

Methods for Estimating the AR (1) Autoregressive Model Parameter for Data Following an Exponential Distribution

E. A. Hussain ^(1*) , A. N. S. Al-Shallawi ⁽²⁾ , A. H. Taha ⁽³⁾ 

^(1,2) Northern Technical University, Administrative Technical College, Mosul, Iraq

⁽³⁾ Northern Technical University, Technical Management Institute, Mosul, Iraq

Article information

Article history:

Received January 19, 2026

Revised: June 15, 2026

Accepted: June 21, 2026

Available online: October 01, 2026

Keywords:

Autoregressive

Bayes

Exponential Distribution

Fisher Information

Correspondence:

Elham Abdulkreem Hussain

e.a.hussain@ntu.edu.iq

Abstract

This study aims to compare the performance of different estimation methods for a first-order autoregressive model (AR(1)) under the assumption that random errors do not follow a normal distribution but rather an exponential one. Four estimation methods were used: Bayesian, Jeffrey, Maximum Likelihood (MLE), and Least Squares (OLS), to evaluate their efficiency in estimating the model parameters.

The study was applied to real temperature data for Mosul, in addition to simulated data generated to verify the reliability of the results. The comparison between the methods was based on Mean Squared Error (MSE) and Akaike Information Criterion (AIC).

The results showed that the best model performance was achieved at a value of $\phi=0$, where the lowest values for both AIC and MSE were obtained in both the real and simulated data, indicating no significant time dependence in the studied series. The results also showed that Bayesian methods have higher flexibility and stability compared to traditional methods, especially under the assumption of an exponential distribution of errors.

DOI: [10.33899/8wv0k731](https://doi.org/10.33899/8wv0k731), ©Authors, 2026, College of Education for Pure Science, University of Mosul.

This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Several studies have addressed the estimation of ARIMA model parameters under non-normal error distributions, including the exponential distribution, to investigate the impact of this assumption on the estimator properties and statistical efficiency. These studies have focused on comparing different estimation methods and using simulation techniques to examine estimator performance. For example:

[27] concluded that using the Student-t distribution in time series models improves prediction accuracy and increases the robustness of the estimate compared to the normal distribution, especially in the presence of extreme values and heavy tails.

[9] showed that using the Student-t distribution in time models improves the robustness of the estimate in the presence of outliers, compared to the normal distribution, which is sensitive to these values.

[19] showed that Bayesian models of autoregressive processes, such as AR(1), allow for the estimation of time-varying parameters, making them better able to represent realistic time series data compared to traditional models.

[26] presented a study in which he developed Bayesian estimators for autoregressive models with explanatory variables in both known and unknown variance cases. He demonstrated their consistency and convergent nature, and simulation results showed that these estimators are robust and highly insensitive to the selection of prior distributions.

[3] presented a comprehensive theoretical framework for non-normal error time series, showing that alternative distributions such as Exponential, Gamma, Laplace, and Cauchy can be used in building autoregressive models, giving greater flexibility in representing real data.

[22] presented a study in 2023 proposing the use of a second-order autoregressive AR(2) model with power-exponential tracking errors as an alternative to the normal distribution.

The performance of the proposed model was evaluated against the conventional model using the AIC and prediction accuracy measures such as RMSE and MAE, based on real and simulated data with varying sample sizes. The results showed that using the power-exponential tracking error distribution provides more efficient and accurate performance in modeling non-normalized time series compared to the normalization assumption.

[2] presented a study demonstrating that combining Bayesian estimation with non-normal distributions (such as exponential and gamma) enables more flexible modeling, where parameters are estimated through a dimensionless distribution. This helps to address asymmetry and heavy tails in the data.

[1] compared different estimators, such as maximum likelihood, percentile maximum likelihood, and least squares, for estimating exponential distribution parameters using Monte Carlo simulation.

[12] discussed methods for estimating generalized exponential distribution parameters, such as the maximum likelihood method and the moment method, comparing them in different contexts.

[11] presented an exponential regression model for handling complex time series data and discussed the estimation of this model's parameters using appropriate statistical methods.

[17] also discussed prediction models such as simple exponential smoothing and the ARIMA (0,2,2) model for predicting lung cancer mortality rates in the United States, identifying the best model based on the lowest mean squared error value.

Time series are considered one of the statistical methods used in analyzing many phenomena that are characterized by time, through which it is possible to know the nature of the changes that occur in the values of those phenomena over time and then determine the results to predict their future values, which is considered the main goal of analyzing these phenomena. Numerous studies have been conducted on time series in many modern disciplines and applications, such as financial time series forecasting, medicine, engineering, signal processing, and the latest in the fields of machine learning and pattern recognition[24],[5].

As for the exponential distribution, it is used in problems related to measuring time, or it can be said that it is used to estimate the time periods between the occurrence of events, such as the duration of a phone call, the duration of a machine repair, the duration of a customer waiting before obtaining the service, etc. Therefore, the distribution depends on time. Then, the distribution parameter is greater than zero [21].

This study aimed to construct an autoregressive (AR) model using time-series data and estimate the model parameters based on four different estimation methods previously described, assuming that random errors followed an exponential distribution rather than a normal distribution. A systematic comparison of these methods was then conducted, evaluating their performance using the mean squared error (MSE) criterion to determine the most efficient and accurate method.

This study also offers an experimental addition by applying these models and estimation methods to real-world climate data representing temperatures in Mosul over a specific time period. Mosul is characterized by unique climatic features, such as temperature fluctuations, significant seasonal variations, and rainfall patterns.

The significance of this addition lies in its ability to test the efficiency of different estimation methods under harsh and semi-arid climatic conditions. This contributes to assessing the suitability of these methods in novel environments and enhances the generalizability of the results.

2. Research Method

Many researchers have written about time series models known as autocorrelation and moving average models, which are denoted by ARMA if the series is stationary, and by ARIMA if the series is unstable, using other distributions such as the Gamma distribution, Normal distribution, Poisson distribution, Bernoulli distribution, etc., and with different orderings for the models. These researchers include, but are not limited to, the following [16],[14], [13], and [18].

Although some of the approaches listed in some of the recent research were not identified in this study, the researcher felt that listing them here can help in the search for the most straightforward approach to utilize.

As is known, the exponential distribution function has several forms, and the form adopted in this research is as in equation (1) with a change in the parameters according to the research goals [23]:

$$f(y|\theta, t_0) = \frac{1}{\theta} \exp \left[- \left(\frac{(y - t_0)}{\theta} \right) \right] \quad y \geq \theta, t_0 > 0 \quad (1)$$

Whereas:

(t_0, θ) : Model parameters

Since the goal of the research is to estimate the parameters of the exponential model according to equation (1) for data tracking the AR (1) model, equation (1) will be as follows[6]:

$$y_t = \phi_1 y_{t-1} + e_t \quad (2)$$

Whereas:

y_t is a time series, e_t is a white disturbance and has an exponential distribution $e_t \sim \text{Exp}(\theta)$ and ϕ_1 : model's parameter. Therefore, the equation becomes as follows:

$$f(y; \theta, \varphi_1) = \frac{1}{\theta} e^{-\frac{(y_t - \varphi_1 y_{t-1})}{\theta}} \quad \theta > 0, \quad 0 \leq \varphi_1 \leq 1 \quad (3)$$

Since the series is stationary when the value of the AR model parameter is $|\varphi_1| < 1$, and since the exponential model relates to time and its parameters start from zero, we will assume that the AR(1) model parameter will be limited to between (0,1).

3. Estimation Methods

Several methods will be presented to estimate the parameters of the exponential distribution, which are maximum likelihood, least squares, Bayesian method using Fisher information (Jeffrey), and Bayesian method using the posterior distribution.

A- Maximum Likelihood Method

Maximum Likelihood (likelihood function) is used to estimate parameters, and these estimators make the logarithm of the likelihood function at its greatest limit [15]. If we have a random sample $(y_1, y_2, \dots, y_{(n-1)})$ from the probability distribution defined in equation (3), the maximum likelihood function will be:

$$L(y_1, y_2, \dots, y_{(n-1)}; \theta, \varphi) = \prod_{t=2}^n \frac{1}{\theta} e^{-\frac{(y_{t-1} - \varphi_1 y_t)}{\theta}} \quad (4)$$

$$= \frac{1}{\theta^{n-1}} e^{-\frac{\sum_{t=2}^n (y_{t-1} - \varphi_1 y_t)}{\theta}}$$

$$\ln L = -\ln(\theta)^{n-1} - \frac{\sum_{t=2}^n (y_{t-1} - \varphi_1 y_t)}{\theta} \quad (5)$$

$$\frac{\partial \ln L}{\partial \theta} = -\frac{(n-1)}{\theta} + \frac{1}{\theta^2} \sum_{t=2}^n (y_{t-1} - \varphi_1 y_t)$$

$$n-1 = \frac{1}{\theta} \sum_{t=2}^n (y_{t-1} - \varphi_1 y_t)$$

$$\theta = \frac{\sum_{t=2}^n y_{t-1}}{n-1} - \varphi_1 \frac{\sum_{t=1}^{n-1} y_t}{n-1}$$

$$\hat{\theta} = \bar{y}_2 - \varphi_1 \bar{y}_1 \quad (6)$$

It can be driven φ_1

$$\frac{\partial \ln L}{\partial \varphi_1} = -\frac{\sum_{t=1}^{n-1} y_t}{\theta} = -\frac{(n-1)\bar{y}_1}{\theta} = -\frac{(n-1)\bar{y}_1}{\bar{y}_2 - \varphi_1 \bar{y}_1}$$

Here, the parameter estimation fails; therefore, maximum likelihood function estimation for the parameter φ_1 can be used according to the formula obtained by [10], which is called the first modified method of maximum likelihood, which is as follows:

$$\hat{\varphi}_{MLE-1} = \frac{y_{(1)} + \bar{y} \log\left(\frac{n}{n+1}\right)}{1 + \log\left(\frac{n}{n+1}\right)} \quad (7)$$

B- Least Squares Method

This classical method relies on reducing the sum of the squares of the residuals to its lowest value [4] see equation (2):

$$Ey_t = \varphi_1 Ey_{t-1} + Ee_t \quad (8)$$

Since e_t is exponentially distributed $e_t \sim \text{Exp}(\theta)$, therefore equation (8) becomes:

$$y_t = \varphi_1 y_{t-1} + \theta \quad (9)$$

From (9)

$$Q = \sum_{t=2}^n [(y_t - (\varphi_1 y_{t-1} + \theta))]^2 \quad (10)$$

It is noted that $t=2$ because the smallest value of y in equation (10) is the same as the sequence y_{t-1} . Taking the derivative of equation (10) with respect to the parameter φ_1 results in:

$$\frac{\partial Q}{\partial \varphi_1} = 2 \sum_{t=2}^n (y_t - (\varphi_1 y_{t-1} + \theta)) (-y_{t-1}) \quad (11)$$

$$\text{let } \frac{\partial Q}{\partial \varphi_1} = 0$$

$$\begin{aligned}
0 &= - \sum_{t=2}^n y_t y_{t-1} + \varphi_1 \sum_{t=2}^n y_{t-1}^2 + \theta \sum_{t=2}^n y_{t-1} \\
\varphi_1 \sum_{t=2}^n y_{t-1}^2 &= \sum_{t=2}^n y_t y_{t-1} - \theta \sum_{t=2}^n y_{t-1} \\
\hat{\varphi}_1 &= \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} - \theta \frac{\sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2}
\end{aligned} \tag{12}$$

derivative equation(10) for θ

$$\frac{\partial Q}{\partial \theta} = 2 \sum_{t=2}^n (y_t - (\varphi_1 y_{t-1} + \theta) (-1))$$

Equal it to zero

$$\begin{aligned}
0 &= \sum_{t=2}^n y_t - \varphi_1 \sum_{t=2}^n y_{t-1} - \sum_{t=2}^n \theta \\
(n-1)\theta &= \sum_{t=2}^n y_t - \varphi_1 \sum_{t=2}^n y_{t-1} \\
\hat{\theta} &= \frac{\sum_{t=2}^n y_t}{(n-1)} - \varphi_1 \frac{\sum_{t=2}^n y_{t-1}}{(n-1)}
\end{aligned} \tag{13}$$

substitute (12) in (13)

$$\begin{aligned}
\varphi_1 &= \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} - \left(\frac{\sum_{t=2}^n y_t}{(n-1)} - \varphi_1 \frac{\sum_{t=2}^n y_{t-1}}{(n-1)} \right) \frac{\sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \\
&= \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} - \frac{\sum_{t=2}^n y_t}{(n-1)} \frac{\sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} + \varphi_1 \frac{(\sum_{t=2}^n y_{t-1})^2}{(n-1) \sum_{t=2}^n y_{t-1}^2} \\
\varphi_1 - \varphi_1 \frac{(\sum_{t=2}^n y_{t-1})^2}{(n-1) \sum_{t=2}^n y_{t-1}^2} &= \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} - \frac{\sum_{t=2}^n y_t}{(n-1)} \frac{\sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2}
\end{aligned}$$

Multiply by $(n-1)$:

$$\begin{aligned}
(n-1)\varphi_1 - \varphi_1 \frac{(\sum_{t=2}^n y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2} &= \frac{(n-1) \sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} - \frac{\sum_{t=2}^n y_t}{\sum_{t=2}^n y_{t-1}^2} \frac{\sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \\
\varphi_1 \left((n-1) - \frac{(\sum_{t=2}^n y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2} \right) &= \frac{(n-1) \sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} - \frac{\sum_{t=2}^n y_t}{\sum_{t=2}^n y_{t-1}^2} \frac{\sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \\
\varphi_1 &= \frac{\frac{(n-1) \sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} - \frac{\sum_{t=2}^n y_t}{\sum_{t=2}^n y_{t-1}^2} \frac{\sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2}}{(n-1) - \frac{(\sum_{t=2}^n y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2}}
\end{aligned}$$

$\sum_{t=2}^n y_{t-1}^2$ Multiply by

$$\hat{\varphi}_1 = \frac{(n-1) \sum_{t=2}^n y_t y_{t-1} - \sum_{t=2}^n y_t \sum_{t=2}^n y_{t-1}}{(n-1) \sum_{t=2}^n y_{t-1}^2 - (\sum_{t=2}^n y_{t-1})^2} \tag{14}$$

C- Bayes Estimator Using Fisher Information (Jeffery)

It is a way to measure the amount of information that an observable random variable X carries about an unknown parameter, such as θ of a distribution representing X [7]. As is known, this method is based on finding the second derivative of the likelihood function, so we obtain the second derivative of equation (5) relative to θ :

$$\begin{aligned}
\frac{\partial^2 \ln l}{\partial \theta^2} &= \frac{(n-1)}{\theta^2} - \frac{2}{\theta^3} \sum_{t=2}^n (y_t - \varphi_1 y_{t-1}) \\
&= \frac{(n-1)\theta - 2 \sum_{t=2}^n (y_t - \varphi_1 y_{t-1})}{\theta^3} \\
I(\theta) &= -E \left(\frac{\partial^2 \ln l}{\partial \theta^2} \right) = \frac{-(n-1)\theta + 2 \sum_{t=2}^n (E y_t - \varphi_1 E y_{t-1})}{\theta^3}
\end{aligned}$$

$$\begin{aligned}
&= \frac{-(n-1)\theta + 2 \sum_{t=2}^n (\theta - \varphi_1 \theta)}{\theta^3} \\
&= \frac{-(n-1)\theta + 2[(n-1)\theta - n\theta\varphi_1]}{\theta^3} \\
&= \frac{-(n-1)\theta + 2\theta(n-1-\varphi_1)}{\theta^3} \\
&= \frac{\theta(n-1-\varphi)}{\theta^3}
\end{aligned}$$

By extracting $(n-1)*\theta$ as a common factor, we get:

$$-E\left(\frac{\partial^2 \ln l}{\partial \theta^2}\right) = \frac{(n-1-\varphi)}{\theta^2} = \frac{c}{\theta^2}$$

$$\sqrt{I(\theta)} = \frac{\sqrt{c}}{\theta} = g(\theta) \quad (15)$$

Whereas:

$$c = (n-1-\varphi)$$

As for Jeffrey's information for the parameter φ_1 , we derive equation (5) relative to this parameter, which yields:

$$\frac{\partial \ln l}{\partial \varphi_1} = \frac{\sum_{t=1}^{n-1} y_{t-1}}{\theta}; \quad \frac{\partial^2 \ln l}{\partial \varphi_1^2} = 0$$

This is a failure in the derivation. If the estimator of the Fisher information is equal to zero, an approximate method is used. There are a number of different ways to do this, such as using the maximum value, the minimum value, or the maximum value of the parameter. Either the value of φ is equal to 1 or zero, depending on the value specified above. Since the minimum value is zero, the Jeffrey value of information is equal to the highest possible value of φ , which is 1, meaning:

$$\sqrt{I(\varphi)} = \varphi = 1 \quad (16)$$

Thus, equation (15) becomes, after substituting the value of φ , as follows: $\sqrt{I(\theta)} = \frac{\sqrt{n}}{\theta}$.

D- Bayes Estimator Using posterior distribution

Bayes' method of parameter estimation relies on the initial information to be estimated. In Bayes' method, the parameters are random variables, not constants as in traditional estimation methods. These variables have a function and a probability distribution known as the initial probability density function (or prior distribution), which represents the initial parameter information. Sample information is expressed as the likelihood function. What is important to us is the estimate of φ , so only the estimator φ will be found, and the estimate of θ will be neglected. The posterior distribution is expressed by combining the two prior distribution functions with the likelihood function [14]:

$$P(\varphi|t) \propto L(t|\varphi)P(\varphi)$$

Where [8] :

$P(\varphi|t)$: represents the posterior distribution of the parameter (φ). $P(\varphi)$: the prior distribution of the parameter (φ), and is equal to 1. $L(t|\varphi)$: was defined previously

Since the maximum likelihood function for the exponential model of the data AR (1) is as in equation (4), and since the posterior distribution is obtained by combining the prior distribution and the maximum likelihood function, here the Jeffrey distribution can be used and considered the prior distribution:

$$p(y|\theta, \varphi) = L * P(\varphi) \quad (17)$$

$$= \frac{1}{\theta^{n-1}} e^{-\frac{\sum_{t=2}^n (y_t - \varphi_1 y_{t-1})}{\theta}} * 1$$

To obtain φ , we integrate equation (17) with respect to θ :

$$P(y_1, y_2, \dots, y_n, \varphi) = \int_0^\infty p(y|\theta, \varphi) d\theta$$

$$P(y_1, y_2, \dots, y_n, \varphi) = \int_0^\infty \frac{1}{\theta^{n-1}} e^{-\frac{\sum_{t=2}^n (y_t - \varphi_1 y_{t-1})}{\theta}} d\theta \quad (18)$$

Suppose that:

$$\sum_{t=2}^n (y_t - \varphi_1 y_{t-1}) = w$$

Equation (18) will become:

$$P(y_1, y_2, \dots, y_n, \varphi) = \int_0^\infty \frac{1}{\theta^{n-1}} e^{-\frac{w}{\theta}} d\theta \quad (19)$$

This is the kernel of the gamma distribution with $\alpha = (n-1)$, $\beta = \sum_{t=2}^n (y_t - \varphi_1 y_{t-1})$, therefore:

$$p(\varphi) = \frac{\Gamma(n-1)}{(\sum_{t=2}^n (y_t - \varphi_1 y_{t-1}))^{(n-1)}} \quad (20)$$

$$\therefore \hat{\varphi} = E(\varphi) = \frac{n-1}{\sum_{t=2}^n (y_t - \varphi_1 y_{t-1})}$$

4. Criteria for selecting the best model (AIC and MSE)

To compare among the statistical estimates according to the above estimation methods, mean square error (MSE) and Akaike Information were used :

a. Mean Square Error (MSE)

MSE is one of the most common measures, according to the following formula:

$$MSE(\hat{\varphi}_1) = \frac{1}{n-1} \sum_{i=1}^R (\hat{\varphi}_{1i} - \varphi_1)^2 \quad (21)$$

Where:

n: sample size.

φ_1 : actual parameter

$\hat{\varphi}_{1i}$: estimated parameter

MSE has been adopted as the primary metric for comparing estimation methods because it is one of the most commonly used measures for assessing the accuracy of estimation and prediction in time series models, and it reflects both bias and variance. Furthermore, MSE gives greater weight to large errors due to the squared deviations, making it sensitive to differences in performance between various methods.

b. Akaike information criterion (AIC) :

AIC is an estimator of prediction error and thereby the relative quality of statistical models for a given set of data.

Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection. AIC value of the model is the following [20] :

$$AIC = 2K - 2\ln(\hat{L}) \quad (22)$$

Where :

K = number of parameters.

L : likelihood function

5. Results And Discussion

In this aspect, we will address two cases. The first case is a practical application on real data, and in the second case, we will resort to generating data through simulation, and then we will work on comparing the results of the real data and the results of the generated data:

5.1 First: Real data study

[1] Daily temperature data in Iraq was chosen for the period 1/1/2023 - 18/8/2023 **with 230 observations**, taken from [25]

as an application of estimation methods. The figure shows that the series is unstable in mean and variance, so the natural logarithm of the data was taken, and then its first difference, as shown in the following figure:

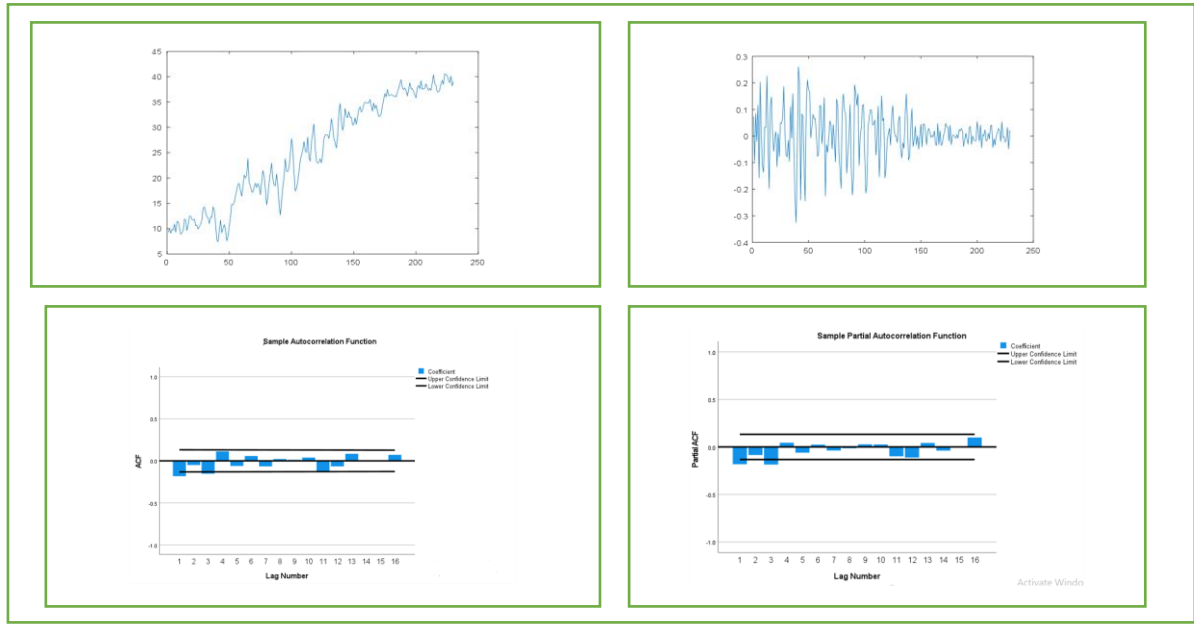


Figure 1: Stationary and non-stationary time series data and ACF, PACF functions

Autocorrelation (ACF) and partial autocorrelation (PACF) functions indicate that the time series may follow a first-order autoregressive model (AR(1)). However, given the potential for this suggestion to be inaccurate, several alternative models will be adopted and compared using the mean squared error (MSE) criterion to select the most efficient model for estimation, as shown below.

Table 1: Comparison and evaluation of statistical models

Model	ϕ	θ	S.E	t-value	Sig.	Ljung-Box	p-value	MSE
ARIMA(1,1,0)	-0.136		0.066	-2.075	0.039	41.265	0.11	6.360
ARIMA(2,1,0)	-0.161		0.065	-2.471	0.014	41.339	0.001	6.170
	-0.195		0.065	-2.987	0.003			
ARIMA(1,1,1)	0.664		0.063	10.525	0.001	25.226	0.066	6.066
		0.958	0.027	35.242	0.001			
ARIMA(0,1,1)		0.263	0.64	4.106	0.001	41.733	0.001	6.245
ARIMA(0,1,2)		0.313	0.065	4.841	0.001	31.020	0.013	5.822
		0.262	0.064	4.070	0.001			

First, the significance of the estimated parameters (ϕ and θ) was tested using the standard error of magnitude (S.E.), t-values, and significance levels (Sig.). Only models with statistically significant parameters (p-value < 0.05) were retained. Second, the fit of the models was examined using the Ljung-Box test. Non-significant values (p-value > 0.05) indicate that the residuals are random and lack autocorrelation, thus demonstrating model fit.

Finally, the accepted models were compared using the MSE criterion, and the model with the lowest MSE value was selected. The table above shows that the appropriate model is the autoregressive ARIMA(1,1,0) model, as it met the basic condition of non-significance in the Ljung-Box test, indicating the absence of autocorrelation in the residuals. In contrast, the other models showed significance in this test, indicating the presence of an unexplained autocorrelation, and therefore, they were excluded.

Applying the estimation methods that the researcher discussed in the theoretical aspect, we tested the methods by taking multiple values of ϕ from 0 to 0.9 and finding the MSE value (which is referred to according to equation (21)) for each estimation method. The results were as in Table 2:

Table 2: Estimation methods and MSE, AIC values for each method for temperatures in Mosul city for the period from 1/1/2023 to 18/8/2023

ϕ	MSE				AIC
	OLS	ML	Jeffery	Bayes	
0	0.1964	0.2332	0.0043	0.0000069	-1921.27
0.1	0.1906	0.2268	0.0035	0.0000136	-1873.48
0.2	0.1849	0.2206	0.0028	0.0000986	-1820.09
0.3	0.1793	0.2144	0.0021	0.0002581	-1759.61
0.4	0.1737	0.2084	0.0016	0.0004861	-1689.87
0.5	0.1683	0.2024	0.0011	0.0007715	-1607.49
0.6	0.1629	0.1965	0.00069	0.0010931	-1506.88
0.7	0.1576	0.1907	0.00039	0.0014034	-1377.60
0.8	0.1525	0.1850	0.00017	0.0015711	-1196.49
0.9	0.1473	0.1794	0.000043	0.0010759	-891.57

Table 2 shows the results of comparing the first-order autoregressive model AR(1) under different values of the parameter ϕ , using both AIC and MSE. The results show that the lowest AIC value was (-1921.27) at $\phi=0$, and the same case also recorded the lowest MSE value of (0.0000069), indicating that this model is the best in terms of statistical fit and prediction accuracy. It is also evident that increasing the value of ϕ leads to higher values for both AIC and MSE, demonstrating a deterioration in the model's performance when time dependence is introduced. These results suggest that the studied time series does not exhibit significant time dependence and tends toward behavior close to random noise.

The results also show a difference in the behavior of the various estimation methods. Bayesian methods provided more stable estimates, especially when using the Jeffrey's method, while MLE and OLS methods showed convergent behavior as the value of ϕ increased. This reflects the flexibility of Bayesian estimation in handling parameter uncertainties, compared to traditional methods.

5.2 Second: Simulation study

A simulation study was conducted using the simulation option in IBM SPSS Statistics 27 within the Monte Carlo simulation framework, where data were generated based on an exponential distribution. The generated sample size was set at 230 observations to ensure adequate representation of the statistical characteristics of the actual data, including the mean and variance. To enhance the reliability of the results, 1000 simulation iterations were performed to assess the stability of the parameters and the efficiency of the estimation methods. The MSE criterion was adopted to compare the results between the original and simulated data, aiming to verify the accuracy of the estimation and the consistency of the results.

We assume that the generated time series follows the AR (1) model.

By applying the estimation methods mentioned in the theoretical aspect. Table 3 shows MSE for each estimation method:

Table 3 shows MSE and AIC for each estimation method

ϕ	MSE				AIC
	Ols	Mle	Jeffery	Bayes	
0	0.09	0.1101	0.0044	0.00000818	-1883.76
0.1	0.087	0.1058	0.0035	0.00001403	-1835.66
0.2	0.081	0.1015	0.0028	0.000107	-1883.37
0.3	0.053	0.0974	0.0021	0.000287	-1720.90
0.4	0.051	0.0933	0.0016	0.0000555	-1650.48
0.5	0.050	0.0893	0.0011	0.0009108	-1567.17
0.6	0.0041	0.0854	0.000698	0.0014	-1465.15
0.7	0.0035	0.0816	0.000393	0.0019	-1333.53
0.8	0.0027	0.0778	0.000174	0.0025	-1147.76
0.9	0.00142	0.0742	0.0000436	0.0032	-829.05

From Table 3, for the simulation data, AIC and MSE parameters for the AR(1) model were calculated under different values of the parameter ϕ . The results showed that the lowest AIC value was -1883.76 at $\phi=0$, and the same condition recorded

the lowest MSE value, indicating that the best-fitting and most accurate model is at $\phi=0$. However, it was observed that the AIC value at $\phi = 0.2$ was very close to the lowest value, demonstrating a convergence in performance between the two models.

These results indicate that the simulation data largely agree with the real data, exhibiting behavior close to random noise, with a low probability of time dependence. Comparing the results of the real data with the simulation data reveals a high degree of agreement, with both cases showing that the best model performance is achieved at $\phi=0$. This strengthens the reliability of the results and confirms that the studied time series does not have a significant time dependence. The convergence of AIC values between $\phi=0$ and some other values in the simulation data also indicates a limited sensitivity of the model to small changes in the parameter.

6. Conclusion

This research yielded a number of important results through time series analysis using the first-order autoregressive model AR(1), assuming that random errors follow an exponential distribution. The results, using both real data from Mosul and simulated data, showed the following:

1. The best model performance was achieved at $\phi=0$, where the lowest values for both the AIC and MSE criteria were recorded. This result indicates that the studied time series does not exhibit significant time dependence and tends toward behavior close to random noise, meaning that introducing the autoregressive component does not lead to a significant improvement in prediction accuracy.
2. The results also showed that increasing the value of ϕ leads to a deterioration in both the statistical fit quality and prediction accuracy, which confirms the unsuitability of higher-order models in this case.
3. On the other hand, a comparison of different estimation methods (Bayesian, Jeffrey, MLE, OLS) showed that Bayesian methods are characterized by higher flexibility and better stability in parameter estimation, especially under the assumption of a non-normal error distribution, while traditional methods exhibited convergent behavior but less flexibility.
4. The simulation data results also showed high agreement with the real data results, which enhances the reliability of the findings and confirms the validity of the methodological framework used in this study.
5. This research highlights the importance of using non-Gaussian distributions, such as the exponential distribution, in modeling time series that do not follow classical assumptions.

7. Recommendations

In light of the findings, several future research directions can be suggested to further develop this field.

1. The study could be expanded to include more complex models, such as ARIMA or non-linear models, to investigate whether these models can capture time patterns not apparent in the AR(1) model.
2. The effect of using other non-normal error distributions, such as the Gamma or Student's t-distribution, could be investigated and compared. Its performance with exponential distributions is particularly noteworthy in data with heavy tails or asymmetry.
3. The proposed methodology can be applied to other types of empirical data, such as economic or medical data, to verify the generality of the findings.

Note: Declaration of generative AI and AI-assisted technologies in the writing process. During the preparation of this work, the author(s) used ChatGPT in order to improve the readability and language of the work. After using this service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

8. Acknowledgements

The authors would like to express their appreciation to Northern Technical University for supplying the materials for this study.

9. Declarations

9.1 Ethics approval and consent to participate

Not applicable.

9.2 Consent for publication

Not applicable.

9.3 Availability of Data and Materials

The data used in this study are publicly available online from open-access sources

9.4 Conflicts of interest

The authors declare that there is no conflict of interest

9.5 Funding

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

10. References

- [1] Al-Naqeeb ,Abdul Khaliq Abdul Jabbar, Hazem Mansour Korkis, & Alaa Majid Hamad .(2010) .Estimating the general exponential distribution parameters using the simulation method. Journal of Economics and Administrative Sciences, 16(57), 21-21. <https://doi.org/10.33095/jeas.v16i57.1440>
- [2] Amin, A. A. (2025). Bayesian modeling and forecasting of seasonal autoregressive models with scale-mixtures of normal errors. *Computational Statistics*, 40(7), 3453-3475. <https://doi.org/10.1007/s00180-025-01617-2>
- [3] Balakrishna, N. (2021). *Non-Gaussian autoregressive-type time series*. Singapore: Springer. <https://doi.org/10.1007/978-981-16-8162-2>
- [4] BS Everitt, DC Howell ,(2005)," Least Squares Estimation", Encyclopedia of Statistics in Behavioral Science, Volume 2, pp. 1041–1045. DOI: 10.1002/0470013192.bsa199
- [5] BuHamra, Sana and Smaoui, Nejib and Gabr, Mahmoud (2003):" The Box--Jenkins analysis and neural networks: prediction and time series modeling" , Elsevier,Vo.27,no.10,p.p(805-815). doi:10.1016/S0307-904X(03)00079-9
- [6] Cheng Junsheng, Yu Dejie, Yang Yu (2006), " A fault diagnosis approach for roller bearings based on EMD method and AR model,Mechanical Systems and Signal Processing ,Volume 20, Issue 2, February 2006, Pages 350-362. doi:10.1016/j.ymssp.2004.11.002
- [7] Frieden, R., & Gatenby, R. A. (2010). Exploratory data analysis using Fisher information. Springer Science & Business Media. Doi:10.1007/978-1-84628-777-0
- [8] Ghosh,Jayantak, Delampady,Mohan , Samanta , Tapas , 2006 , " An Introduction to Bayesian Analysis Theory and Method " , Springer. https://doi.org/10.1007/978-0-387-35433-0_6
- [9] Harvey, A., & Luati, A. (2014). Filtering with heavy tails. Journal of the American Statistical Association, 109(507), 1112-1122.<https://doi.org/10.1080/01621459.2014.887011>
- [10] Hazem,Imad Aboudi , Adwar, Etaf Abdel Ahad (2011) : "Reliability estimates for the exponential distribution with two parameters", Al-Mansour Magazine Issue(15) 2011,p.p (131-159). DOI: Not available
- [11] Hesamian, G., Torkian, F., Johannssen, A., & Chukhrova, N. (2023). An Exponential Autoregressive Time Series Model for Complex Data. *Mathematics*, 11(19), 4022. <https://doi.org/10.3390/math11194022>
- [12] Hossain, S. A. (2018). Estimating the Parameters of a Generalized Exponential Distribution. *Journal of Statistical Theory and Applications*, 17(3), 537-553. doi.org/10.2991/jsta.2018.17.3.9
- [13] Hung, E., Mantziou, A., & Reinert, G. (2025). A Bayesian mixture model for Poisson network autoregression. *Social Network Analysis and Mining*, 15(1), 70. <https://doi.org/10.1007/s13278-025-01485-0>
- [14] Ibazizenm Mohamed & Fellag, Hocine (2003): Bayesian estimation of an AR(1) process with exponential white noise , *Statistics: A Journal of Theoretical and Applied Statistics*, 37:5, 365-372 <http://dx.doi.org/10.1080/0233188031000078042>
- [15] Jiao, J., Venkat, K., Han, Y., & Weissman, T. (2017). Maximum likelihood estimation of functionals of discrete distributions. *IEEE Transactions on Information Theory*, 63(10), 6774-6798. doi: 10.1109/TIT.2017.2733537.
- [16] Khan, S. (2020). ARIMA model for accurate time series stocks forecasting. *International Journal of Advanced Computer Science and Applications*. DOI: 10.14569/IJACSA.2020.0110765
- [17] Kubuafor, E., Baidoo, D., Okeke, O. J., Amevor, R., Arhin, G., & Korley, J. T. (2025). Evaluation of Time Series Forecasting Models for Predicting Lung Cancer Mortality Rates in the United States: A Comparison with Altuhaifa (2023) Study. *arXiv preprint arXiv:2508.16052*. <https://doi.org/10.48550/arXiv.2508.16052>
- [18] Mahdavi, A., & Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13), 6543-6557. <https://doi.org/10.1080/03610926.2015.1130839>
- [19] Mark, C., Metzner, C., & Fabry, B. (2014). Bayesian inference of time varying parameters in autoregressive processes. *arXiv preprint arXiv:1405.1668*. <https://doi.org/10.48550/arXiv.1405.1668>
- [20] McElreath, R. (2016). Statistical Rethinking: A Bayesian Course with Examples in R and Stan. *Stat. Rethink. Bayesian Course Ex. R Stan Authorfunder Rights Reserv. No Reuse Allow. Permis. Multimodal Actions Huperzine Proc. Natl. Acad. Sci. US Am*, 110, E746-755. DOI: 10.1201/9781315372495
- [21] Norton, Matthew; Khokhlov, Valentyn; Uryasev, Stan (2021). "Calculating CVaR and bPOE for common probability distributions with application to portfolio optimization and density estimation" ,*Annals of Operations Research*. Springer. 299 (1–2): 1281–1315. <https://doi.org/10.1007/s10479-019-03373-1>
- [22] Oyinloye, A. A., Ayodele, O. J., & Abifade, V. O. (2023). Modeling power exponential error innovations with autoregressive process. *IJMSS*, 11(3), 13. <https://doi.org/10.37745/IJMSS.13/VOL11N21321>
- [23] Triana, Y., & Purwadi, J. (2019, November). Exponential distribution parameter estimation with bayesian self method in survival analysis. In *Journal of Physics: Conference Series* (Vol. 1373, No. 1). IOP Publishing.

DOI:10.1088/1742-6596/1373/1/012050

- [24] Wackerly, Dennis; Mendenhall, William; Scheaffer, Richard L. (2008). Mathematical Statistics with Applications (7 ed.). Belmont, CA, USA: Thomson Higher Education. ISBN 978-0-495-38508-0
DOI: Not available
- [25] World Bank open data <https://data.worldbank.org/>
- [26] Yang, K., & Wang, D. (2017). Bayesian estimation for first-order autoregressive model with explanatory variables. Communications in Statistics-Theory and Methods, 46(22), 11214-11227.
<https://doi.org/10.1080/03610926.2016.1260736>
- [27] Zhu, D., & Galbraith, J. W. (2010). Modeling and forecasting with Student-t distributions. *International Journal of Forecasting*, 26(3), 660–670. <https://doi.org/10.1016/j.ijforecast.2009.09.008>

طرائق تقدير معلمة نموذج الانحدار الذاتي AR(1) لبيانات تتبع التوزيع الاسي

الهام عبدالكريم حسين⁽¹⁾، احمد نجم شيت الشلاوي⁽²⁾، عمار حازم طه⁽³⁾

^(1,2) الجامعة التقنية الشمالية، الكلية التقنية الادارية، الموصل، العراق

⁽³⁾ الجامعة التقنية الشمالية، معهد الادارة التقني، الموصل، العراق

المستخلص:

هدف هذه الدراسة إلى مقارنة أداء طرق التقدير المختلفة لنموذج الانحدار الذاتي من الدرجة الأولى (AR(1)) بافتراض أن الأخطاء العشوائية لا تتبع التوزيع الطبيعي، بل التوزيع الاسي. استخدمت أربع طرق تقدير: بايزي، وجيفري، والامكان الاعظم (MLE)، والمربعات الصغرى (OLS)، لتقييم كفاءتها في تقدير معالم النموذج.

طبقت الدراسة على بيانات حقيقية تمثلت بدرجات الحرارة لمدينة الموصل، بالإضافة إلى بيانات محاكاة تم توليدها للتحقق من موثوقية النتائج. استندت المقارنة بين الطرق إلى متوسط مربع الخطأ (MSE) ومعيار معلومات أكايكي (AIC). أظهرت النتائج أن أفضل أداء للنموذج تحقق عند قيمة $\phi=0$ ، حيث سجلت أدنى قيم لكل من AIC و MSE في كل من البيانات الحقيقية والمحاكاة، مما يشير إلى عدم وجود اعتماد زمني كبير في السلاسل المدروسة. كما أظهرت النتائج أن الأساليب البايزية تتمتع بمرونة واستقرار أعلى مقارنة بالأساليب التقليدية، خاصة في ظل افتراض التوزيع الاسي للأخطاء.