varphi z_{t-j-1}^*] - \bar{\epsilon}_c)] = 0, \tag{2.14}$$

$$\sum_{t=2}^{n} M_2(\epsilon_t^*) = \sum_{t=2}^{n} [M_2(\bar{\epsilon}_c) + M_2^{'}(\bar{\epsilon}_c)(\sum_{j=0}^{t-2} \theta^j [z_{t-j}^* - \varphi z_{t-j-1}^*] - \bar{\epsilon}_c)] = 0. \tag{2.15}$$

Consequently, the straight estimated values of Equations (2.12) and (2.13), that we call them non-recursive estimators, are as follows:

$$\hat{\varphi} = \frac{\sum_{t=2}^{n} M_1(\bar{\epsilon}_c) + M_1'(\bar{\epsilon}_c)[z_t^* + \hat{\theta}\epsilon_{t-1}^* - \bar{\epsilon}_c]}{\sum_{t=2}^{n} M_1'(\bar{\epsilon}_c) z_{t-1}^*}, \tag{2.16}$$

$$\hat{\theta} = -\frac{\sum_{t=2}^{n} M_2(\bar{\epsilon}_c) + M_2'(\bar{\epsilon}_c)[z_t^* - \hat{\varphi} z_{t-1}^* - \bar{\epsilon}_c]}{\sum_{t=2}^{n} M_2'(\bar{\epsilon}_c)\epsilon_{t-1}^*}. \tag{2.17}$$

From Equation (2.5), the estimates of residuals distribution parameters are obtained as:

$$\frac{\partial}{\partial \eta_j} A(\eta) = \frac{\sum_{t=2}^{n} T_j(\epsilon_t)}{n-1}, \quad j = 1, 2, ..., k. \tag{2.18}$$

In the same way, the modified MLEs are computed in the Weibull family. Suppose that the residuals distribution belongs to Weibull family with the following presentation:

$$f^{\lambda,\kappa}(\epsilon_t) = \lambda \exp\{-\lambda H(\epsilon_t; \kappa)\} h(\epsilon_t; \kappa), \quad \kappa, \lambda, \epsilon_t > 0.$$

In this equation, $H(\epsilon_t; \kappa)$ is a non-negative monotonically increasing function and $h(\epsilon_t; \kappa)$ is the derivative of $H(\epsilon_t; \kappa)$. Therefore, the log-likelihood of $(z_2, z_3, ..., z_n)$ given $Z_1 = z_1$ and $\epsilon_1 = E(\epsilon_t)$ is:

$$\ell(\lambda, \kappa, \varphi, \theta) = (n-1)\log\lambda - \lambda \sum_{t=2}^{n} H(\epsilon_t; \kappa) + \sum_{t=2}^{n} \log h(\epsilon_t; \kappa).$$

By deriving the log-likelihood function with respect to the unknown parameters and equalizing it to zero, the subsequent Equations (2.19)-(2.22) are obtained:

$$\frac{\partial}{\partial \varphi} \ell(\lambda, \kappa, \varphi, \theta) = \sum_{t=2}^{n} N_1(\epsilon_t^*) = 0, \tag{2.19}$$

$$\frac{\partial}{\partial \theta} \ell(\lambda, \kappa, \varphi, \theta) = \sum_{t=2}^{n} N_2(\epsilon_t^*) = 0, \tag{2.20}$$

$$\frac{\partial}{\partial \lambda} \ell(\lambda, \kappa, \varphi, \theta) = \frac{n-1}{\lambda} - \sum_{t=2}^{n} H(\epsilon_t; \kappa) = 0, \tag{2.21}$$

$$\frac{\partial}{\partial \kappa} \ell(\lambda, \kappa, \varphi, \theta) = -\lambda \sum_{t=2}^{n} \frac{\partial}{\partial \kappa} H(\epsilon_t; \kappa) + \sum_{t=2}^{n} \frac{\partial}{\partial \kappa} \log h(\epsilon_t; \kappa) = 0, \tag{2.22}$$

where

$$N_1(\epsilon_t^*) = -\lambda \frac{\partial}{\partial \varphi} H(\epsilon_t^* + \mu_\epsilon; \kappa) + \frac{\partial}{\partial \varphi} \log h(\epsilon_t^* + \mu_\epsilon; \kappa),$$

and

$$N_2(\epsilon_t^*) = -\lambda \frac{\partial}{\partial \theta} H(\epsilon_t^* + \mu_\epsilon; \kappa) + \frac{\partial}{\partial \theta} \log h(\epsilon_t^* + \mu_\epsilon; \kappa).$$

Since there are no explicit solutions for (2.19) and (2.20), we substitute the linearization form of the functions $N_1(\epsilon_t^*)$ and $N_2(\epsilon_t^*)$ in the equal Equations as follows:

$$N_1(\epsilon_t^*) \simeq N_1(\bar{\epsilon}_c) + N_1^{'}(\bar{\epsilon}_c)(\epsilon_t^* - \bar{\epsilon}_c),$$
$$N_2(\epsilon_t^*) \simeq N_2(\bar{\epsilon}_c) + N_2^{'}(\bar{\epsilon}_c)(\epsilon_t^* - \bar{\epsilon}_c).$$

The MLEs of unknown parameters and model coefficients are obtained by solving the estimating equations whose results are as below:

$$\hat{\varphi} = \frac{\sum_{t=2}^{n} N_1(\bar{\epsilon}_c) + N_1^{'}(\bar{\epsilon}_c)[z_t^* + \hat{\theta}\epsilon_{t-1}^* - \bar{\epsilon}_c]}{\sum_{t=2}^{n} N_1^{'}(\bar{\epsilon}_c) z_{t-1}^*}, \tag{2.23}$$

$$\hat{\theta} = -\frac{\sum_{t=2}^{n} N_2(\bar{\epsilon}_c) + N_2^{'}(\bar{\epsilon}_c)[z_t^* - \hat{\varphi} z_{t-1}^* - \bar{\epsilon}_c]}{\sum_{t=2}^{n} N_2^{'}(\bar{\epsilon}_c)\epsilon_{t-1}^*}, \tag{2.24}$$

$$\hat{\lambda} = (n-1)(\sum_{t=2}^{n} H(\epsilon_t;\kappa))^{-1}, \tag{2.25}$$

$$-\lambda \sum_{t=2}^{n} \frac{d}{d\kappa} H(\epsilon_t;\kappa) + \sum_{t=2}^{n} \frac{d}{d\kappa} \log h(\epsilon_t;\kappa) = 0. \tag{2.26}$$

Coefficients φ and θ can also be calculated by recursive relationships. Thus, some residual distributions are considered for estimating the first-order ARMA model parameters by the gained theoretical results. They are gamma, log-normal and inverse Gaussian from the exponential family and the Weibull and Rayleigh from the Weibull family. Consider model (2.1), where the residuals are i.i.d random variables with gamma distribution, $G(\alpha,\beta)$,

$$f_1^{\omega}(\epsilon_t) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}(\epsilon_t)^{\alpha-1}\exp\{-\frac{\epsilon_t}{\beta}\}, \quad \epsilon_t, \alpha, \beta > 0, \quad t = 2, 3, .., n.$$

Canonical form of the density function is:

$$f_1^{\eta}(\epsilon_t) = \exp\{\eta_{11}\epsilon_t + \eta_{12}\log(\epsilon_t) - A_1(\eta)\},$$

where $\eta_{11} = -1/\beta$, $\eta_{12} = \alpha - 1$, $h_1(\epsilon_t) = 1$ and

$$A_1(\eta) = \log\Gamma(\alpha) + \alpha\log\beta = \log\Gamma(\eta_{12}+1) + (\eta_{12}+1)\log(-\frac{1}{\eta_{11}}).$$

The unknown parameters in the presented model are estimated using Equations (2.16)-(2.18). The calculations are straightforward and the non-recursive estimators are:

$$\hat{\varphi}_{n,1} = \frac{\left(\frac{(\bar{\epsilon}_c+\mu_\epsilon)^2}{\hat{\beta}_n(\hat{\alpha}_n-1)} - (\bar{\epsilon}_c + \mu_\epsilon)\right)\sum_{t=2}^{n} z_{t-1}^*}{\sum_{t=2}^{n} {z_{t-1}^*}^2} + \frac{\sum_{t=2}^{n} z_{t-1}^*(z_t^* + \hat{\theta}_{n,1}\epsilon_{t-1}^* - \bar{\epsilon}_c)}{\sum_{t=2}^{n} {z_{t-1}^*}^2}, \tag{2.27}$$

$$\hat{\theta}_{n,1} = \frac{\left((\bar{\epsilon}_c + \mu_\epsilon) - \frac{(\bar{\epsilon}_c+\mu_\epsilon)^2}{\hat{\beta}_n(\hat{\alpha}_n-1)}\right)\sum_{t=2}^{n} \epsilon_{t-1}^*}{\sum_{t=2}^{n} {\epsilon_{t-1}^*}^2} - \frac{\sum_{t=2}^{n} \epsilon_{t-1}^*(z_t^* - \hat{\varphi}_{n,1} z_{t-1}^* - \bar{\epsilon}_c)}{\sum_{t=2}^{n} {\epsilon_{t-1}^*}^2}, \tag{2.28}$$

$$\hat{\beta}_n = \frac{\sum_{t=2}^{n}(z_t - \hat{\varphi}_{n,1}z_{t-1} + \hat{\theta}_{n,1}\epsilon_{t-1})}{(n-1)\hat{\alpha}_n},$$
$$\Gamma_d(\alpha) = \frac{\sum_{t=2}^{n}\log(z_t - \hat{\varphi}_{n,1}z_{t-1} + \hat{\theta}_{n,1}\epsilon_{t-1})}{n-1} - \log(\hat{\beta}_n),$$

where $\Gamma_d(\alpha)$ is di-gamma function, $(\partial/\partial\alpha)\log\Gamma(\alpha)$, and $\mu_\epsilon = \alpha\beta$. In practical situations, the recursive form of the estimators are more useful. To this end, using Equations (2.14) and (2.15), the estimators in an equivalent form come from the following relations:

$$\hat{\varphi}_{n,1} = \frac{\left(\frac{(\bar{\epsilon}_c+\mu_\epsilon)^2}{\hat{\beta}_n(\hat{\alpha}_n-1)} - (\bar{\epsilon}_c + \mu_\epsilon)\right)\sum_{t=2}^{n}\sum_{j=0}^{t-2}\hat{\theta}_{n,1}^{j}z^*_{t-j-1}}{\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,1}^{j}z^*_{t-j-1})^2} + \frac{\sum_{t=2}^{n}\left[(\sum_{j=0}^{t-2}\hat{\theta}_{n,1}^{j}z^*_{t-j-1})(\sum_{j=0}^{t-2}\hat{\theta}_{n,1}^{j}z^*_{t-j} - \bar{\epsilon}_c)\right]}{\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,1}^{j}z^*_{t-j-1})^2},$$

$$\sum_{t=2}^{n}\Bigg[\Big(\frac{-(\bar{\epsilon}_c + \mu_\epsilon)^2}{\hat{\beta}_n} + (\bar{\epsilon}_c + \mu_\epsilon)(\hat{\alpha}_n - 1)\Big)\sum_{j=0}^{t-2}j\theta^{j-1}(z^*_{t-j} - \hat{\varphi}_{n,1}z^*_{t-j-1}) -$$
$$(\hat{\alpha}_n - 1)\Big(\sum_{j=0}^{t-2}j\theta^{j-1}(z^*_{t-j} - \hat{\varphi}_{n,1}z^*_{t-j-1})\Big)\Big(\sum_{j=0}^{t-2}\theta^{j}(z^*_{t-j} - \hat{\varphi}_{n,1}z^*_{t-j-1}) - \bar{\epsilon}_c\Big)\Bigg] = 0,$$
$$\hat{\beta}_n = \frac{\sum_{t=2}^{n}\left[\sum_{j=0}^{t-2}\hat{\theta}_{n,1}^{j}(z_{t-j} - \hat{\varphi}_{n,1}z_{t-j-1}) + \hat{\theta}_{n,1}^{t-1}\mu_\epsilon\right]}{(n-1)\hat{\alpha}_n},$$
$$\Gamma_d(\alpha) = \frac{\sum_{t=2}^{n}\log\left[\sum_{j=0}^{t-2}\hat{\theta}_{n,1}^{j}(z_{t-j} - \hat{\varphi}_{n,1}z_{t-j-1}) + \hat{\theta}_{n,1}^{t-1}\mu_\epsilon\right]}{n-1} - \log\hat{\beta}_n. \tag{2.29}$$

The assumed ARMA process, with respect to the stationarity conditions and considering $\sum_{h=0}^{\infty}|\gamma(h)| < \infty$ where $\gamma(h) = cov(z_t, z_{t-h})$, is ergodic for the mean (Wei (2006)). A process is said to be ergodic for the mean if the time series average converges to the population mean. Then accordingly, $\sum_t(z_t - \mu_z) = o_p(1)$ and $\bar{z}$ is an unbiased and consistent estimator for μ_z. Moreover, by considering the weak law of large numbers, we conclude that $\bar{\epsilon}_c \xrightarrow{p} 0$. By looking over the aforementioned points, Equations (2.27) and (2.28) can be reduced to simpler formulas. Thus, the equations will be converted to the subsequent relations:

$$\hat{\varphi}_{n,1} \cong \frac{\sum_{t=2}^{n} z^*_{t-1}(z^*_t + \hat{\theta}_{n,1}\epsilon^*_{t-1})}{\sum_{t=2}^{n} {z^*_{t-1}}^2}, \tag{2.30}$$
$$\hat{\theta}_{n,1} \cong \frac{\sum_{t=2}^{n} \epsilon^*_{t-1}(z^*_t - \hat{\varphi}_{n,1}z^*_{t-1})}{\sum_{t=2}^{n} {\epsilon^*_{t-1}}^2}. \tag{2.31}$$

Reduced forms provide acceptable approximations in parameter estimation. From Equations (2.30) and (2.31) and by replacing definition of ϵ^*_{t-1} of Equation (2.11), φ and θ are simply estimable. These estimates become more accurate as the sample size increases. By applying the recursive equations (2.2) and (2.11), and considering the fact that $\epsilon_t = \epsilon^*_t + \mu_\epsilon$, μ_ϵ is estimated. As a result, by substituting ϵ_1 by μ_ϵ and using the estimated values of φ and θ, the residual vector of ϵ_t can be approximated. Emphasizing that the method is not exact, relations (2.30) and (2.31) can be calculated for the other distributions mentioned in this paper. Let us obtain the coefficients and parameter estimators for other non-negative distributions. If ϵ_t's are distributed as log-normal distribution, $LN(M, S)$, we have:

$$\begin{aligned} f_2^{\omega}(\epsilon_t) &= \frac{1}{\epsilon_t\sqrt{2\pi S^2}}\exp\{\frac{-1}{2S^2}(\log\epsilon_t - M)^2\}, \quad \epsilon_t, S > 0, \quad M \in R, \quad t = 2, 3, ..., n, \\ &= \frac{1}{\epsilon_t\sqrt{2\pi}}\exp\{\eta_{21}\Big(\log\epsilon_t\Big)^2 + \eta_{22}\log\epsilon_t - A_2(\eta)\}, \end{aligned}$$

where $\eta_{21} = -1/2S^2$, $\eta_{22} = M/S^2$, $h_2(\epsilon_t) = 1/(\epsilon_t\sqrt{2\pi})$ and

$$A_2(\eta) = \frac{M^2}{2S^2} + \frac{1}{2}\log S^2 = -\frac{\eta_{22}^2}{4\eta_{21}} + \frac{1}{2}\log(-\frac{1}{2\eta_{21}}).$$

Afterwards, the modified MLEs are obtained by (2.16)-(2.18) as follows:

$$\begin{aligned} \hat{\varphi}_{n,2} &= \frac{\Big((\bar{\epsilon}_c + \mu_\epsilon)(\hat{S}_n^2 - \hat{M}_n + \log(\bar{\epsilon}_c + \mu_\epsilon)\Big)\sum_{t=2}^n z_{t-1}^*}{\Big(-\hat{S}_n^2 + \hat{M}_n - \log(\bar{\epsilon}_c + \mu_\epsilon) + 1\Big)\sum_{t=2}^n {z_{t-1}^*}^2} + \frac{\sum_{t=2}^n z_{t-1}^*(z_t^* + \hat{\theta}_{n,2}\epsilon_{t-1}^* - \bar{\epsilon}_c)}{\sum_{t=2}^n {z_{t-1}^*}^2}, \\ \hat{\theta}_{n,2} &= -\Big[\frac{\Big((\bar{\epsilon}_c + \mu_\epsilon)(-\hat{S}_n^2 + \hat{M}_n - \log(\bar{\epsilon}_c + \mu_\epsilon)\Big)\sum_{t=2}^n \epsilon_{t-1}^*}{(\hat{S}_n^2 - \hat{M}_n + \log(\bar{\epsilon}_c + \mu_\epsilon) - 1)\sum_{t=2}^n {\epsilon_{t-1}^*}^2} + \frac{\sum_{t=2}^n \epsilon_{t-1}^*(z_t^* - \hat{\varphi}_{n,2}z_{t-1}^* - \bar{\epsilon}_c)}{\sum_{t=2}^n {\epsilon_{t-1}^*}^2}\Big], \\ \hat{S}_n^2 &= \frac{\sum_{t=2}^n\Big(\log(z_t - \hat{\varphi}_{n,2}z_{t-1} + \hat{\theta}_{n,2}\epsilon_{t-1}) - \hat{M}_n\Big)^2}{n-1}, \\ \hat{M}_n &= \frac{\sum_{t=2}^n \log(z_t - \hat{\varphi}_{n,2}z_{t-1} + \hat{\theta}_{n,2}\epsilon_{t-1})}{n-1}, \end{aligned} \tag{2.32}$$

where $\mu_\epsilon = \exp\{M + \frac{1}{2}S^2\}$. Furthermore based on recursive Equations (2.14) and (2.15), we have the followings relations:

$$\begin{aligned} \hat{\varphi}_{n,2} &= \frac{\Big((\bar{\epsilon}_c + \mu_\epsilon)(\hat{S}_n^2 - \hat{M}_n + \log(\bar{\epsilon}_c + \mu_\epsilon)\Big)\sum_{t=2}^n\sum_{j=0}^{t-2}\hat{\theta}_{n,2}^j z_{t-j-1}^*}{\Big(-\hat{S}_n^2 + \hat{M}_n - \log(\bar{\epsilon}_c + \mu_\epsilon) + 1\Big)\sum_{t=2}^n(\sum_{j=0}^{t-2}\hat{\theta}_{n,2}^j z_{t-j-1}^*)^2} \\ &+ \frac{\sum_{t=2}^n\Big[(\sum_{j=0}^{t-2}\hat{\theta}_{n,2}^j z_{t-j-1}^*)(\sum_{j=0}^{t-2}\hat{\theta}_{n,2}^j z_{t-j}^* - \bar{\epsilon}_c)\Big]}{\sum_{t=2}^n(\sum_{j=0}^{t-2}\hat{\theta}_{n,2}^j z_{t-j-1}^*)^2}, \\ &\sum_{t=2}^n\Big[\Big((\bar{\epsilon}_c + \mu_\epsilon)(-1 - \frac{1}{\hat{S}_n^2}\log(\bar{\epsilon}_c + \mu_\epsilon) + \frac{\hat{M}_n}{\hat{S}_n^2}\Big)\sum_{j=0}^{t-2} j\theta^{j-1}(z_{t-j}^* - \hat{\varphi}_{n,2}z_{t-j-1}^*) - \\ &(-1 - \frac{1}{\hat{S}_n^2}\log(\bar{\epsilon}_c + \mu_\epsilon) + \frac{\hat{M}_n}{\hat{S}_n^2} + \frac{1}{\hat{S}_n^2})\Big(\sum_{j=0}^{t-2} j\theta^{j-1}(z_{t-j}^* - \hat{\varphi}_{n,2}z_{t-j-1}^*)\Big) \end{aligned}$$

$$\times\Big(\sum_{j=0}^{t-2}\theta^j(z^*_{t-j}-\hat{\varphi}_{n,2}z^*_{t-j-1})-\bar{\epsilon}_c\Big)\Big]=0,$$

$$\hat{S}^2_n=\frac{\sum_{t=2}^n\Big[\log(\sum_{j=0}^{t-2}\hat{\theta}^j_{n,2}(z_{t-j}-\hat{\varphi}_{n,2}z_{t-j-1})+\hat{\theta}^{t-1}_{n,2}\mu_\epsilon)-\hat{M}_n\Big]^2}{n-1},$$

$$\hat{M}_n=\frac{\sum_{t=2}^n\log\Big(\sum_{j=0}^{t-2}\hat{\theta}^j_{n,2}(z_{t-j}-\hat{\varphi}_{n,2}z_{t-j-1})+\hat{\theta}^{t-1}_{n,2}\mu_\epsilon\Big)}{n-1}. \tag{2.33}$$

The inverse Gaussian distribution is another distribution that belongs to the two-parameter exponential family with natural parameters $\eta_{31}=-\vartheta/(2\tau^2)$ and $\eta_{32}=-\vartheta/2$,

$$f_3^\omega(\epsilon_t)=\sqrt{\frac{\vartheta}{2\pi\epsilon_t^3}}\exp\{-\frac{\vartheta(\epsilon_t-\tau)^2}{2\tau^2\epsilon_t}\}\quad \epsilon_t,\vartheta,\tau>0,\quad t=2,3,...,n,$$
$$=\frac{1}{\sqrt{2\pi\epsilon_t^3}}\exp\{\eta_{31}\epsilon_t+\eta_{32}\frac{1}{\epsilon_t}-A_3(\eta)\},$$

where $h(\epsilon_t)=\frac{1}{\sqrt{2\pi\epsilon_t^3}}$ and $A_3(\eta)=2\sqrt{\eta_{31}\eta_{32}}-\frac{1}{2}\log(-2\eta_{32})$. In accordance with previous methods, the modified MLEs by (2.16)-(2.18) are:

$$\hat{\varphi}_{n,3}=\frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)\Big(3\hat{\tau}^2_n(\bar{\epsilon}_c+\mu_\epsilon)+\hat{\vartheta}_n(\bar{\epsilon}_c+\mu_\epsilon)^2-\hat{\vartheta}_n\hat{\tau}^2_n\Big)\Big)\sum_{t=2}^n z^*_{t-1}}{\Big(\Big(-3(\bar{\epsilon}_c+\mu_\epsilon)+2\hat{\vartheta}_n\Big)\hat{\tau}^2_n\Big)\sum_{t=2}^n {z^*_{t-1}}^2}$$
$$+\frac{\sum_{t=2}^n z^*_{t-1}(z^*_t+\hat{\theta}_{n,3}\epsilon^*_{t-1}-\bar{\epsilon}_c)}{\sum_{t=2}^n {z^*_{t-1}}^2},$$

$$\hat{\theta}_{n,3}=-\Big[\frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)\Big(-3\hat{\tau}^2_n(\bar{\epsilon}_c+\mu_\epsilon)-\hat{\vartheta}_n(\bar{\epsilon}_c+\mu_\epsilon)^2+\hat{\vartheta}_n\hat{\tau}^2_n\Big)\Big)\sum_{t=2}^n \epsilon^*_{t-1}}{\Big(\Big(3(\bar{\epsilon}_c+\mu_\epsilon)-2\hat{\vartheta}_n\Big)\hat{\tau}^2_n\Big)\sum_{t=2}^n {\epsilon^*_{t-1}}^2}$$
$$+\frac{\sum_{t=2}^n \epsilon^*_{t-1}(z^*_t-\hat{\varphi}_{n,3}z^*_{t-1}-\bar{\epsilon}_c)}{\sum_{t=2}^n {\epsilon^*_{t-1}}^2}\Big],$$

$$\hat{\tau}_n=\frac{\sum_{t=2}^n(z_t-\hat{\varphi}_{n,3}z_{t-1}+\hat{\theta}_{n,3}\epsilon_{t-1})}{n-1},$$

$$\hat{\vartheta}_n=\frac{n-1}{\sum_{t=2}^n(\frac{1}{z_t-\hat{\varphi}_{n,3}z_{t-1}+\hat{\theta}_{n,3}\epsilon_{t-1}}-\frac{1}{\hat{\tau}_n})}. \tag{2.34}$$

Additionally, in recursive form using (2.14) and (2.15) we obtain:

$$\hat{\varphi}_{n,3}=\frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)(3\hat{\tau}^2_n(\bar{\epsilon}_c+\mu_\epsilon)+\hat{\vartheta}_n(\bar{\epsilon}_c+\mu_\epsilon)^2-\hat{\vartheta}_n\hat{\tau}^2_n)\Big)\sum_{t=2}^n\sum_{j=0}^{t-2}\hat{\theta}^j_{n,3}z^*_{t-j-1}}{\Big(\Big(-3(\bar{\epsilon}_c+\mu_\epsilon)+2\hat{\vartheta}_n\Big)\hat{\tau}^2_n\Big)\sum_{t=2}^n(\sum_{j=0}^{t-2}\hat{\theta}^j_{n,3}z^*_{t-j-1})^2}+$$

$$\frac{\sum_{t=2}^{n}\left[(\sum_{j=0}^{t-2}\hat{\theta}_{n,3}^{j}z_{t-j-1}^{*})(\sum_{j=0}^{t-2}\hat{\theta}_{n,3}^{j}z_{t-j}^{*}-\bar{\epsilon}_c)\right]}{\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,3}^{j}z_{t-j-1}^{*})^2},$$

$$\begin{aligned}
&\sum_{t=2}^{n}\Big[\Big(\frac{(\bar{\epsilon}_c+\mu_\epsilon)(-3\hat{\tau}_n^2(\bar{\epsilon}_c+\mu_\epsilon)-\hat{\vartheta}_n(\bar{\epsilon}_c+\mu_\epsilon)^2+\hat{\vartheta}_n\hat{\tau}_n^2)}{\hat{\tau}_n^2}\Big)\sum_{j=0}^{t-2}j\theta^{j-1}(z_{t-j}^{*}-\hat{\varphi}_{n,3}z_{t-j-1}^{*})- \\
&\quad\Big(-3(\bar{\epsilon}_c+\mu_\epsilon)+2\hat{\vartheta}_n\Big)\Big(\sum_{j=0}^{t-2}j\theta^{j-1}(z_{t-j}^{*}-\hat{\varphi}_{n,3}z_{t-j-1}^{*})\Big)\Big(\sum_{j=0}^{t-2}\theta^{j}(z_{t-j}^{*}-\hat{\varphi}_{n,3}z_{t-j-1}^{*})-\bar{\epsilon}_c\Big)\Big]=0, \\
&\hat{\tau}_n=\frac{\sum_{t=2}^{n}\Big(\sum_{j=0}^{t-2}\hat{\theta}_{n,3}^{j}(z_{t-j}-\hat{\varphi}_{n,3}z_{t-j-1})+\hat{\theta}_{n,3}^{t-1}\mu_\epsilon\Big)}{n-1}, \\
&\hat{\vartheta}_n=\frac{n-1}{\sum_{t=2}^{n}\Big(\frac{1}{\sum_{j=0}^{t-2}\hat{\theta}_{n,3}^{j}(z_{t-j}-\hat{\varphi}_{n,3}z_{t-j-1})+\hat{\theta}_{n,3}^{t-1}\mu_\epsilon}-\frac{1}{\hat{\tau}_n}\Big)},
\end{aligned}\tag{2.35}$$

where $\mu_\epsilon=\tau$.

Likewise, we can estimate the parameters of the first-order ARMA model in the case of Weibul family. Consider the first-order ARMA model with Weibull distribution for residuals, $W(\kappa,\xi)$,

$$f_4^{\omega}(\epsilon_t)=\frac{1}{\xi^{\kappa}}(\kappa\epsilon_t^{\kappa-1})\exp\{-(\frac{\epsilon_t}{\xi})^{\kappa}\},\quad \epsilon_t,\kappa,\xi>0,\quad t=2,3,...,n,$$

where $H(\epsilon_t;\kappa)=\epsilon_t^{\kappa}$ and $\lambda_{41}=1/\xi^{\kappa}$. The modified MLEs are acquired using Equations (2.23)-(2.26). The resulting relationships are shown as below:

$$\begin{aligned}
\hat{\varphi}_{n,4}&=\frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)\Big(\hat{\kappa}_n(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n}-(\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}\Big)\Big)\sum_{t=2}^{n}z_{t-1}^{*}}{\Big((\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}+\hat{\kappa}_n(\hat{\kappa}_n-1)(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n}\Big)\sum_{t=2}^{n}{z_{t-1}^{*}}^2}+\frac{\sum_{t=2}^{n}z_{t-1}^{*}(z_t^{*}+\hat{\theta}_{n,4}\epsilon_{t-1}^{*}-\bar{\epsilon}_c)}{\sum_{t=2}^{n}{z_{t-1}^{*}}^2}, \\
\hat{\theta}_{n,4}&=-\Big[\frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)\Big(-\hat{\kappa}_n(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n}+(\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}\Big)\Big)\sum_{t=2}^{n}\epsilon_{t-1}^{*}}{\Big(-(\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}-\hat{\kappa}_n(\hat{\kappa}_n-1)(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n}\Big)\sum_{t=2}^{n}{\epsilon_{t-1}^{*}}^2}+\frac{\sum_{t=2}^{n}\epsilon_{t-1}^{*}(z_t^{*}-\hat{\varphi}_{n,4}z_{t-1}^{*}-\bar{\epsilon}_c)}{\sum_{t=2}^{n}{\epsilon_{t-1}^{*}}^2}\Big], \\
\hat{\xi}_n&=\Big[\frac{\sum_{t=2}^{n}(z_t-\hat{\varphi}_{n,4}z_{t-1}+\hat{\theta}_{n,4}\epsilon_{t-1})^{\hat{\kappa}_n}}{n-1}\Big]^{\frac{1}{\hat{\kappa}_n}}, \\
&\frac{\sum_{t=2}^{n}(z_t-\hat{\varphi}_{n,4}z_{t-1}+\hat{\theta}_{n,4}\epsilon_{t-1})^{\kappa}\log(z_t-\hat{\varphi}_{n,4}z_{t-1}+\hat{\theta}_{n,4}\epsilon_{t-1})}{\sum_{t=2}^{n}(z_t-\hat{\varphi}_{n,4}z_{t-1}+\hat{\theta}_{n,4}\epsilon_{t-1})^{\kappa}}-\frac{1}{\kappa} \\
&-\frac{\sum_{t=2}^{n}\log(z_t-\hat{\varphi}_{n,4}z_{t-1}+\hat{\theta}_{n,4}\epsilon_{t-1})}{n-1}=0,
\end{aligned}\tag{2.36}$$

where $\mu_\epsilon=\xi\Gamma(1+1/\kappa)$. Moreover, according to relations (2.2) and (2.11), recursive equations can be calculated as depicted in equation (2.37):

$$\hat{\varphi}_{n,4} = \frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)\Big(\hat{\kappa}_n(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n}-(\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}\Big)\Big)\sum_{t=2}^{n}\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}z^*_{t-j-1}}{\Big((\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}+\hat{\kappa}_n(\hat{\kappa}_n-1)(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n}\Big)\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}z^*_{t-j-1})^2}$$
$$+\frac{\sum_{t=2}^{n}\Big[(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}z^*_{t-j-1})(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}z^*_{t-j}-\bar{\epsilon}_c)\Big]}{\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}z^*_{t-j-1})^2},$$

$$\sum_{t=2}^{n}\Big[\Big(-\hat{\kappa}_n(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n+1}+(\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}(\bar{\epsilon}_c+\mu_\epsilon)\Big)\sum_{j=0}^{t-2}j\theta^{j-1}(z^*_{t-j}-\hat{\varphi}_{n,4}z^*_{t-j-1})-$$
$$\Big((\hat{\kappa}_n-1)\hat{\xi}_n^{\hat{\kappa}_n}+\hat{\kappa}_n(\hat{\kappa}_n-1)(\bar{\epsilon}_c+\mu_\epsilon)^{\hat{\kappa}_n}\Big)\Big(\sum_{j=0}^{t-2}j\theta^{j-1}(z^*_{t-j}-\hat{\varphi}_{n,4}z^*_{t-j-1})\Big)\times$$
$$\Big(\sum_{j=0}^{t-2}\theta^{j}(z^*_{t-j}-\hat{\varphi}_{n,4}z^*_{t-j-1})-\bar{\epsilon}_c\Big)\Big]=0,$$

$$\hat{\xi}_n=\Big[\frac{\sum_{t=2}^{n}\Big(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}(z_{t-j}-\hat{\varphi}_{n,4}z_{t-j-1})+\hat{\theta}_{n,4}^{t-1}\mu_\epsilon\Big)^{\hat{\kappa}_n}}{n-1}\Big]^{\frac{1}{\hat{\kappa}_n}},$$

$$\frac{\sum_{t=2}^{n}\Big(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}(z_{t-j}-\hat{\varphi}_{n,4}z_{t-j-1})+\hat{\theta}_{n,4}^{t-1}\mu_\epsilon\Big)^{\kappa}\log\Big(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}(z_{t-j}-\hat{\varphi}_{n,4}z_{t-j-1})+\hat{\theta}_{n,4}^{t-1}\mu_\epsilon\Big)}{\sum_{t=2}^{n}\Big(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}(z_{t-j}-\hat{\varphi}_{n,4}z_{t-j-1})+\hat{\theta}_{n,4}^{t-1}\mu_\epsilon\Big)^{\kappa}}-\frac{1}{\kappa}$$
$$-\frac{\sum_{t=2}^{n}\log\Big(\sum_{j=0}^{t-2}\hat{\theta}_{n,4}^{j}(z_{t-j}-\hat{\varphi}_{n,4}z_{t-j-1})+\hat{\theta}_{n,4}^{t-1}\mu_\epsilon\Big)}{n-1}=0. \tag{2.37}$$

A popular distribution in the Weibull family is the Reyligh distribution and occurs when $\kappa=2$, $H(\epsilon_t;2)=\epsilon_t^2/2$ and $\lambda_{51}=1/\sigma^2$. The density of this distribution is:

$$f_5^{\omega}(\epsilon_t)=\frac{\epsilon_t}{\sigma^2}\exp\Big\{-\frac{\epsilon_t^2}{2\sigma^2}\Big\},\quad \epsilon_t,\sigma>0,\quad t=2,3,...,n,$$

and $\mu_\epsilon=\sigma\sqrt{\pi/2}$. Then, the modified MLEs by Equations (2.23)-(2.26) are:

$$\hat{\varphi}_{n,5}=\frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)\Big((\bar{\epsilon}_c+\mu_\epsilon)^2-\hat{\sigma}_n^2\Big)\Big)\sum_{t=2}^{n}z^*_{t-1}}{\Big(\hat{\sigma}_n^2+(\bar{\epsilon}_c+\mu_\epsilon)^2\Big)\sum_{t=2}^{n}{z^*_{t-1}}^2}+\frac{\sum_{t=2}^{n}z^*_{t-1}(z^*_t+\hat{\theta}_{n,5}\epsilon^*_{t-1}-\bar{\epsilon}_c)}{\sum_{t=2}^{n}{z^*_{t-1}}^2},$$
$$\hat{\theta}_{n,5}=-\Big[\frac{\Big((\bar{\epsilon}_c+\mu_\epsilon)\Big(-(\bar{\epsilon}_c+\mu_\epsilon)^2+\hat{\sigma}_n^2\Big)\Big)\sum_{t=2}^{n}\epsilon^*_{t-1}}{\Big(-\hat{\sigma}_n^2-(\bar{\epsilon}_c+\mu_\epsilon)^2\Big)\sum_{t=2}^{n}{\epsilon^*_{t-1}}^2}+\frac{\sum_{t=2}^{n}\epsilon^*_{t-1}(z^*_t-\hat{\varphi}_{n,5}z^*_{t-1}-\bar{\epsilon}_c)}{\sum_{t=2}^{n}{\epsilon^*_{t-1}}^2}\Big],$$
$$\hat{\sigma}_n=\Big[\frac{\sum_{t=2}^{n}(z_t-\hat{\varphi}_{n,5}z_{t-1}+\hat{\theta}_{n,5}\epsilon_{t-1})^2}{2(n-1)}\Big]^{\frac{1}{2}}. \tag{2.38}$$

Using Equations (2.2) and (2.11), we have:

$$\hat{\varphi}_{n,5} = \frac{\Big((\bar{\epsilon}_c + \mu_\epsilon)\Big((\bar{\epsilon}_c + \mu_\epsilon)^2 - \hat{\sigma}_n^2\Big)\Big)\sum_{t=2}^{n}\sum_{j=0}^{t-2}\hat{\theta}_{n,5}^{j} z_{t-j-1}^*}{\Big(\hat{\sigma}_n^2 + (\bar{\epsilon}_c + \mu_\epsilon)^2\Big)\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,5}^{j} z_{t-j-1}^*)^2}$$
$$+ \frac{\sum_{t=2}^{n}\Big[(\sum_{j=0}^{t-2}\hat{\theta}_{n,5}^{j} z_{t-j-1}^*)(\sum_{j=0}^{t-2}\hat{\theta}_{n,5}^{j} z_{t-j}^* - \bar{\epsilon}_c)\Big]}{\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,5}^{j} z_{t-j-1}^*)^2},$$
$$\sum_{t=2}^{n}\Big[\Big(-(\bar{\epsilon}_c + \mu_\epsilon)^3 + \hat{\sigma}_n^2(\bar{\epsilon}_c + \mu_\epsilon)\Big)\sum_{j=0}^{t-2} j\theta^{j-1}(z_{t-j}^* - \hat{\varphi}_{n,4} z_{t-j-1}^*) -$$
$$\Big(\hat{\sigma}_n^2 + (\bar{\epsilon}_c + \mu_\epsilon)^2\Big)\Big(\sum_{j=0}^{t-2} j\theta^{j-1}(z_{t-j}^* - \hat{\varphi}_{n,5} z_{t-j-1}^*)\Big)\Big(\sum_{j=0}^{t-2}\theta^{j}(z_{t-j}^* - \hat{\varphi}_{n,5} z_{t-j-1}^*) - \bar{\epsilon}_c\Big)\Big] = 0,$$
$$\hat{\sigma}_n = \Big[\frac{\sum_{t=2}^{n}(\sum_{j=0}^{t-2}\hat{\theta}_{n,5}^{j}(z_{t-j} - \hat{\varphi}_{n,5} z_{t-j-1}) + \hat{\theta}_{n,4}^{t-1}\mu_\epsilon)^2}{2(n-1)}\Big]^{\frac{1}{2}}. \tag{2.39}$$

Since the introduced estimators are the modified versions of the MLEs for estimating the first-order ARMA model parameters, we expect them to reflect the desirable properties of the MLEs. One of the most important properties to be assessed for the modified estimators is their asymptotic normality. Asymptotic theory gives us a more complete picture of the estimator's behavior as the sample size increases. The asymptotic distribution of the modified MLEs vector is provided in Theorem 2.1.

Theorem 2.1 (Asymptotic distribution)**.** *Suppose that* $\{Z_t\}_{t=1}^n$ *is a stationary time series modeled by Equation* (2.1)*, and the* ϵ_t*'s are i.i.d random variables generated by f(.). Then the asymptotic distribution of the estimator* γ_n *is as follows:*

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} N(0, nC).$$

where $C = I_n^{-1}(\gamma_0)J_n(\gamma_0)I_n^{-1}(\gamma_0)$, $I_n(\gamma) = -E\Big[\frac{\partial^2 \ell_n^f(\gamma)}{\partial\gamma\partial\gamma^T}\Big]$,$J_n(\gamma) = E\Big[\Big(\frac{\partial \ell_n^f(\gamma)}{\partial\gamma}\Big)\Big(\frac{\partial \ell_n^f(\gamma)}{\partial\gamma^T}\Big)\Big]$ *and* γ_0 *is the true value of parameter.*

Proof. Assuming the regularity conditions are established, we write the Taylor expansion of $(\partial/\partial\gamma)\ell_n^f(\hat{\gamma}_n) = (\partial/\partial\gamma)(\sum_{t=2}^{n}\log f^{\hat{\gamma}_n}(\epsilon_t))$ around γ_0 as follows:

$$\frac{\partial}{\partial\gamma}\ell_n^f(\hat{\gamma}_n) = \frac{\partial}{\partial\gamma}\ell_n^f(\gamma)|_{\gamma_0} + \frac{\partial^2}{\partial\gamma\partial\gamma^T}\ell_n^f(\gamma)(\hat{\gamma}_n - \gamma)|_{\gamma_0} + o_p(1).$$

The notation $o_p(1)$ demonstrates a quantity which is convergent to zero in probability when n tends to infinity. Since $\frac{\partial}{\partial\gamma}\ell_n^f(\hat{\gamma}_n) = 0$, the relationship can be rewritten as:

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) = -\frac{n^{\frac{1}{2}}\frac{\partial}{\partial\gamma}\ell_n^f(\gamma)|_{\gamma_0}}{\frac{\partial^2}{\partial\gamma\partial\gamma^T}\ell_n^f(\gamma)|_{\gamma_0}} + o_p(1).$$

Using the weak law of large numbers, it can be concluded that

$$-\frac{\partial^2}{\partial\gamma\partial\gamma^T}\ell_n^f(\gamma)|_{\gamma_0} \xrightarrow{p} I_n(\gamma_0).$$

On the basis of the regularity conditions, the derivative and integration operations are interchangeable, so we have $E\Big[\frac{\partial}{\partial\gamma}\ell_n^f(\gamma)|_{\gamma_0}\Big] = 0$. Now from the multivariate central limit theorem, we have $\frac{\partial}{\partial\gamma}\ell_n^f(\gamma)|_{\gamma_0} \xrightarrow{d} N(0, J_n(\gamma_0))$, and, consequently,

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} N(0, nI_n^{-1}(\gamma_0)J_n(\gamma_0)I_n^{-1}(\gamma_0)).$$

In other words,

$$\hat{\gamma_n} \sim N(\gamma_0, I_n^{-1}(\gamma_0)J_n(\gamma_0)I_n^{-1}(\gamma_0)).$$

□

3 Simulation

In this section, simulation studies have been used to investigate and analyze the results of section 2. We have shown that the modified MLEs obtained in the previous section are appropriate estimators for the first-order ARMA model parameters. The studies have been conducted in two parts. In the first part, we have focused on examining the estimators' capability to estimate different values of model coefficients and parameters based on different distributions. In the second part, different estimation methods, including recursive, non-recursive and reduced forms, have been compared in terms of capability and accuracy in estimating model parameters. For both parts, the time series vector data have been generated in the same way. To construct the time series vector first, the samples of residuals under non-negative distributions, discussed in this article, were randomly generated in different sizes. Following that, for each sample using various values of coefficients φ and θ, the time series observation vectors were produced. Then, the model coefficients and parameters were estimated by modified method based on just generated time series vector or generated residual and time series vectors.

This simulation was coded in R software in which various packages, such as rootSolve and BB, were utilized to solve a nonlinear equations system in some cases.

In Tables (1)-(7) for each estimator, a column was made to demonstrate the mean square error values (MSE). The MSE values were used to measure the average squared difference between the estimated and the actual values. Furthermore, in order to study the results more accurately, approximate confidence intervals were calculated based on the modified MLEs and their results were also listed in Tables (1)-(7).

The approximate ML confidence intervals (MLCI) are some asymptotic versions of confidence sets that are used in more complicated situations. Such confidence intervals are created entirely on the basis of the asymptotic properties of the MLEs. If $X_1, X_2, ..., X_n$ are i.i.d of $f(x|\Theta)$ in which $\Theta \in \mathbb{R}^k$ and $\hat{\theta}_i$ is the MLE of θ_i, a two-side MLCI for θ_i in the confidence level of $(1-\alpha)$, considering the asymptotic normal distribution of the MLEs, can be constructed as:

$$\hat{\theta}_i - Z_{\frac{\alpha}{2}} \sqrt{\widehat{var}(\hat{\theta}_i)} < \theta_i < \hat{\theta}_i + Z_{\frac{\alpha}{2}} \sqrt{\widehat{var}(\hat{\theta}_i)}, \qquad i = 1, 2, ..., k.$$

The true variance of the MLE is approximated with respect to the asymptotic efficiency property of the MLEs. With this approch, the Cramer-Rao Lower Bond can be used as an estimation of the variance of the MLE. Accordingly, $\widehat{var}(\hat{\theta}_i) = \lambda I_n^{-1}(\Theta)\lambda^T$ where $\lambda = (\lambda_1, \lambda_2, ..., \lambda_k)$, $\lambda_i = \frac{\partial}{\partial \theta_i}\Theta$ and $I_n(\Theta) = -\frac{\partial}{\partial \Theta \partial \Theta^T} \ell_n^f(\Theta)|_{\Theta=\hat{\Theta}}$ is an approximation for the Fisher information matrix. Description of this finding can be found in Casella and Berger (2001), Cramér (1946) and Huber (1967).

In the first part, since the reduced and recursive methods are constructed based on the non-recursive method in parameter estimation, the parameters were approximated using non-recursive method. Tables (1)-(4) display the non-recursive parameter estimation results along with the corresponding MSEs, MLCIs and lengths of the MLCIs (LCI). Each of the Tables (1)-(4) includes two sections: a and b; each pair of a and b belongs to the parameter estimation results in the first-order ARMA model under an identical residual distribution. The model coefficients selected for estimation in these tables are the combination of even and odd numbers with different signs. It is not possible to investigate all states of φ and θ, due to their high diversity. Therefore, it is attempted to consider the modalities that cover all situations well as much as possible. In this case $\varphi = -0.5, -0.1, 0.2, 0.8$ and $\theta = -0.6, -0.4, 0.3, 0.5$ were considered. The number of replicates were 10^4 and the sample sizes were $n = 50, 150, 500$.

It can be seen in Tables (1) to (4) that the proximity or distance of the model coefficients, as well as their signs, do not affect the performance of the modified estimators in estimating the distribution parameters and model coefficients. For example, in the case of $\varphi = 0.2$ and $\theta = 0.3$, the estimated values and their MSEs have similar accuracy to those of $\varphi = 0.8$ and $\theta = -0.6$. On the other hand, the accuracy of the estimated values of the model coefficients does not change by varying the residual distributions. Comparing the results of Tables (1.b) to (4.b) implies that the accuracy of the estimators and their MSEs is approximately the same in estimating the model coefficients according to the different residual distributions. These results demonstrate the capability of the proposed estimators in estimating the various model coefficient values and distribution parameters.

In addition, all the calculated MLCIs in all the tables contain the correct values of their respective model parameters and coefficients. In other words, there is not any case that a parameter or a coefficient would be outside of its confidence interval. Even in estimating φ coefficient, whose LCIs are shorter than those of other parameters, there is no coefficient value outside the confidence interval. Cencerning the accuracy of estimators, using the results of Tables (1)-(4), it can be found that when the sample size increases, the estimated values get closer to the true parameter values. In this case, both the MSEs and LCIs decline with increasing the sample size. A notable case observed in the Tables is the large values of MSEs and LCIs in estimating the shape parameter of the inverse Gaussian distribution (ϑ) compared to the similar values calculated in estimating the other distribution parameters (Table (3.a)). Large amounts of the MSEs and LCIs in estimating the shape parameter of the inverse Gaussian distribution decrease rapidly with increasing the sample size.

In the second part, time series data were generated on the basis of log-normal distribution residuals with parameters $M = 0$ and $S = 0.75$ to compare different parameter estimation approaches. The results of estimating the model parameters and coefficients using each of the non-recursive, recursive and reduced methods have been summarized in Tables (5), (6) and (7). The coefficients for these analyses were $\varphi = -0.6, 0.3$, and $\theta = -0.7, 0.2$ and the sample sizes were $n = 50, 350, 1000$. Similar to previous situation, the number of iterations was 10^4.

It can be concluded from the results given in Tables (5) and (6) that, although the non-recursive estimators are slightly more accurate in estimating the distribution parameters than the recursive estimators, there is generally no significant difference in the analogous estimated values. In other words, both recursive and non-recursive estimation techniques work with approximately equal precision in terms of parameter estimations. Minor differences in the accuracy of estimating the distribution parameters between recursive and non-recursive methods can be better seen in the first 6 columns of the Tables 5 and 6 when the sample size is 50. A comparison of the columns 1 and 4 in both tables shows that the estimated values of the lognormal distribution parameters, M and S, by non-recursive method are a bit closer to the true parameter values than the parameters estimated by the recursive method. Moreover, MSEs of the estimators listed in columns 2 and 5 of both tables are greater in the recursive state than in the non-recursive state. This discrepancy in estimating the distribution parameters by recursive and non-recursive methods decreases with increasing the sample size. However, both methods have similar accuracy in estimating model coefficients.

Table (7) shows the results of the parameter estimation by the reduced method. The estimated values of the model coefficients that are demonstrated in the columns 7 and 10 of this table are as good as the estimated values based on the recursive and non-recursive methods. Nonetheless, the estimated values of the distribution parameters using reduced method are more distanced from the actual parameter values than the estimated values of the recursive and non-recursive methods. This distance is not large and decreases as the sample size increases. In all cases, the estimated MSE values by reduced model are greater than MSEs measured by the two other methods. The greater MSEs suggest a higher error rate in estimating model parameters using the reduced approach than recursive and non-recursive approaches. Furthermore, the confidence intervals constructed based on the results of all three approaches cover the actual values of the distribution parameters and model coefficients.

Overall, it can be inferred that the distribution parameters and model coefficients are well estimated with a slight difference using all three methods. In Tables (1)-(7), the general trend improved estimators' accuracy with increasing the sample size. Moreover, the modified estimators derived from different residual distributions, estimate the model coefficients with the same precision.

4 Case Study and Results

To demonstrate the capability of the modified MLEs in estimating the first-order ARMA model parameters in non-normal mode a numerical example has been presented. The data consist of annual percentages of the unemployment rate among the population of the United States in the range of 15-64 age group from January 1960 to 2018 that can be accessed from Federal Reserve Economic Data of the US.

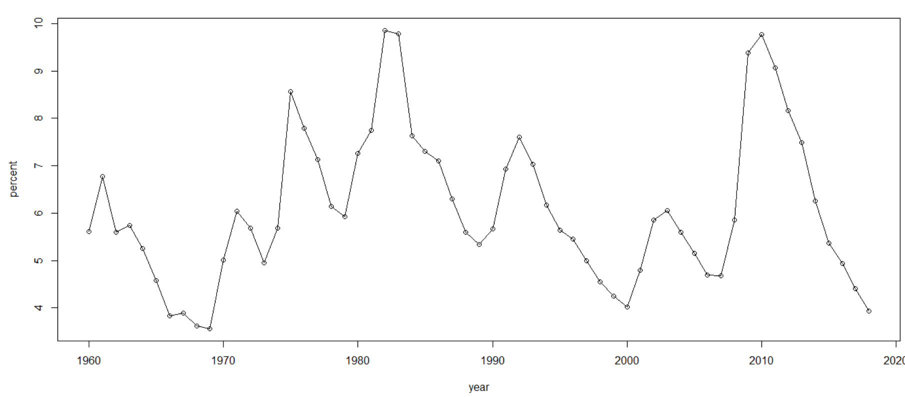

Figure 1: Annual Unemployment Rate in the US

According to the graph shown in Figure (1), over the 58-year period, there was not a significant upward or downward trend in the dataset. However, we can use some statistical tests such as the KPSS test (Kwiatkowski et al., 1992) to analyze the absence of trend and non-stationarity of the dataset more precisely. The null hypothesis of this test is that the dataset is stationary, and the alternative hypothesis is that the dataset is not stationary. In the significant level of $\alpha = 0.05$, the calculated p-value of the test is 0.112 that is greater than α. Therefore, the assumption of non-stationarity of the time series is rejected.

To identify the proper order of the ARMA model, one method is to analyze the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs (Figure (2)). In the sample ACF graph, relatively strong autocorrelations at lags 1 and 2 dominate. Nevertheless, again, there is a clear sign of the damped oscillatory. In the PACF graph, lags 1 and 2 have significant values and, by ignoring the partial autocorrelation in lag 11, it can be concluded that the partial autocorrelation is interrupted after two significant lags. From the ACF and PACF graphs, it seems that an AR(2) model should be selected for data modeling. On the other hand, in the ACF graph, after lag 2 and up to 13th lag, the sample autocorrelations do not show significant values, whereas the autocorrelations are significant in lags 13, 14, and 15. Therefore, it is better to use other criteria for a more accurate analysis of the model order.

One of the most widely used criteria for selecting the ARMA model order is Akaike information criterion (AIC) (Akaike et al., 1973).

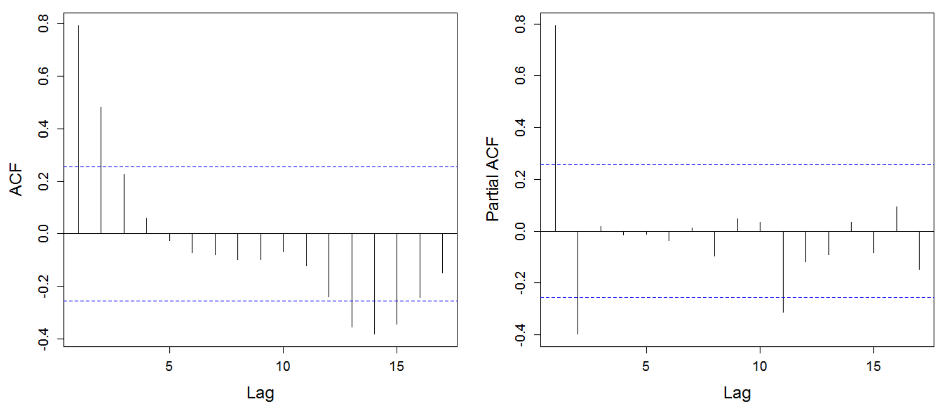

Figure 2: The sample ACF and PACF of unemployment rate series

This criterion states that the model which minimizes

$$\text{AIC} = -2 \log \text{likelihood} + 2k,$$

should be chosen, where $k = p + q + 1$ and p and q are the orders of the AR and MA parts, respectively. AICc is another criterion that has been introduced to overcome the AIC bias problem (Hurvich & Tsai, 1989). AICc is defined as follows:

$$\text{AICc} = -2 \log \text{likelihood} + 2k \frac{n}{n-k-1}.$$

The Schwarz Bayesian Information Criterion (BIC) (Schwarz, 1978) is another approach to determine the ARMA model order that is closely related to the AIC,

$$\text{BIC} = -2 \log \text{likelihood} + k \log(n).$$

According to Table (8), the minimum obtained values of the information criteria are for the first-order ARMA model and, with a very slight difference, to AR(2) model. Therefore, the first-order ARMA model is our selection for data modeling. The estimated parameters by the ML method and under the assumption of residual normality are $\hat{\phi} = 0.6782$, $\hat{\theta} = 0.5159$, and, consequently, the estimated model is $z_t = 0.6782z_{t-1} + \epsilon_t - 0.5159\epsilon_{t-1}$.

We use residual analysis to assess the fitness of the estimated model. An important diagnostic tool for examining the independence of the residuals is the sample ACF graph. By comparing the ACF values of residuals in Figure (3) with the standard error lines $\pm 2/\sqrt{n}$, the lack of meaningful autocorrolation in the estimated model residuals can be concluded. The density of the residuals can be assess by a histogram (Figure (4)). The histogram shows the non-normality and slight right skewness of residual density. For more confidence about the non-normality of the residuals, we use the Shapiro-Wilk normality test in the confidence level of 0.95 (Shapiro & Wilk, 1965). Regarding the p-value of the test (1.80×10^{-8}) the null hypothesis that the residuals are normally distributed is rejected.

Table 8: The estimated values of information criteria

ARMA(p,q)	AIC	AICc	BIC
ARMA(0,1)	1.118	1.159	0.188
ARMA(0,2)	0.892	0.938	-0.002
ARMA(1,0)	0.938	0.979	0.008
ARMA(1,1)	0.754	0.801	-0.139
ARMA(1,2)	0.786	0.839	-0.072
ARMA(2,0)	0.762	0.808	-0.132
ARMA(2,1)	0.779	0.832	-0.079
ARMA(2,2)	0.829	0.891	0.005

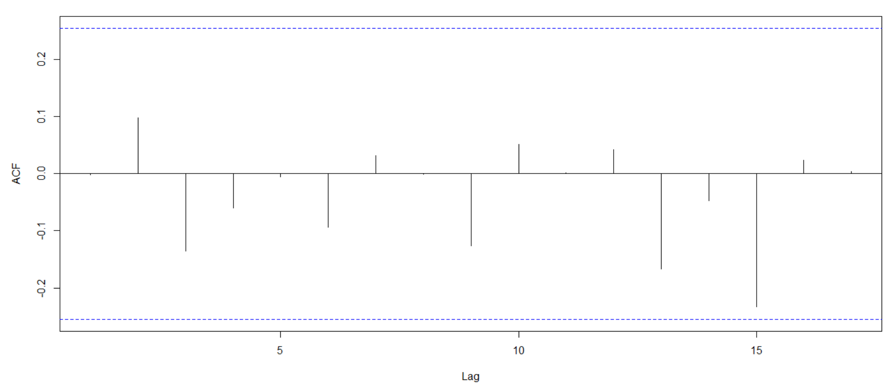

Figure 3: The sample ACF of residuals of the estimated first-order ARMA model

Now, by considering the positive domain of the residuals, we assign each of the distributions described in section 2 as candidate residual distributions and estimate

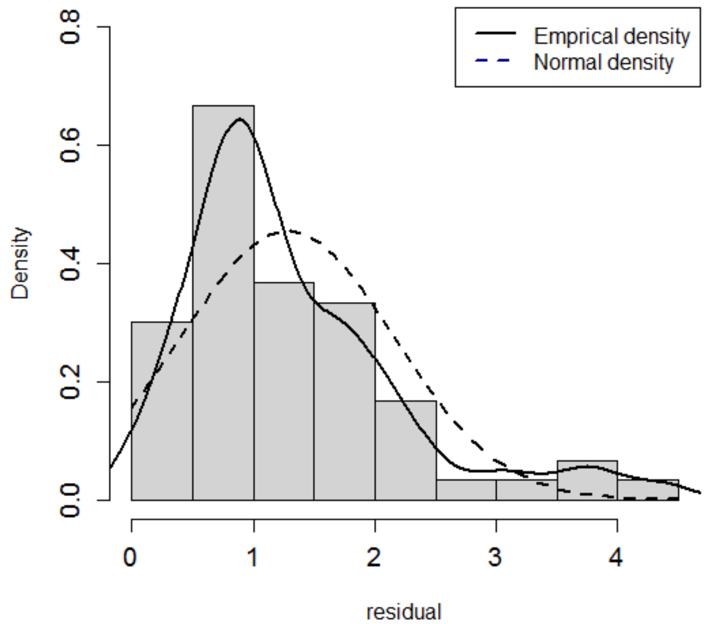

Figure 4: Histogram of residuals of the estimated first-order ARMA model

the model parameters by the recursive modified MLEs (Table (9)). By comparing the results of Table (9), it is observed that the estimated values of the first-order ARMA model coefficients based on different residual distributions are slightly different. Thus, we expect to obtian a similar result from the reduced method, which estimates the coefficients of the first-order ARMA model without considering the parameters of residual distribution. The estimated values of the first-order ARMA model coefficients using the reduced method are $\hat{\phi} = 0.6822$ and $\hat{\theta} = 0.4927$. These values are appriximately equal to the estimated values of the model coefficients with the recursive method based on different of residual the distribution. Therefore, it can be concluded that the reduced method with less computational volume than the recursive method is capable to estimate the model coefficients.

Table 9: The estimated values of the first-order ARMA model parameters by modified MLEs

Residual distribution	First parameter	Second parameter	φ	θ
$G(\alpha,\beta)$	2.654	0.472	0.6722	0.4925
$LN(M,S)$	0.027	0.664	0.6821	0.4824
$INVG(\tau,\varphi)$	1.255	2.126	0.6827	0.4923
$W(\kappa,\xi)$	1.638	1.410	0.6791	0.4931
$R(\sigma)$	-	1.063	0.6818	0.4891

Information criteria and the Kolmogorov-Smirnov (K-S) test are the methods that we use to select the appropriate model from among the estimated models to fit to unemployment rate data in the United States. Parameter K, defined in information criteria, is the number of the estimated parameters now. The lowest value of the calculated information criteria belongs to the gamma and then to the log-normal distributions (Table (10)). Furthermore, The null hypotheses that the residuals come from the estimated gamma and log-normal distributions are not rejected because the p-values of K-S test are greater than $\alpha = 0.05$.

Table 10: The estimated values of information criteria and p-values of K-S test

Residual distribution	AIC	AICc	BIC	p-value of K-S test
$G(\alpha, \beta)$	139	139.74	138.08	0.806
$LN(M, S)$	141	141.73	140.07	0.582
$INVG(\tau, \varphi)$	154	154.75	153.06	0.043
$W(\kappa, \xi)$	157	157.74	156.08	0.040
$R(\sigma)$	168	168.44	167.29	0.031

Therefore, model $z_t = 0.6722z_{t-1} + \epsilon_t - 0.4925\epsilon_{t-1}$ with residuals distributed of $G(2.654, 0.472)$ and model $z_t = 0.6821z_{t-1} + \epsilon_t - 0.4824\epsilon_{t-1}$ with residuals distributed of $LN(0.027, 0.664)$ are suitable for modeling the data of unemployment rate in the United States. Figure (5), reveals that the gamma density is the closest density to the residual empirical density compared with other estimated ones.

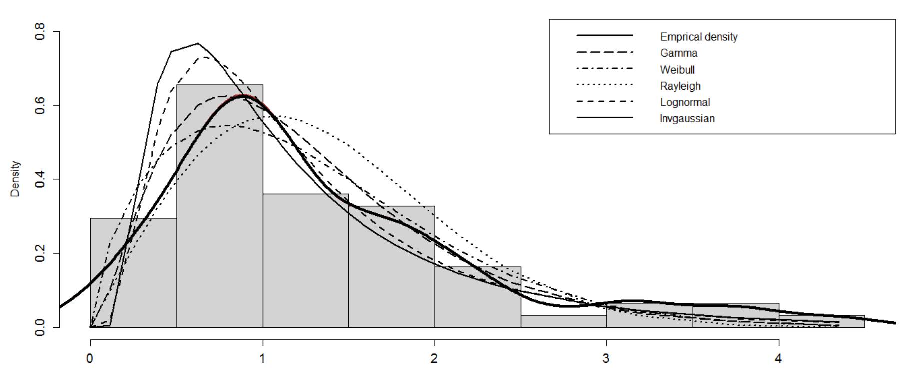

Figure 5: Histogram of the residuals of the estimated first-order ARMA model with curves of the estimated residual densities

Table 1.a: Modified MLEs for the first-order GARMA models

			$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = G(2,3)$							
φ	θ	n	$\hat{\alpha}_n$	MSE	MLCI	LCI	$\hat{\beta}_n$	MSE	MLCI	LCI
-0.5	-0.6	50	2.265	0.098	(1.88,2.08)	0.20	2.947	0.125	(2.88,3.25)	0.37
		150	2.043	0.0066	(1.91,2.09)	0.17	3.042	0.0086	(2.87,3.19)	0.32
		500	2.009	0.00087	(1.97,2.07)	0.10	3.043	0.0015	(2.88,3.05)	0.18
-0.5	-0.4	50	2.241	0.067	(1.91,2.12)	0.21	2.962	0.106	(2.82,3.18)	0.36
		150	2.019	0.0071	(1.91,2.09)	0.18	3.073	0.0098	(2.86,3.18)	0.32
		500	2.009	0.00091	(1.94,2.05)	0.11	3.045	0.0020	(2.91,3.09)	0.17
-0.5	0.3	50	2.246	0.078	(1.87,2.07)	0.20	2.989	0.129	(2.88,3.25)	0.37
		150	2.031	0.0063	(1.90,2.09)	0.19	3.064	0.0095	(2.88,3.20)	0.32
		500	2.029	0.00096	(1.95,2.06)	0.11	3.029	0.0018	(2.89,3.07)	0.18
-0.5	0.5	50	2.267	0.090	(1.88,2.09)	0.21	2.952	0.133	(2.86,3.23)	0.36
		150	2.039	0.0087	(1.95,2.14)	0.19	3.044	0.0099	(2.83,3.14)	0.31
		500	2.003	0.00093	(1.95,2.05)	0.10	3.046	0.0016	(2.91,3.08)	0.17
-0.1	-0.6	50	2.401	0.073	(1.94,2.15)	0.21	2.892	0.098	(2.80,3.15)	0.35
		150	2.108	0.0076	(1.93,2.11)	0.18	3.003	0.0090	(2.85,3.17)	0.32
		500	2.061	0.00088	(1.95,2.05)	0.10	3.025	0.0015	(2.91,3.08)	0.17
-0.1	-0.4	50	2.438	0.096	(1.89,2.10)	0.21	2.821	0.129	(2.86,3.23)	0.37
		150	2.070	0.0050	(1.92,2.10)	0.18	3.024	0.0078	(2.84,3.16)	0.32
		500	2.053	0.00090	(1.94,2.04)	0.10	3.000	0.0022	(2.92,3.10)	0.18
-0.1	0.3	50	2.469	0.085	(1.93,2.14)	0.21	2.791	0.114	(2.81,3.16)	0.35
		150	2.080	0.0083	(1.90,2.08)	0.18	3.011	0.0091	(2.89,3.21)	0.33
		500	2.03	0.00083	(1.95,2.05)	0.10	3.003	0.0016	(2.90,3.08)	0.18
-0.1	0.5	50	2.385	0.090	(1.92,2.13)	0.21	2.876	0.120	(2.82,3.18)	0.36
		150	2.051	0.0066	(1.90,2.08)	0.18	3.045	0.0096	(2.88,3.20)	0.32
		500	2.037	0.00093	(1.96,2.06)	0.10	3.030	0.0010	(2.89,3.07)	0.18
0.2	-0.6	50	2.470	0.091	(1.87,2.08)	0.21	2.849	0.124	(2.88,3.25)	0.37
		150	2.113	0.0069	(1.94,2.13)	0.18	3.031	0.0084	(2.83,3.14)	0.31
		500	2.063	0.00087	(1.94,2.05)	0.11	3.041	0.0024	(2.92,3.09)	0.17
0.2	-0.4	50	2.424	0.076	(1.90,2.11)	0.21	2.918	0.115	(2.85,3.22)	0.37
		150	2.090	0.0080	(1.97,2.16)	0.18	3.037	0.0086	(2.79,3.10)	0.31
		500	2.035	0.00092	(1.95,2.05)	0.10	3.061	0.0016	(2.91,3.10)	0.19
0.2	0.3	50	2.350	0.074	(1.93,2.14)	0.21	2.886	0.094	(2.80,3.16)	0.36
		150	2.060	0.0065	(1.94,2.13)	0.19	3.026	0.0085	(2.83,3.14)	0.31
		500	2.032	0.00075	(1.94,2.04)	0.10	3.033	0.0018	(2.91,3.10)	0.19
0.2	0.5	50	2.374	0.065	(1.89,2.10)	0.21	2.915	0.090	(2.84,3.20)	0.36
		150	2.060	0.0073	(1.86,2.05)	0.18	3.032	0.0095	(2.92,3.25)	0.33
		500	2.049	0.00072	(1.97,2.07)	0.10	3.020	0.0021	(2.91,3.07)	0.16
0.8	-0.6	50	2.499	0.096	(1.90,2.10)	0.20	2.810	0.143	(2.87,3.24)	0.37
		150	2.064	0.0083	(1.97,2.16)	0.18	3.015	0.0095	(2.80,3.11)	0.31
		500	2.053	0.00095	(1.95,2.05)	0.10	3.001	0.0017	(2.91,3.09)	0.17
0.8	-0.4	50	2.360	0.059	(1.87,2.08)	0.21	2.893	0.108	(2.87,3.24)	0.37
		150	2.080	0.0063	(1.95,2.13)	0.18	2.995	0.0079	(2.82,3.13)	0.31
		500	2.005	0.00093	(1.94,2.05)	0.11	3.036	0.0017	(2.92,3.10)	0.18
0.8	0.3	50	2.362	0.090	(1.87,2.08)	0.21	2.923	0.150	(2.89,3.26)	0.37
		150	2.207	0.0039	(1.88,2.06)	0.18	2.978	0.0068	(2.88,3.21)	0.32
		500	2.019	0.00082	(1.95,2.05)	0.10	2.995	0.0014	(2.90,3.09)	0.19
0.8	0.5	50	2.402	0.108	(1.95,2.16)	0.21	2.897	0.114	(2.80,3.15)	0.35
		150	2.061	0.0067	(1.91,2.09)	0.17	3.020	0.0090	(2.87,3.20)	0.32
		500	2.037	0.00091	(1.93,2.03)	0.10	3.010	0.0014	(2.94,3.12)	0.18

Table 1.b: Modified MLEs for the first-order GARMA models

			$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = G(2,3)$							
φ	θ	n	$\hat{\varphi}_{n,1}$	MSE	MLCI	LCI	$\hat{\theta}_{n,1}$	MSE	MLCI	LCI
-0.5	-0.6	50	-0.502	0.000040	(-0.51,-0.48)	0.03	-0.597	0.022	(-0.68,-0.50)	0.17
		150	-0.496	0.000037	(-0.51,-0.49)	0.02	-0.595	0.0071	(-0.63,-0.55)	0.08
		500	-0.499	0.000036	(-0.51,-0.49)	0.02	-0.598	0.0021	(-0.61,-0.57)	0.04
-0.5	-0.4	50	-0.497	0.000039	(-0.51,-0.48)	0.03	-0.397	0.022	(-0.47,-0.29)	0.18
		150	-0.496	0.000035	(-0.51,-0.48)	0.03	-0.396	0.0070	(-0.44,-0.35)	0.08
		500	-0.499	0.000032	(-0.51,-0.49)	0.02	-0.398	0.0022	(-0.42,-0.38)	0.04
-0.5	0.3	50	-0.495	0.000041	(-0.51,-0.48)	0.03	0.311	0.020	(0.22,0.39)	0.17
		150	-0.495	0.000038	(-0.51,-0.49)	0.02	0.304	0.0066	(0.26,0.35)	0.09
		500	-0.497	0.000038	(-0.51,-0.49)	0.02	0.305	0.0019	(0.28,0.32)	0.04
-0.5	0.5	50	-0.492	0.000036	(-0.51,-0.48)	0.03	0.512	0.020	(0.42,0.59)	0.17
		150	-0.495	0.000037	(-0.51,-0.49)	0.02	0.503	0.0068	(0.46,0.54)	0.08
		500	-0.502	0.000035	(-0.51,-0.49)	0.02	0.501	0.0017	(0.48,0.52)	0.04
-0.1	-0.6	50	-0.104	0.000048	(-0.11,-0.08)	0.03	-0.579	0.026	(-0.67,-0.49)	0.18
		150	-0.106	0.000046	(-0.11,-0.09)	0.02	-0.601	0.0072	(-0.63,-0.54)	0.09
		500	-0.105	0.000032	(-0.11,-0.09)	0.02	-0.601	0.0016	(-0.62,-0.57)	0.05
-0.1	-0.4	50	-0.104	0.000048	(-0.11,-0.08)	0.03	-0.382	0.027	(-0.46,-0.29)	0.17
		150	-0.103	0.000040	(-0.11,-0.08)	0.03	-0.402	0.0061	(-0.44,-0.35)	0.09
		500	-0.101	0.000038	(-0.11,-0.09)	0.02	-0.402	0.0020	(-0.42,-0.37)	0.05
-0.1	0.3	50	-0.095	0.000046	(-0.11,-0.08)	0.03	0.326	0.022	(0.23,0.41)	0.18
		150	-0.095	0.000041	(-0.11,-0.09)	0.02	0.307	0.0073	(0.26,0.34)	0.08
		500	-0.097	0.000035	(-0.11,-0.09)	0.02	0.304	0.0022	(0.28,0.32)	0.04
-0.1	0.5	50	-0.093	0.000037	(-0.12,-0.09)	0.03	0.523	0.025	(0.43,0.60)	0.17
		150	-0.097	0.000035	(-0.11,-0.09)	0.02	0.503	0.0066	(0.46,0.55)	0.09
		500	-0.096	0.000032	(-0.11,-0.09)	0.02	0.506	0.0025	(0.48,0.52)	0.04
0.2	-0.6	50	0.191	0.000042	(0.18,0.21)	0.03	-0.585	0.024	(-0.68,-0.50)	0.18
		150	0.193	0.000039	(0.18,0.21)	0.03	-0.603	0.0064	(-0.63,-0.54)	0.09
		500	0.194	0.000031	(0.19,0.21)	0.02	-0.604	0.0018	(-0.62,-0.58)	0.04
0.2	-0.4	50	0.191	0.000038	(0.18,0.21)	0.03	-0.384	0.026	(-0.47,-0.28)	0.19
		150	0.193	0.000031	(0.19,0.21)	0.02	-0.403	0.0068	(-0.43,-0.35)	0.08
		500	0.198	0.000033	(0.19,0.21)	0.02	-0.406	0.0021	(-0.41,-0.37)	0.04
0.2	0.3	50	0.201	0.000036	(0.18,0.21)	0.03	0.320	0.026	(0.23,0.41)	0.18
		150	0.200	0.000033	(0.19,0.21)	0.02	0.305	0.0069	(0.26,0.34)	0.08
		500	0.201	0.000031	(0.19,0.21)	0.02	0.303	0.0018	(0.28,0.32)	0.04
0.2	0.5	50	0.202	0.000041	(0.18,0.21)	0.03	0.525	0.024	(0.43,0.61)	0.18
		150	0.201	0.000039	(0.19,0.21)	0.02	0.502	0.0066	(0.46,0.55)	0.09
		500	0.201	0.000036	(0.19,0.21)	0.02	0.511	0.0019	(0.48,0.52)	0.04
0.8	-0.6	50	0.798	0.000026	(0.793,0.805)	0.012	-0.553	0.025	(-0.65,-0.46)	0.19
		150	0.799	0.000021	(0.793,0.805)	0.012	-0.595	0.0069	(-0.63,-0.54)	0.09
		500	0.799	0.000021	(0.795,0.805)	0.010	-0.599	0.0018	(-0.62,-0.58)	0.04
0.8	-0.4	50	0.798	0.000026	(0.792,0.805)	0.013	-0.369	0.023	(-0.45,-0.27)	0.18
		150	0.799	0.000022	(0.793,0.805)	0.012	-0.392	0.0071	(-0.43,-0.34)	0.09
		500	0.800	0.000020	(0.794,0.805)	0.011	-0.400	0.0023	(-0.42,-0.37)	0.05
0.8	0.3	50	0.786	0.000026	(0.792,0.804)	0.012	0.292	0.034	(0.24,0.43)	0.19
		150	0.791	0.000020	(0.794,0.805)	0.011	0.294	0.0065	(0.26,0.35)	0.09
		500	0.796	0.000018	(0.794,0.805)	0.011	0.294	0.0018	(0.28,0.32)	0.04
0.8	0.5	50	0.798	0.000024	(0.793,0.804)	0.011	0.531	0.023	(0.44,0.62)	0.18
		150	0.799	0.000020	(0.794,0.805)	0.011	0.505	0.0064	(0.47,0.55)	0.08
		500	0.8001	0.000019	(0.794,0.805)	0.011	0.503	0.0020	(0.48,0.52)	0.04

Table 2.a: Modified MLEs for the first-order LNARMA models

$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = LN(2, 0.5)$										
φ	θ	n	$\hat{M}_n$	MSE	MLCI	LCI	$\hat{S}_n$	MSE	MLCI	LCI
-0.5	-0.6	50	1.969	0.066	(1.83,2.12)	0.29	0.532	0.029	(0.29,0.77)	0.48
		150	1.990	0.016	(1.91,2.07)	0.16	0.511	0.0071	(0.39,0.63)	0.24
		500	1.994	0.0050	(1.94,2.03)	0.09	0.504	0.0018	(0.44,0.56)	0.12
-0.5	-0.4	50	1.979	0.068	(1.83,2.12)	0.29	0.533	0.024	(0.28,0.78)	0.50
		150	1.998	0.014	(1.91,2.07)	0.16	0.506	0.0062	(0.39,0.63)	0.24
		500	2.001	0.0050	(1.95,2.04)	0.09	0.502	0.0016	(0.44,0.56)	0.12
-0.5	0.3	50	1.979	0.061	(1.82,2.11)	0.29	0.530	0.030	(0.28,0.77)	0.49
		150	1.997	0.015	(1.91,2.07)	0.16	0.505	0.0060	(0.39,0.62)	0.23
		500	2.002	0.0053	(1.95,2.04)	0.09	0.501	0.0017	(0.44,0.56)	0.12
-0.5	0.5	50	1.972	0.064	(1.83,2.13)	0.30	0.532	0.032	(0.27,0.78)	0.51
		150	1.997	0.014	(1.91,2.07)	0.16	0.504	0.0034	(0.39,0.62)	0.23
		500	1.998	0.0053	(1.95,2.04)	0.09	0.503	0.0056	(0.44,0.56)	0.12
-0.1	-0.6	50	1.988	0.055	(1.85,2.13)	0.28	0.518	0.023	(0.29,0.77)	0.48
		150	1.987	0.015	(1.92,2.08)	0.16	0.513	0.0062	(0.39,0.62)	0.23
		500	1.989	0.0049	(1.95,2.04)	0.09	0.510	0.0018	(0.44,0.57)	0.13
-0.1	-0.4	50	1.983	0.046	(1.83,2.13)	0.30	0.517	0.023	(0.28,0.77)	0.49
		150	1.986	0.013	(1.92,2.08)	0.16	0.513	0.0054	(0.38,0.62)	0.24
		500	2.001	0.0054	(1.95,2.04)	0.09	0.510	0.0020	(0.44,0.56)	0.12
-0.1	0.3	50	1.997	0.061	(1.84,2.13)	0.29	0.511	0.029	(0.28,0.76)	0.48
		150	1.998	0.015	(1.91,2.07)	0.16	0.505	0.0062	(0.39,0.62)	0.23
		500	1.995	0.0053	(1.95,2.04)	0.09	0.503	0.0021	(0.44,0.56)	0.12
-0.1	0.5	50	2.000	0.054	(1.85,2.13)	0.28	0.509	0.020	(0.26,0.77)	0.51
		150	1.992	0.015	(1.91,2.07)	0.16	0.510	0.0052	(0.39,0.62)	0.23
		500	2.001	0.0050	(1.95,2.04)	0.09	0.507	0.0017	(0.44,0.57)	0.13
0.2	-0.6	50	2.007	0.059	(1.85,2.14)	0.29	0.516	0.027	(0.27,0.75)	0.48
		150	2.000	0.015	(1.92,2.08)	0.16	0.504	0.0060	(0.39,0.62)	0.23
		500	2.000	0.0050	(1.95,2.04)	0.09	0.501	0.0017	(0.44,0.56)	0.12
0.2	-0.4	50	2.002	0.058	(1.85,2.14)	0.29	0.517	0.026	(0.28,0.76)	0.48
		150	2.000	0.016	(1.92,2.08)	0.16	0.503	0.0058	(0.38,0.62)	0.24
		500	1.999	0.0048	(1.96,2.04)	0.08	0.503	0.0016	(0.44,0.57)	0.13
0.2	0.3	50	1.999	0.063	(1.84,2.13)	0.29	0.507	0.018	(0.25,0.77)	0.52
		150	1.990	0.016	(1.92,2.08)	0.16	0.511	0.0069	(0.39,0.62)	0.23
		500	2.001	0.0050	(1.95,2.04)	0.09	0.506	0.0022	(0.44,0.56)	0.12
0.2	0.5	50	1.993	0.060	(1.86,2.14)	0.28	0.523	0.027	(0.26,0.76)	0.50
		150	1.997	0.014	(1.91,2.08)	0.17	0.506	0.0053	(0.39,0.63)	0.24
		500	2.000	0.0051	(1.95,2.03)	0.08	0.502	0.0019	(0.44,0.57)	0.13
0.8	-0.6	50	2.009	0.062	(1.85,2.14)	0.29	0.509	0.025	(0.28,0.75)	0.47
		150	1.994	0.014	(1.93,2.08)	0.15	0.507	0.0056	(0.38,0.62)	0.24
		500	2.000	0.0049	(1.95,2.04)	0.09	0.503	0.0017	(0.44,0.56)	0.12
0.8	-0.4	50	2.024	0.059	(1.88,2.16)	0.28	0.506	0.024	(0.28,0.76)	0.48
		150	2.001	0.015	(1.91,2.07)	0.16	0.503	0.0060	(0.39,0.62)	0.23
		500	2.002	0.0050	(1.96,2.04)	0.08	0.501	0.0018	(0.44,0.57)	0.13
0.8	0.3	50	2.010	0.057	(1.87,2.15)	0.28	0.509	0.025	(0.27,0.77)	0.50
		150	2.002	0.015	(1.93,2.09)	0.16	0.503	0.0059	(0.39,0.62)	0.23
		500	2.002	0.0052	(1.96,2.04)	0.08	0.502	0.0019	(0.44,0.56)	0.12
0.8	0.5	50	2.007	0.056	(1.87,2.15)	0.28	0.517	0.026	(0.27,0.77)	0.50
		150	1.999	0.016	(1.93,2.08)	0.15	0.502	0.0052	(0.38,0.62)	0.24
		500	2.003	0.0056	(1.95,2.04)	0.09	0.501	0.0017	(0.43,0.56)	0.13

Table 2.b: Modified MLEs for the first-order LNARMA models

			$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = LN(2, 0.5)$							
φ	θ	n	$\hat{\varphi}_{n,2}$	MSE	MLCI	LCI	$\hat{\theta}_{n,2}$	MSE	MLCI	LCI
-0.5	-0.6	50	-0.500	0.000021	(-0.509,-0.490)	0.018	-0.598	0.028	(-0.73,-0.45)	0.28
		150	-0.499	0.000016	(-0.508,-0.491)	0.017	-0.603	0.0065	(-0.68,-0.51)	0.17
		500	-0.499	0.000014	(-0.508,-0.491)	0.017	-0.600	0.0019	(-0.65,-0.54)	0.11
-0.5	-0.4	50	-0.500	0.000024	(-0.508,-0.491)	0.017	-0.405	0.026	(-0.54,-0.26)	0.28
		150	-0.500	0.000022	(-0.508,-0.491)	0.017	-0.398	0.0067	(-0.48,-0.30)	0.18
		500	-0.500	0.000020	(-0.508,-0.492)	0.016	-0.396	0.0018	(-0.44,-0.34)	0.10
-0.5	0.3	50	-0.500	0.000022	(-0.508,-0.491)	0.017	0.305	0.027	(0.18,0.44)	0.26
		150	-0.500	0.000018	(-0.508,-0.491)	0.017	0.302	0.0067	(0.22,0.39)	0.17
		500	-0.499	0.000020	(-0.507,-0.492)	0.015	0.302	0.0020	(0.25,0.35)	0.10
-0.5	0.5	50	-0.500	0.000023	(-0.508,-0.491)	0.017	0.502	0.027	(0.37,0.63)	0.26
		150	-0.499	0.000018	(-0.508,-0.491)	0.017	0.502	0.0063	(0.41,0.59)	0.18
		500	-0.499	0.000017	(-0.507,-0.492)	0.015	0.502	0.0019	(0.44,0.54)	0.10
-0.1	-0.6	50	-0.099	0.000032	(-0.117,-0.082)	0.035	-0.577	0.019	(-0.72,-0.44)	0.28
		150	-0.101	0.000026	(-0.118,-0.081)	0.037	-0.598	0.0063	(-0.68,-0.51)	0.17
		500	-0.101	0.000024	(-0.115,-0.085)	0.030	-0.596	0.0027	(-0.65,-0.55)	0.10
-0.1	-0.4	50	-0.1003	0.000034	(-0.116,-0.083)	0.033	-0.388	0.020	(-0.52,-0.24)	0.28
		150	-0.1002	0.000030	(-0.118,-0.081)	0.037	-0.398	0.0064	(-0.48,-0.30)	0.18
		500	-0.1005	0.000026	(-0.114,-0.085)	0.029	-0.399	0.0028	(-0.45,-0.35)	0.10
-0.1	0.3	50	-0.0997	0.000027	(-0.115,-0.084)	0.031	0.319	0.019	(0.18,0.45)	0.27
		150	-0.1004	0.000027	(-0.118,-0.081)	0.037	0.305	0.0067	(0.20,0.39)	0.19
		500	-0.1001	0.000023	(-0.115,-0.084)	0.031	0.299	0.0012	(0.25,0.35)	0.10
-0.1	0.5	50	-0.099	0.000033	(-0.114,-0.085)	0.029	0.521	0.020	(0.37,0.66)	0.29
		150	-0.098	0.000029	(-0.117,-0.082)	0.035	0.503	0.0067	(0.42,0.59)	0.17
		500	-0.1000	0.000025	(-0.114,-0.086)	0.028	0.500	0.0026	(0.45,0.55)	0.10
0.2	-0.6	50	0.199	0.000028	(0.188,0.211)	0.023	-0.570	0.027	(-0.71,-0.43)	0.28
		150	0.199	0.000027	(0.187,0.212)	0.025	-0.596	0.0066	(-0.67,-0.50)	0.17
		500	0.199	0.000025	(0.188,0.211)	0.023	-0.596	0.0019	(-0.65,-0.55)	0.10
0.2	-0.4	50	0.199	0.000028	(0.188,0.211)	0.023	-0.379	0.0265	(-0.51,-0.23)	0.28
		150	0.199	0.000026	(0.186,0.213)	0.027	-0.396	0.0064	(-0.47,-0.30)	0.17
		500	0.199	0.000028	(0.187,0.212)	0.025	-0.397	0.0019	(-0.44,-0.34)	0.10
0.2	0.3	50	0.200	0.000029	(0.185,0.214)	0.029	0.324	0.020	(0.17,0.46)	0.29
		150	0.199	0.000030	(0.187,0.212)	0.025	0.301	0.0076	(0.23,0.39)	0.16
		500	0.200	0.000027	(0.189,0.210)	0.021	0.303	0.0027	(0.24,0.35)	0.11
0.2	0.5	50	0.199	0.000027	(0.185,0.214)	0.029	0.515	0.028	(0.39,0.65)	0.26
		150	0.199	0.000026	(0.187,0.212)	0.025	0.503	0.0066	(0.42,0.59)	0.17
		500	0.200	0.000024	(0.188,0.211)	0.023	0.501	0.0020	(0.45,0.55)	0.10
0.8	-0.6	50	0.798	0.000016	(0.795,0.805)	0.010	-0.568	0.029	(-0.70,-0.43)	0.27
		150	0.799	0.000013	(0.795,0.804)	0.009	-0.598	0.0068	(-0.67,-0.50)	0.17
		500	0.799	0.000011	(0.795,0.804)	0.009	-0.601	0.0020	(-0.64,-0.54)	0.10
0.8	-0.4	50	0.799	0.000015	(0.795,0.806)	0.011	-0.360	0.029	(-0.50,-0.22)	0.28
		150	0.799	0.000011	(0.795,0.804)	0.009	-0.394	0.0071	(-0.48,-0.30)	0.18
		500	0.799	0.000012	(0.795,0.804)	0.009	-0.394	0.0020	(-0.44,-0.34)	0.10
0.8	0.3	50	0.799	0.000017	(0.795,0.804)	0.009	0.333	0.029	(0.20,0.46)	0.26
		150	0.799	0.000012	(0.795,0.803)	0.008	0.306	0.0040	(0.22,0.39)	0.17
		500	0.799	0.000010	(0.795,0.804)	0.009	0.303	0.0018	(0.25,0.35)	0.10
0.8	0.5	50	0.799	0.000016	(0.795,0.804)	0.009	0.528	0.026	(0.39,0.66)	0.27
		150	0.800	0.000010	(0.795,0.804)	0.009	0.505	0.0074	(0.43,0.59)	0.16
		500	0.800	0.000012	(0.796,0.805)	0.009	0.505	0.0019	(0.45,0.55)	0.10

Table 3.a: Modified MLEs for the first-order INVGARMA models

$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = INVG(1,3)$										
φ	θ	n	$\hat{\tau}_n$	MSE	MLCI	LCI	$\hat{\vartheta}_n$	MSE	MLCI	LCI
-0.5	-0.6	50	1.005	0.032	(0.84,1.17)	0.32	2.923	0.57	(1.83,4.19)	2.36
		150	0.998	0.0064	(0.91,1.09)	0.18	2.979	0.16	(2.32,3.67)	1.35
		500	1.001	0.0030	(0.95,1.05)	0.10	3.003	0.041	(2.62,3.37)	0.74
-0.5	-0.4	50	1.001	0.038	(0.83,1.15)	0.32	2.939	0.52	(1.79,4.11)	2.31
		150	1.003	0.0065	(0.90,1.08)	0.18	2.959	0.15	(2.31,3.66)	1.35
		500	0.999	0.0033	(0.95,1.05)	0.10	2.991	0.045	(2.62,3.36)	0.74
-0.5	0.3	50	0.997	0.033	(0.84,1.16)	0.32	2.932	0.56	(1.81,4.16)	2.34
		150	1.003	0.0059	(0.91,1.09)	0.18	2.994	0.11	(2.30,3.65)	1.35
		500	0.997	0.0032	(0.95,1.05)	0.10	2.985	0.043	(2.62,3.36)	0.74
-0.5	0.5	50	1.001	0.036	(0.84,1.17)	0.33	2.877	0.55	(1.78,4.08)	2.30
		150	1.005	0.0060	(0.90,1.09)	0.19	2.973	0.15	(2.31,3.67)	1.36
		500	1.001	0.0032	(0.94,1.05)	0.11	2.985	0.041	(2.61,3.35)	0.74
-0.1	-0.6	50	1.013	0.0352	(0.85,1.18)	0.33	2.965	0.52	(1.81,4.15)	2.34
		150	1.003	0.0059	(0.90,1.09)	0.19	3.003	0.14	(2.32,3.69)	1.37
		500	1.001	0.0029	(0.95,1.05)	0.10	3.010	0.052	(2.61,3.35)	0.74
-0.1	-0.4	50	1.000	0.034	(0.84,1.16)	0.32	2.995	0.54	(1.80,4.13)	2.33
		150	1.007	0.0061	(0.91,1.10)	0.19	2.997	0.14	(2.32,3.69)	1.37
		500	1.003	0.0033	(0.95,1.05)	0.10	2.998	0.042	(2.63,3.37)	0.74
-0.1	0.3	50	1.015	0.035	(0.83,1.15)	0.32	2.990	0.56	(1.82,4.17)	2.35
		150	1.004	0.0064	(0.90,1.09)	0.19	2.981	0.13	(2.30,3.65)	1.35
		500	1.003	0.0033	(0.95,1.05)	0.10	2.992	0.042	(2.62,3.37)	0.75
-0.1	0.5	50	1.020	0.033	(0.84,1.16)	0.32	3.011	0.60	(1.80,4.14)	2.34
		150	1.005	0.0060	(0.91,1.10)	0.19	3.013	0.13	(2.31,3.66)	1.35
		500	0.999	0.0033	(0.96,1.05)	0.09	3.004	0.047	(2.62,3.36)	0.74
0.2	-0.6	50	1.021	0.033	(0.84,1.16)	0.32	3.040	0.52	(1.80,4.13)	2.33
		150	1.002	0.0060	(0.91,1.09)	0.18	3.016	0.14	(2.32,3.68)	1.36
		500	1.004	0.0034	(0.95,1.05)	0.10	3.002	0.042	(2.63,3.37)	0.74
0.2	-0.4	50	1.017	0.037	(0.85,1.17)	0.32	2.954	0.57	(1.81,4.15)	2.34
		150	1.003	0.0056	(0.91,1.09)	0.18	2.992	0.14	(2.31,3.66)	1.35
		500	1.003	0.0033	(0.95,1.04)	0.09	2.996	0.040	(2.62,3.35)	0.73
0.2	0.3	50	1.022	0.0367	(0.85,1.17)	0.32	2.968	0.54	(1.80,4.13)	2.33
		150	1.000	0.0061	(0.91,1.10)	0.19	2.988	0.13	(2.32,3.68)	1.36
		500	1.002	0.0033	(0.95,1.05)	0.10	2.987	0.040	(2.62,3.37)	0.75
0.2	0.5	50	1.023	0.035	(0.84,1.17)	0.33	2.978	0.51	(1.80,4.12)	2.32
		150	1.002	0.0060	(0.90,1.09)	0.19	2.997	0.13	(2.30,3.65)	1.35
		500	0.999	0.0033	(0.95,1.05)	0.10	3.002	0.041	(2.62,3.36)	0.74
0.8	-0.6	50	1.043	0.036	(0.85,1.18)	0.33	3.178	0.58	(1.81,4.16)	2.35
		150	1.006	0.0064	(0.91,1.09)	0.18	2.977	0.12	(2.32,3.68)	1.36
		500	1.002	0.0033	(0.95,1.05)	0.10	2.999	0.044	(2.62,3.37)	0.75
0.8	-0.4	50	1.038	0.035	(0.84,1.16)	0.32	3.119	0.53	(1.80,4.14)	2.34
		150	1.003	0.0064	(0.91,1.09)	0.18	2.973	0.13	(2.33,3.70)	1.37
		500	1.003	0.0032	(0.95,1.05)	0.10	2.990	0.040	(2.62,3.36)	0.74
0.8	0.3	50	1.031	0.034	(0.85,1.17)	0.32	3.016	0.49	(1.78,4.08)	2.30
		150	1.002	0.0056	(0.91,1.09)	0.18	2.986	0.14	(2.31,3.67)	1.36
		500	1.005	0.0032	(0.95,1.05)	0.10	2.983	0.042	(2.61,3.35)	0.74
0.8	0.5	50	1.033	0.032	(0.84,1.17)	0.33	3.040	0.55	(1.80,4.12)	2.32
		150	1.007	0.0062	(0.90,1.09)	0.19	2.950	0.14	(2.31,3.67)	1.36
		500	1.002	0.0030	(0.95,1.05)	0.10	2.966	0.046	(2.62,3.37)	0.75

Table 3.b: Modified MLEs for the first-order INVGARMA models

$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = INVG(1,3)$										
φ	θ	n	$\hat{\varphi}_{n,3}$	MSE	MLCI	LCI	$\hat{\theta}_{n,3}$	MSE	MLCI	LCI
-0.5	-0.6	50	-0.499	0.000023	(-0.511,-0.48)	0.022	-0.593	0.026	(-0.67,-0.50)	0.17
		150	-0.499	0.000029	(-0.509,-0.490)	0.018	-0.600	0.0041	(-0.64,-0.56)	0.08
		500	-0.499	0.000023	(-0.509,-0.490)	0.018	-0.598	0.0017	(-0.62,-0.57)	0.05
-0.5	-0.4	50	-0.500	0.000026	(-0.511,-0.488)	0.022	-0.402	0.030	|(-0.47,-0.30)	0.17
		150	-0.499	0.000024	(-0.509,-0.490)	0.019	-0.396	0.0038	(-0.44,-0.36)	0.08
		500	-0.500	0.000027	(-0.509,-0.490)	0.019	-0.401	0.0019	(-0.42,-0.38)	0.04
-0.5	0.3	50	-0.500	0.000024	(-0.511,-0.489)	0.022	0.300	0.026	(0.22,0.39)	0.17
		150	-0.499	0.000023	(-0.509,-0.490)	0.019	0.302	0.0039	(0.26,0.34)	0.08
		500	-0.499	0.000024	(-0.508,-0.491)	0.017	0.299	0.0018	(0.28,0.32)	0.04
-0.5	0.5	50	-0.500	0.000028	(-0.510,-0.489)	0.020	0.500	0.028	(0.42,0.58)	0.16
		150	-0.499	0.000022	(-0.509,-0.490)	0.019	0.505	0.0039	(0.46,0.54)	0.08
		500	-0.500	0.000019	(-0.508,-0.491)	0.017	0.501	0.0019	(0.47,0.52)	0.05
-0.1	-0.6	50	-0.100	0.000036	(-0.119,-0.080)	0.039	-0.585	0.028	(-0.67,-0.50)	0.17
		150	-0.100	0.000034	(-0.121,-0.078)	0.043	-0.596	0.0037	(-0.63,-0.55)	0.08
		500	-0.100	0.000029	(-0.115,-0.084)	0.030	-0.600	0.0018	(-0.62,-0.58)	0.04
-0.1	-0.4	50	-0.100	0.000039	(-0.120,-0.079)	0.041	-0.396	0.027	(-0.47,-0.30)	0.17
		150	-0.100	0.000030	(-0.119,-0.080)	0.038	-0.392	0.0041	(-0.43,-0.35)	0.08
		500	-0.100	0.000022	(-0.116,-0.083)	0.033	-0.396	0.0019	(-0.42,-0.37)	0.05
-0.1	0.3	50	-0.100	0.000043	(-0.117,-0.082)	0.034	0.314	0.028	(0.22,0.38)	0.16
		150	-0.100	0.000032	(-0.118,-0.081)	0.037	0.302	0.0041	(0.26,0.35)	0.09
		500	-0.099	0.000031	(-0.115,-0.084)	0.030	0.304	0.0019	(0.28,0.32)	0.04
-0.1	0.5	50	-0.100	0.000033	(-0.119,-0.080)	0.039	0.517	0.025	(0.43,0.60)	0.17
		150	-0.100	0.000031	(-0.117,-0.083)	0.033	0.505	0.0038	(0.46,0.54)	0.09
		500	-0.100	0.000029	(-0.114,-0.085)	0.028	0.502	0.0020	(0.48,0.52)	0.04
0.2	-0.6	50	0.199	0.000036	(0.179,0.220)	0.041	-0.577	0.024	(-0.66,-0.48)	0.18
		150	0.200	0.000035	(0.184,0.215)	0.031	-0.596	0.0039	(-0.63,-0.55)	0.08
		500	0.200	0.000031	(0.186,0.213)	0.026	-0.596	0.0020	(-0.62,-0.57)	0.05
0.2	-0.4	50	0.199	0.000034	(0.180,0.219)	0.039	-0.382	0.029	(-0.46,-0.27)	0.19
		150	0.199	0.000031	(0.182,0.217)	0.034	-0.396	0.0038	(-0.43,-0.34)	0.09
		500	0.199	0.000029	(0.186,0.213)	0.026	-0.396	0.0019	(-0.41,-0.37)	0.04
0.2	0.3	50	0.199	0.000036	(0.181,0.218)	0.037	0.322	0.029	(0.23,0.41)	0.18
		150	0.200	0.000032	(0.183,0.216)	0.032	0.302	0.0040	(0.26,0.35)	0.09
		500	0.199	0.000030	(0.185,0.214)	0.028	0.302	0.0020	(0.28,0.32)	0.04
0.2	0.5	50	0.199	0.000033	(0.184,0.215)	0.030	0.520	0.026	(0.42,0.59)	0.17
		150	0.200	0.000032	(0.183,0.216)	0.032	0.504	0.0040	(0.46,0.55)	0.09
		500	0.200	0.000033	(0.186,0.213)	0.027	0.501	0.0020	(0.48,0.52)	0.04
0.8	-0.6	50	0.798	0.000019	(0.792,0.807)	0.014	-0.562	0.027	(-0.67,-0.47)	0.20
		150	0.799	0.000014	(0.793,0.806)	0.012	-0.595	0.0044	(-0.63,-0.54)	0.09
		500	0.799	0.000013	(0.794,0.805)	0.011	-0.596	0.0019	(-0.61,-0.57)	0.04
0.8	-0.4	50	0.798	0.000018	(0.791,0.808)	0.016	-0.363	0.027	(-0.47,-0.26)	0.21
		150	0.799	0.000014	(0.793,0.806)	0.012	-0.395	0.0039	(-0.43,-0.33)	0.10
		500	0.799	0.000014	(0.794,0.805)	0.010	-0.396	0.0020	(-0.42,-0.37)	0.05
0.8	0.3	50	0.799	0.000016	(0.792,0.807)	0.014	0.331	0.026	(0.22,0.44)	0.22
		150	0.799	0.000012	(0.794,0.805)	0.011	0.304	0.0038	(0.26,0.36)	0.10
		500	0.800	0.000013	(0.795,0.804)	0.009	0.305	0.0020	(0.28,0.32)	0.04
0.8	0.5	50	0.798	0.000018	(0.793,0.806)	0.013	0.533	0.025	(0.43,0.63)	0.20
		150	0.799	0.000013	(0.794,0.805)	0.011	0.506	0.0039	(0.46,0.55)	0.09
		500	0.799	0.000012	(0.795,0.804)	0.009	0.503	0.0018	(0.48,0.52)	0.04

Table 4.a: Modified MLEs for the first-order WARMA models

$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = W(1.5, 2)$										
φ	θ	n	$\hat{\kappa}_n$	MSE	MLCI	LCI	$\hat{\xi}_n$	MSE	MLCI	LCI
-0.5	-0.6	50	1.605	0.132	(1.28,1.95)	0.67	2.038	0.154	(1.70,2.42)	0.72
		150	1.520	0.039	(1.34,1.71)	0.37	2.008	0.048	(1.80,2.22)	0.42
		500	1.508	0.0014	(1.41,1.60)	0.19	2.004	0.0019	(1.88,2.12)	0.24
-0.5	-0.4	50	1.621	0.135	(1.27,1.94)	0.67	2.062	0.147	(1.69,2.41)	0.72
		150	1.515	0.044	(1.36,1.72)	0.36	2.008	0.057	(1.81,2.23)	0.42
		500	1.498	0.0014	(1.41,1.60)	0.19	1.991	0.0020	(1.88,2.11)	0.23
-0.5	0.3	50	1.603	0.134	(1.28,1.95)	0.66	2.057	0.155	(1.69,2.41)	0.72
		150	1.515	0.037	(1.35,1.71)	0.36	2.005	0.047	(1.80,2.22)	0.42
		500	1.511	0.0013	(1.41,1.60)	0.19	2.009	0.0018	(1.88,2.12)	0.24
-0.5	0.5	50	1.601	0.127	(1.27,1.93)	0.66	2.048	0.154	(1.68,2.40)	0.72
		150	1.524	0.042	(1.35,1.71)	0.36	2.012	0.055	(1.80,2.22)	0.42
		500	1.508	0.0013	(1.41,1.60)	0.19	2.002	0.0017	(1.89,2.12)	0.23
-0.1	-0.6	50	1.614	0.135	(1.28,1.94)	0.66	2.058	0.159	(1.69,2.43)	0.74
		150	1.506	0.039	(1.34,1.70)	0.36	1.995	0.053	(1.79,2.22)	0.43
		500	1.508	0.0015	(1.40,1.60)	0.20	2.004	0.0021	(1.88,2.12)	0.24
-0.1	-0.4	50	1.624	0.126	(1.26,1.91)	0.65	2.056	0.160	(1.67,2.41)	0.74
		150	1.524	0.040	(1.35,1.71)	0.36	2.013	0.055	(1.80,2.23)	0.43
		500	1.508	0.0016	(1.41,1.61)	0.20	2.004	0.0022	(1.89,2.12)	0.23
-0.1	0.3	50	1.605	0.126	(1.28,1.94)	0.66	2.040	0.176	(1.70,2.43)	0.73
		150	1.523	0.039	(1.34,1.71)	0.37	2.011	0.055	(1.80,2.23)	0.43
		500	1.513	0.0017	(1.41,1.61)	0.20	2.005	0.0021	(1.89,2.12)	0.23
-0.1	0.5	50	1.609	0.131	(1.27,1.93)	0.66	2.019	0.163	(1.68,2.41)	0.73
		150	1.524	0.039	(1.34,1.71)	0.37	2.009	0.050	(1.80,2.23)	0.43
		500	1.508	0.0013	(1.41,1.60)	0.19	2.001	0.0018	(1.88,2.12)	0.24
0.2	-0.6	50	1.615	0.122	(1.26,1.91)	0.65	2.054	0.160	(1.67,2.40)	0.73
		150	1.524	0.040	(1.34,1.70)	0.36	2.013	0.053	(1.80,2.23)	0.43
		500	1.508	0.0015	(1.40,1.60)	0.20	2.002	0.0020	(1.88,2.12)	0.24
0.2	-0.4	50	1.593	0.132	(1.28,1.94)	0.66	2.032	0.149	(1.67,2.40)	0.73
		150	1.515	0.041	(1.35,1.71)	0.36	2.003	0.050	(1.79,2.22)	0.43
		500	1.505	0.0013	(1.41,1.61)	0.20	2.001	0.0013	(1.90,2.13)	0.23
0.2	0.3	50	1.590	0.125	(1.26,1.92)	0.66	2.037	0.154	(1.67,2.39)	0.72
		150	1.522	0.040	(1.35,1.71)	0.36	2.015	0.056	(1.79,2.22)	0.43
		500	1.504	0.0018	(1.40,1.60)	0.20	2.001	0.0018	(1.88,2.11)	0.23
0.2	0.5	50	1.611	0.116	(1.26,1.91)	0.65	2.063	0.155	(1.67,2.40)	0.73
		150	1.515	0.042	(1.35,1.71)	0.36	2.012	0.057	(1.80,2.23)	0.43
		500	1.508	0.0015	(1.41,1.61)	0.20	2.003	0.0019	(1.89,2.13)	0.24
0.8	-0.6	50	1.585	0.129	(1.26,1.92)	0.66	2.031	0.149	(1.66,2.39)	0.73
		150	1.515	0.039	(1.35,1.72)	0.37	2.006	0.051	(1.80,2.23)	0.43
		500	1.509	0.0013	(1.40,1.59)	0.19	2.005	0.0019	(1.87,2.11)	0.24
0.8	-0.4	50	1.610	0.140	(1.29,1.95)	0.66	2.044	0.167	(1.69,2.41)	0.72
		150	1.519	0.040	(1.35,1.71)	0.36	2.008	0.056	(1.80,2.22)	0.42
		500	1.512	0.0016	(1.41,1.60)	0.19	2.009	0.0018	(1.88,2.12)	0.24
0.8	0.3	50	1.602	0.143	(1.28,1.95)	0.67	2.045	0.153	(1.69,2.41)	0.72
		150	1.520	0.044	(1.34,1.70)	0.36	2.005	0.060	(1.79,2.21)	0.42
		500	1.506	0.0014	(1.40,1.59)	0.19	2.002	0.0018	(1.88,2.11)	0.23
0.8	0.5	50	1.588	0.119	(1.26,1.92)	0.66	2.023	0.153	(1.68,2.41)	0.73
		150	1.517	0.047	(1.34,1.70)	0.36	2.006	0.060	(1.80,2.22)	0.42
		500	1.504	0.0014	(1.41,1.60)	0.19	1.999	0.0019	(1.89,2.12)	0.23

Table 4.b: Modified MLEs for the first-order WARMA models

φ	θ	n	$\hat{\varphi}_{n,4}$	MSE	MLCI	LCI	$\hat{\theta}_{n,4}$	MSE	MLCI	LCI
			$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = W(1.5, 2)$							
-0.5	-0.6	50	-0.500	0.000039	(-0.511,-0.488)	0.022	-0.581	0.023	(-0.68,-0.48)	0.20
		150	-0.5001	0.000036	(-0.511,-0.488)	0.022	-0.594	0.0065	(-0.64,-0.55)	0.09
		500	-0.5000	0.000037	(-0.509,-0.490)	0.019	-0.597	0.0019	(-0.62,-0.58)	0.04
-0.5	-0.4	50	-0.5001	0.000036	(-0.511,-0.488)	0.023	-0.366	0.026	(-0.48,-0.28)	0.20
		150	-0.499	0.000033	(-0.511,-0.489)	0.022	-0.396	0.0065	(-0.43,-0.34)	0.09
		500	-0.5002	0.000035	(-0.509,-0.490)	0.019	-0.399	0.0020	(-0.42,-0.38)	0.04
-0.5	0.3	50	-0.499	0.000037	(-0.511,-0.489)	0.022	0.320	0.025	(0.22,0.42)	0.20
		150	-0.499	0.000035	(-0.510,-0.489)	0.021	0.304	0.0062	(0.25,0.35)	0.10
		500	-0.5001	0.000031	(-0.509,-0.490)	0.019	0.303	0.0020	(0.28,0.32)	0.04
-0.5	0.5	50	-0.499	0.000035	(-0.511,-0.488)	0.023	0.517	0.023	(0.41,0.62)	0.21
		150	-0.499	0.000034	(-0.510,-0.489)	0.021	0.505	0.0069	(0.46,0.55)	0.09
		500	-0.5001	0.000031	(-0.509,-0.490)	0.019	0.501	0.0018	(0.48,0.52)	0.04
-0.1	-0.6	50	-0.1003	0.000048	(-0.114,-0.085)	0.029	-0.575	0.023	(-0.68,-0.48)	0.20
		150	-0.1004	0.000035	(-0.113,-0.086)	0.027	-0.600	0.0065	(-0.64,-0.54)	0.10
		500	-0.1006	0.000032	(-0.112,-0.087)	0.025	-0.597	0.0020	(-0.62,-0.57)	0.05
-0.1	-0.4	50	-0.103	0.000046	(-0.114,-0.085)	0.029	-0.379	0.026	(-0.48,-0.28)	0.20
		150	-0.103	0.000047	(-0.114,-0.086)	0.028	-0.399	0.0066	(-0.44,-0.35)	0.09
		500	-0.102	0.000043	(-0.112,-0.087)	0.025	-0.403	0.0017	(-0.42,-0.37)	0.05
-0.1	0.3	50	-0.098	0.000047	(-0.114,-0.086)	0.028	0.314	0.024	(0.21,0.41)	0.20
		150	-0.0986	0.000045	(-0.114,-0.087)	0.027	0.308	0.0067	(0.25,0.35)	0.10
		500	-0.0989	0.000046	(-0.112,-0.087)	0.025	0.303	0.0019	(0.28,0.32)	0.04
-0.1	0.5	50	-0.098	0.000044	(-0.114,-0.085)	0.029	0.514	0.022	(0.41,0.62)	0.21
		150	-0.099	0.000041	(-0.114,-0.086)	0.028	0.506	0.0058	(0.46,0.55)	0.09
		500	-0.1003	0.000038	(-0.112,-0.087)	0.025	0.501	0.0020	(0.48,0.52)	0.04
0.2	-0.6	50	0.199	0.000048	(0.187,0.212)	0.025	-0.575	0.024	(-0.68,-0.48)	0.20
		150	0.2002	0.000047	(0.187,0.212)	0.025	-0.592	0.0068	(-0.64,-0.55)	0.09
		500	0.199	0.000043	(0.188,0.211)	0.023	-0.597	0.0019	(-0.61,-0.57)	0.04
0.2	-0.4	50	0.199	0.000050	(0.187,0.212)	0.025	-0.388	0.024	(-0.48,-0.28)	0.20
		150	0.199	0.000046	(0.186,0.213)	0.027	-0.397	0.0066	(-0.44,-0.34)	0.10
		500	0.199	0.000043	(0.188,0.211)	0.023	-0.397	0.0020	(-0.42,-0.37)	0.05
0.2	0.3	50	0.199	0.000047	(0.186,0.213)	0.026	0.316	0.022	(0.22,0.42)	0.20
		150	0.2002	0.000043	(0.187,0.213)	0.025	0.305	0.0061	(0.26,0.36)	0.10
		500	0.199	0.000041	(0.188,0.212)	0.024	0.301	0.0019	(0.28,0.32)	0.04
0.2	0.5	50	0.199	0.000050	(0.186,0.213)	0.027	0.522	0.022	(0.42,0.62)	0.20
		150	0.199	0.000049	(0.187,0.212)	0.025	0.504	0.0059	(0.45,0.55)	0.10
		500	0.2001	0.000045	(0.188,0.212)	0.024	0.502	0.0019	(0.48,0.52)	0.04
0.8	-0.6	50	0.798	0.000026	(0.793,0.806)	0.013	-0.588	0.021	(-0.68,-0.47)	0.20
		150	0.799	0.000021	(0.794,0.805)	0.011	-0.597	0.0067	(-0.63,-0.54)	0.09
		500	0.799	0.000019	(0.794,0.805)	0.011	-0.599	0.0018	(-0.62,-0.58)	0.04
0.8	-0.4	50	0.799	0.000025	(0.793,0.806)	0.013	-0.381	0.021	(-0.48,-0.28)	0.20
		150	0.799	0.000020	(0.793,0.806)	0.013	-0.398	0.0066	(-0.44,-0.34)	0.10
		500	0.800	0.000018	(0.794,0.805)	0.011	-0.396	0.0019	(-0.42,-0.38)	0.04
0.8	0.3	50	0.798	0.000023	(0.793,0.806)	0.013	0.316	0.023	(0.22,0.42)	0.20
		150	0.799	0.000019	(0.794,0.805)	0.011	0.303	0.0058	(0.26,0.35)	0.09
		500	0.800	0.000016	(0.794,0.805)	0.011	0.302	0.0018	(0.28,0.32)	0.04
0.8	0.5	50	0.798	0.000023	(0.793,0.806)	0.013	0.511	0.026	(0.41,0.62)	0.21
		150	0.799	0.000017	(0.794,0.805)	0.011	0.503	0.0067	(0.46,0.55)	0.09
		500	0.799	0.000017	(0.795,0.806)	0.011	0.501	0.0020	(0.48,0.52)	0.04

Table 5: Modified MLEs by non-recursive method for the first-order LNARMA models

						$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = LN(0, 0.75)$									
φ	θ	n	$\hat{M}_n$	MSE	MLCI	$\hat{S}_n$	MSE	MLCI	$\hat{\varphi}_{n,2}$	MSE	MLCI	$\hat{\theta}_{n,2}$	MSE	MLCI
-0.6	-0.7	50	-0.019	0.061	(-0.21,0.18)	0.78	0.039	(0.61,0.96)	-0.599	0.000050	(-0.609,-0.590)	-0.69	0.017	(-0.77,-0.62)
		350	-0.006	0.012	(-0.08,0.07)	0.76	0.010	(0.69,0.83)	-0.600	0.000043	(-0.608,-0.591)	-0.69	0.0025	(-0.71,-0.68)
		1000	-0.005	0.0045	(-0.05,0.04)	0.75	0.0039	(0.72,0.80)	-0.600	0.000042	(-0.607,-0.592)	-0.70	0.0010	(-0.71,-0.69)
-0.6	0.2	50	-0.015	0.059	(-0.21,0.18)	0.78	0.038	(0.60,0.95)	-0.600	0.000047	(-0.610,-0.589)	0.20	0.017	(0.13,0.28)
		350	-0.008	0.012	(-0.08,0.07)	0.77	0.0011	(0.70,0.84)	-0.600	0.000046	(-0.607,-0.592)	0.19	0.0026	(0.18,0.21)
		1000	-0.003	0.0040	(-0.05,0.04)	0.75	0.0033	(0.71,0.80)	-0.599	0.000053	(-0.606,-0.593)	0.20	0.0010	(0.19,0.21)
0.3	-0.7	50	0.001	0.061	(-0.19,0.19)	0.77	0.038	(0.59,0.95)	0.299	0.000074	(0.287,0.313)	-0.68	0.017	(-0.75,-0.61)
		350	-0.007	0.015	(-0.08,0.07)	0.77	0.012	(0.70,0.84)	0.299	0.000077	(0.289,0.311)	-0.69	0.0030	(-0.71,-0.67)
		1000	-0.005	0.0046	(-0.05,0.04)	0.76	0.0041	(0.72,0.80)	0.299	0.000059	(0.291,0.308)	-0.69	0.0010	(-0.71,-0.69)
0.3	0.2	50	0.002	0.059	(-0.19,0.19)	0.77	0.038	(0.59,0.94)	0.299	0.000069	(0.286,0.313)	0.22	0.017	(0.15,0.30)
		350	-0.004	0.014	(-0.08,0.07)	0.76	0.010	(0.69,0.83)	0.299	0.000065	(0.287,0.313)	0.20	0.0028	(0.18,0.22)
		1000	-0.004	0.0040	(-0.05,0.04)	0.75	0.0031	(0.71,0.79)	0.299	0.000058	(0.289,0.310)	0.20	0.0010	(0.19,0.21)

Table 6: Modified MLEs by recursive method for the first-order LNARMA models

						$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = LN(0, 0.75)$									
φ	θ	n	$\hat{M}_n$	MSE	MLCI	$\hat{S}_n$	MSE	MLCI	$\hat{\varphi}_{n,2}$	MSE	MLCI	$\hat{\theta}_{n,2}$	MSE	MLCI
-0.6	-0.7	50	-0.024	0.074	(-0.23,0.20)	0.78	0.050	(0.59,0.98)	-0.600	0.000055	(-0.610,-0.589)	-0.69	0.021	(-0.78,-0.61)
		350	-0.015	0.013	(-0.09,0.06)	0.78	0.011	(0.70,0.84)	-0.599	0.000058	(-0.608,-0.591)	-0.69	0.0027	(-0.72,-0.68)
		1000	-0.011	0.0050	(-0.05,0.03)	0.76	0.0045	(0.72,0.80)	-0.599	0.000045	(-0.607,-0.593)	-0.70	0.0010	(-0.71,-0.69)
-0.6	0.2	50	-0.021	0.083	(-0.23,0.19)	0.79	0.049	(0.59,0.98)	-0.598	0.000067	(-0.608,-0.590)	0.20	0.025	(0.11,0.29)
		350	-0.011	0.014	(-0.09,0.07)	0.77	0.012	(0.69,0.84)	-0.598	0.000046	(-0.607,-0.592)	0.19	0.0029	(0.18,0.22)
		1000	-0.001	0.0046	(-0.05,0.04)	0.75	0.0040	(0.71,0.80)	-0.599	0.000045	(-0.606,-0.593)	0.20	0.0010	(0.19,0.21)
0.3	-0.7	50	0.025	0.073	(-0.19,0.22)	0.79	0.045	(0.57,0.96)	0.293	0.000083	(0.287,0.313)	-0.67	0.022	(-0.76,-0.59)
		350	-0.008	0.012	(-0.09,0.07)	0.77	0.012	(0.70,0.84)	0.299	0.000059	(0.289,0.311)	-0.69	0.0026	(-0.71,-0.68)
		1000	-0.004	0.0054	(-0.05,0.04)	0.75	0.0047	(0.72,0.80)	0.299	0.000066	(0.291,0.308)	-0.69	0.0010	(-0.70,-0.69)
0.3	0.2	50	0.021	0.076	(-0.22,0.21)	0.78	0.047	(0.58,0.97)	0.293	0.000058	(0.287,0.312)	0.22	0.024	(0.12,0.30)
		350	-0.012	0.015	(-0.08,0.07)	0.76	0.012	(0.69,0.83)	0.299	0.000070	(0.285,0.314)	0.20	0.0029	(0.18,0.22)
		1000	-0.009	0.0045	(-0.05,0.04)	0.76	0.0036	(0.71,0.80)	0.299	0.000055	(0.289,0.310)	0.21	0.0010	(0.19,0.21)

Table 7: Modified MLEs by reduced method for the first-order LNARMA models

						$z_t = \varphi z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t = LN(0, 0.75)$									
φ	θ	n	$\hat{M}_n$	MSE	MLCI	$\hat{S}_n$	MSE	MLCI	$\hat{\varphi}_{n,2}$	MSE	MLCI	$\hat{\theta}_{n,2}$	MSE	MLCI
-0.6	-0.7	50	-0.046	0.15	(-0.32,0.27)	0.79	0.083	(0.51,1.07)	-0.600	0.000066	(-0.608,-0.591)	-0.71	0.11	(-0.86,-0.57)
		350	-0.016	0.016	(-0.09,0.07)	0.77	0.013	(0.69,0.85)	-0.600	0.000067	(-0.608,-0.591)	-0.68	0.0034	(-0.72,-0.68)
		1000	-0.008	0.0057	(-0.05,0.04)	0.78	0.0048	(0.71,0.80)	-0.599	0.000066	(-0.607,-0.592)	-0.70	0.0012	(-0.71,-0.69)
-0.6	0.2	50	-0.043	0.15	(-0.31,0.27)	0.78	0.083	(0.50,1.06)	-0.600	0.000070	(-0.609,-0.590)	0.20	0.11	(0.04,0.33)
		350	-0.014	0.014	(-0.08,0.08)	0.78	0.012	(0.69,0.84)	-0.600	0.000059	(-0.608,-0.591)	0.20	0.0030	(0.18,0.22)
		1000	-0.007	0.0060	(-0.05,0.04)	0.76	0.0047	(0.71,0.80)	-0.599	0.000065	(-0.606,-0.593)	0.20	0.0010	(0.19,0.21)
0.3	-0.7	50	0.048	0.16	(-0.24,0.32)	0.79	0.084	(0.46,1.02)	0.297	0.000076	(0.287,0.312)	-0.66	0.10	(-0.81,-0.52)
		350	-0.011	0.014	(-0.08,0.09)	0.78	0.012	(0.68,0.83)	0.299	0.000079	(0.288,0.312)	-0.69	0.0032	(-0.71,-0.67)
		1000	-0.005	0.0058	(-0.05,0.04)	0.76	0.0048	(0.71,0.80)	0.299	0.000063	(0.291,0.308)	-0.70	0.0011	(-0.71,-0.69)
0.3	0.2	50	0.052	0.16	(-0.26,0.31)	0.78	0.074	(0.49,1.04)	0.291	0.000083	(0.288,0.311)	0.21	0.12	(0.07,0.35)
		350	-0.010	0.016	(-0.09,0.07)	0.77	0.012	(0.69,0.85)	0.300	0.000076	(0.284,0.315)	0.20	0.0035	(0.18,0.22)
		1000	-0.006	0.0056	(-0.05,0.04)	0.76	0.0047	(0.71,0.80)	0.300	0.000070	(0.289,0.311)	0.20	0.0010	(0.19,0.21)

5 Discussion and Future Works

We calculated the Modified MLEs to estimate the parameters of the first-order ARMA models when the residuals come from the two parametric distribution families: exponential and Weibull. Modified estimators have been presented in three forms: non-recursive, recursive and reduced. The recursive form obtained by recursive definition of the residuals based on time series observations is a more functional approach than the non-recursive form. However, it has more complex equations than the non-recursive form. While, the reduced form follows simpler equations compared to the recursive form and are computable in functional situations.

Estimating the first-order ARMA model parameters with non-normal residuals using modified MLEs is important in different fields. Model selection is one of the fields in which model parameters need to be estimated. After diagnosing the non-normality of the data distribution, the main point that should be discussed is the proper distribution of the residuals. Selecting the most appropriate distribution for the residuals under consideration is in the scope of model selection. The introduced model selection methods are split into two categories, criteria and statistical hypothesis tests, both of which are mainly based on Kullback-Leibler divergence criterion (KL) (Kullback and Leibler (1951)). According to the KL criterion, the appropriate fitted distribution to the data is the one with the minimum discrepancy from the true distribution of the data. Minimizing the KL criterion is equivalent to maximizing the portion of this criterion which leads to the calculation of the likelihood function and the estimation of the model parameters using the ML procedure. It is therefore very important to provide solutions to estimate the parameters of different models using the ML method, particularly in non-normal mode, where the selection of the model is desired. Modified MLEs, by estimating the parameters of the first-order ARMA models, enable us to select the optimal model of the residuals and there by improve the modeling and forecasting of time series.

Although obtaining estimators is important for modeling time series by the first-order ARMA model, further aspects of estimators need to be discussed to complete the analysis. Among them, the analysis of features such as the consistency and efficiency of the modified MLEs can be suggested. Furthermore, since in model diagnostic discussions the distribution of the data generator is assumed unknown, it seems necessary to re-analyze the important features of the estimators such as their consistency, efficiency and asymptotic distribution. Another subject that should be discussed is the robustness of the estimators. By considering the dependency of the structure of the modified MLEs on the sample mean of Z_t, these estimators are definitely not robust against outliers. Therefore, a broad discussion upon robustness should be considered in future studies.

6 Conclusion

The aim of this study was to estimate the parameters of the first-order ARMA model while the residuals that generate the process were non-normally distributed. MLEs were ourestimators of interest, but due to the complexity of the likelihood functions, there was no explicit solution for them. For solving this problem, we used a modified method that was based on linearization of the log-likelihood function in complex phrases using first-order Taylor's series expansion. The residuals were selected from exponential and Weibull families. Then the coefficients and residual distribution parameters of the first-order ARMA model were estimated in several subsets of these families. Ultimately, the simulation study confirmed the theoretical conclusions. The novelty of this study is the approximation of MLEs in a modified way to estimate the first-order ARMA model parameters in non-normal mode. It facilitates time series modeling based on the first-order ARMA models in non-normal mode. However, the parameter estimation methods produced in this study may need more development. To complete this procedure, further supplementary concepts such as robustness of the proposed estimators should be addressed in future works.

Acknowledgements

The authors thank two referees for constructive suggestions.

References

Akaike, H., Petrov, B. N., and Csaki, F. (1973). Second international symposium on information theory. Akadémiai Kiadó, Budapest.

Bayer, F. M., Cintra, R. J., and Cribari-Neto, F. (2018). Beta seasonal autoregressive moving average models. ArXiv E-Prints, 1806, arXiv:1806.07921.

Box, G. E., and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, **26**(2), 211–243.

Box, G. E. P., and Jenkins, G. M. (1976). Time series analysis forecasting and control (Revised ed).

Braga, D., and Calmon, W. (2017). Periodic Gamma Autoregressive Model: An application to the Brazilian hydroelectric system. *RAIRO - Operations Research*, **51**(2), 469–483.

Casella, G., and Berger, R. L. (2001). Statistical Inference (2nd edition). Cengage Learning.

Cramér, H. (1946). Mathematical methods of statistics. Princeton University Press.

Cryer, J. D., and Chan, K. S. (2008). Time Series Analysis: With Applications in R (2nd ed.). Springer-Verlag.

Duca, V. E. L. de A., Souza, R. C., Ferreira, P. G. C., and Oliveira, F. L. C. (2019). Simulation of time series using periodic gamma autoregressive models. *International Transactions in Operational Research*, **26**(4), 1315–1338.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, **38**(1), 34–105.

Fernandez, B., and Salas, J. D. (1986). Periodic gamma autoregressive processes for operational hydrology. *Water Resources Research*, **22**(10), 1385–1396.

Gaver, D. P., and Lewis, P. A. W. (1980). First-order autoregressive gamma sequences and point processes. *Advances in Applied Probability*, **12**(3), 727–745.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1: Statistics.

Hurvich, C. M., and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, **76**(2), 297–307.

Kendall, M. G., and Hill, A. B. (1953). The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, **116**(1), 11–34.

Kullback, S., and Leibler, R. A. (1951). On information and sufficiency, *Annals Maths. Statist*, **22**, 79–86.

Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?, *Journal of Econometrics*, **54**(1–3), 159–178.

Li, W. K., and McLeod, A. I. (1988). ARMA modelling with non-Gaussian innovations. *Journal of Time Series Analysis*, **9**(2), 155–168.

Mandelbrot, B. (1967). The variation of some other speculative prices. *The Journal of Business*, **40**(4), 393–413.

Mandelbrot, B. B. (1997). The variation of certain speculative prices. In Fractals and scaling in finance (pp. 371–418). Springer.

Mandelbrot, B., and Taylor, H. M. (1967). On the distribution of stock price differences. *Operations Research*, **15**(6), 1057–1062.

Sarlak, N., and Sorman, A. (2007). Gamma autoregressive models and application on Kizilirmak Basin.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, **6**(2), 461–464.

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, **52**(3–4), 591–611.

Spierdijk, L. (2016). Confidence intervals for ARMA–GARCH Value-at-Risk: The case of heavy tails and skewness. *Computational Statistics and Data Analysis*, **100**, 545–559.

Tiku, M. L., Wong, W. K., and Bian, G. (1999). Time series models with asymmetric innovations. *Communications in Statistics - Theory and Methods*, **28**(6), 1331–1360.

Trindade, A. A., Zhu, Y., and Andrews, B. (2010). Time series models with asymmetric Laplace innovations. *Journal of Statistical Computation and Simulation*, **80**(12), 1317–1333.

Organization for Economic Co-operation and Development. (2020, January 1). Unemployment Rate: Aged 15-64: All Persons for the United States. FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series.

Wei, W. W. (2006). Time series analysis. In The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2.

Weiss, G. (1975). Time-Reversibility of linear stochastic processes.

Weiss, G. (1977). Shot noise models for the generation of synthetic streamflow data. *Water Resources Research*, **13**(1), 101–108.

Yakowitz, S. J. (1973). A stochastic model for daily river flows in an arid region. *Water Resources Research*, **9**(5), 1271–1285.

Zamani Mehreyan, S., and Sayyareh, A. (2017). Separated hypotheses testing for autoregressive models with non-negative residuals. *Journal of Statistical Computation and Simulation*, **87**(4), 689–711.

Zhang, Z., and Li, W. K. (2019). An Experiment on Autoregressive and Threshold Autoregressive Models with Non-Gaussian Error with Application to Realized Volatility. *Economies*, **7**(2), 58.