Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/130316

Measures of Sample Skewness and Kurtosis for AR (1) Model with Missing Data with Applications in Economic Data

*Mahmoud M. Abdelwahab**

Department of Basic Sciences, Higher Institute of Administrative Sciences, Osim, Cairo, Egypt

Received: 11 Jul. 2023, Revised: 22 Oct. 2023, Accepted: 30 Jan. 2024.
Published online: 1 May 2024

Abstract: Many statistical functions require that a distribution be normal or nearly normal. In time series models, data is assumed to follow a normal distribution. Two numerical measures of shape skewness and kurtosis can be used to test for normality. Autoregressive models are one of the models used to estimate time series data. The missing data in time series models affect in terms of their following a normal distribution. In this paper, derive the moments, skewness and kurtosis of AR (1) Model with Missing data without constant term and used the parameter of the model by using ordinary least squares (OLS), Yule Walker (YW) and weighted least squares (WLS). Moreover, Monte Carlo simulation at various sample sizes and different proportions of missing data for comparative study skewness and kurtosis between ordinary least squares (OLS), Yule-Walker (YW) and weighted least squares (WLS). In addition, time series real data with missing data was measures of kurtosis could be used to compare between these methods.

Keywords: The moment, Time series, AR (1) Model, Missing data, Skewness and Kurtosis.

1 Introduction

Jones (1962) measured the case of periodic sampling where the observed data consist of repeated in two groups, the first group consecutive observations followed by second group missed observations. Following Parzen (1963) introduced time series model with missing observations as a specific case of stationary process. He observed that the data $\left\{y_1, y_2, \cdots\cdots, y_p\right\}$ can be represented as,

$$x_k = \eta + \sum_{i=1}^{n} \rho_i x_{k-i} + \varepsilon_k \qquad \varepsilon_k \sim iid(0, \sigma^2) \qquad (1)$$

$$y_k = a_k x_k \qquad k = 1,2, \cdots\cdots, n \qquad (2)$$

Under the following assumptions:

$$\mathrm{E}\left(\varepsilon_k\right) = 0 \qquad (\mathrm{I})$$

$$\mathrm{E}\left(y_k\right) = a_k \mathrm{E}\left(x_k\right) \qquad (\mathrm{II})$$

$$\mathrm{E}\left(x_{k-i}\varepsilon_k\right) = 0 \qquad (\mathrm{III})$$

Where, $\{a_k, k = 1,2, \cdots\cdots, n\}$ represent the state of observation such that,

*Corresponding author e-mail: mma_fars@yahoo.com

$$a_k = \begin{cases} 1 & \text{if } x_k \text{ is obsrved,} \\ 0 & \text{O / W} \end{cases}$$

Typical examples of $\{a_k\}$ are a stochastic markov process and periodic deterministic case. If $\{x_k\}$ is a stationary process, it means the start from the infinite past.

Scheinok (1965) and Bloomfield (1970) investigated the case where x_k is observed when a ''success'' is achieved on a Bernoulli trial. These references concentrate on non-parametric spectral analysis of time series with missing values. Dunsmuir and Robinson (1981) suggested the estimator of AR (1) model with missing observations without constant by using Yule-Walker method. In a time series regression with missing observations, Shively (1993) created several tests for autoregressive disturbances, where the disturbance terms are produced by the AR(1) and AR(p) processes with a potential seasonal component. Takeuchi (1995) obtained another estimator for AR (1) model with missing observations without constant by using OLS method. Park and Fuller (1995) introduced a weighted symmetric estimator for AR (1) model. Vrbik (2005) described a simple technique for computing the first few moments of a large class of sample statistics related to the first-order autoregressive model with normally distributed error terms. Youssef (2006) presented a performance of alternative predictors for the unit root process. Abdelwahab, et al. (2012) proposed a general form of moments for AR (1) and AR (2) model with and without constant term. Abdelwahab (2016) presented estimator of AR (1) and AR (2) model with missing observations with and without constant term and studied the properties of this estimator. Saadatmand, et al. (2017) examined two approaches to estimating the missing value with respect to Pitman's measure of proximity and considered estimate of a missing value for AR (1) with exponential innovations. Abdelwahab and Issa (2019) obtained the Forms of the moments of AR(P) model with missing observations and they suggested general form of mean and variance for this model. Issa and Abdelwahab (2020) derived OLS estimator for AR (1) panel data model with missing data. Issa (2021) suggested a new form of the estimator of AR (1) model with constant term with missing observations by using Ordinary Least Squares (OLS) method. In (2022) used another method to Estimate the parameter for AR(1) Model With Incomplete Data by using weighted Least Squares (OLS) method and he studied the properties of OLS and WLS estimators are discussed. Enany, et al. (2023) introduced a closed form estimator for ρ in case of missing observations using maximum likelihood (ML) of AR(1) model without constant term with missing observations when y_0 is random.

The aims of research is to derive the moment, skewness and kurtosis of AR (1) Model with missing data without constant term. In addition, To compare skewness and kurtosis with other estimate techniques at various sample sizes and missing data percentages, a Monte Carlo simulation analysis was done. The remainder of the essay is structured as follows: Provide the estimation techniques for the AR (1) model with missing data in section (2). In Section (3), the moment of AR (1) model with missing data without constant term and derivation the skewness and kurtosis of AR (1) with missing data without constant term. In section (4), simulation studies are carried out to compare between the skewness and kurtosis with other estimation methods OLS estimator, YW estimator, and weighted least squares (WLS). In section (5) applied studies by using real data. In section (6), conclusion of the theory and simulation study has been presented.

2 Materials and Methods

In this section. we will describe the three different estimation methods for the AR (1) Model with Missing Data used in this paper.

Yule-Walker Estimator of AR (1) Model

Dunsmuir and Robinson (1981) suggested estimator of AR (1) model with missing observations without constant term be using Yule-Walker method when |ρ|<1 is given by

$$\hat{\rho}_{YW} = \frac{\sum_{k=2}^{n} y_k y_{k-1} / \sum_{k=2}^{n} a_k a_{k-1}}{\sum_{k=1}^{n-1} y_{k-1}^2 / \sum_{k=1}^{n-1} a_k} \quad (3)$$

Least Square Estimator of AR (1) Model

Takeuchi (1995) derived the estimator of AR (1) model with missing observations with constant term be using Least Square method when $|\rho| < 1$ is given by

$$\hat{\rho}_{ols} = \frac{\sum_{k=2}^{n} a_k y_k y_{k-1}}{\sum_{k=2}^{n} a_k y_{k-1}^2} \tag{4}$$

Weighted Least Square Estimator of AR (1) model

Issa (2022) derived the estimator of AR (1) model with missing data by using weighted Least Square method when $|\rho| < 1$ *and* $|\gamma| < 1$ is given by:

$$\hat{\rho}_{WLS} = \frac{\sum_{k=2}^{n} w_k a_k y_k y_{k-1}}{\sum_{k=2}^{n} w_k a_k y_{k-1}^2} \tag{5}$$

Where

$$w_k = |y_{k-1}|^{-2\gamma} \tag{6}$$

$$w_k = |y_{k-1}|^{\gamma-1} \tag{7}$$

Where, γ is the coefficient of heteroscedasticity according to Brewer (2002) which is used in regression models. We will be revised γ in case of AR(1) model with missing observations to minimize the residual sum of squares.
By substituting equation (6) in equation (5), to get;

$$\hat{\rho}_{WLS.I} = \frac{\sum_{k=2}^{n} a_k |y_{k-1}|^{-2\gamma} y_k y_{k-1}}{\sum_{k=2}^{n} a_k |y_{k-1}|^{-2\gamma} y_{k-1}^2} \tag{8}$$

When $\gamma = 0$, then $\hat{\rho}_{WLS.I} = \hat{\rho}_{OLS}$
Moreover, Substitute the equation (7) in equation (5), then;

$$\hat{\rho}_{WLS.II} = \frac{\sum_{k=2}^{n} a_k |y_{k-1}|^{\gamma-1} y_k y_{k-1}}{\sum_{k=2}^{n} a_k |y_{k-1}|^{\gamma-1} y_{k-1}^2} \tag{9}$$

When $\gamma = 1$, then $\hat{\rho}_{WLS.II} = \hat{\rho}_{OLS}$

3 Results and Discussion

The moment of the autoregressive models with missing data. Will be introduced in the following section

- **The First Moment About Zero of AR (1) Model**

The AR (1) with the missing data by using equation (1) takes the form:

$$y_k = \rho a_k x_{k-1} + a_k \varepsilon_k, \quad k = 1,2,\cdots\cdots,n \tag{10}$$

By taking the expectation of equation (10), we will get

$$E(y_k) = \rho E(a_k x_{k-1}) + E(a_k \varepsilon_k)$$

By using assumption (I) in model, we will get

$$E(y_k) - \rho E(y_k) = 0$$
$$(1-\rho)E(y_k) = 0$$

$$E(y_k) = 0 = \mu_{1.AR(1)} \tag{11}$$

- **The Second Moment About Zero of AR (1) Model**

We will square the equation (10) and taking the expectation, we will get

$$y_k^2 = [\rho a_k x_{k-1} + a_k \varepsilon_k]^2$$

$$y_k^2 = (\rho a_k x_{k-1})^2 + 2a_k \varepsilon_k \rho x_{k-1} + a_k^2 \varepsilon_k^2$$

$$E(y_k^2) = \frac{a_k \sigma^2 - a_k \sigma^2 \rho}{(1-\rho^2)(1-\rho)}$$

$$E(y_k^2) = \mu_{2.AR(1)} = \frac{a_k \sigma^2}{1-\rho^2} \quad (12)$$

- **The Third Moment about Zero of AR (1) Model.**

Now, the third moment of y_k will be derived

$$E[y_k^3] = E[\rho a_k x_{k-1} + a_k \varepsilon_k]^3$$

$$E[y_k^3] = \rho^3 E[a_k^3 x_{k-1}^3] + 3a_k \rho^2 E[a_k^2 x_{k-1}^2] E[\varepsilon_k] + 3a_k \rho E[a_k x_{k-1}] E[\varepsilon_k^2] + a_k^3 E[\varepsilon_k^3] \quad (13)$$

From the assumptions (I) and (II) of the model, the expected value of the model can be calculated

$E(y_k^3) = E(a_k x_{k-1}^3)$

$E(y_k) = E(y_{k-1}) = \mu_{1.AR(1)},$

$E(y_k^2) = \mu_{2.AR(1)}$

We get:

$$\mu_{3.AR(1)} = E[y_k^3] = 0 \quad (14)$$

Where

$$\mu_{1.AR(1)} = 0 \text{ and } \mu_{2.AR(1)} = \frac{a_k \sigma^2}{1-\rho^2}$$

- **The Fourth Moment About Zero of AR (1) Model.**

The fourth moment of y_k will be derived by taking the expectation of equation (10), The fourth moments can be calculated as follows:

$$E[y_k^4] = E[\rho a_k x_{k-1} + a_k \varepsilon_k]^4$$

$$E[y_k^4] = \rho^4 E[a_k^4 x_{k-1}^4] + a_k^4 E[\varepsilon_k^4] + 4a_k \rho E[a_k x_{k-1}] E[\varepsilon_k^3] + 4\rho^3 E[a_k^3 x_{k-1}^3] E[\varepsilon_k] + 6a_k \rho^2 E[a_k^2 x_{k-1}^2] E[\varepsilon_k^2] \quad (15)$$

By using the assumption (I) and (II), Equation (15) can be rewritten as:

$$E[y_k^4] = \mu_{4.AR(1)} = \frac{a_k \sigma^4 + 6a_k \sigma^2 \rho^2 \mu_{2.AR(1)}}{1-\rho^4} \quad (16)$$

Lemma (3.1) The Skewness of AR (1) with missing data

The Skewness is measures often used to describe a probability distribution by:

$$\alpha_{3.AR(1)} = \frac{\mu_{3.AR(1)}}{\mu_{2.AR(1)}^{\frac{3}{2}}}, \quad (17)$$

By substituting equation (12) and (14) in equation (17), we get:

$$\alpha_{3.AR(1)} = 0 \quad (18)$$

Lemma (3.2) The Kurtosis of AR (1) with missing data

- The kurtosis can be defined as follows:

$$\alpha_{4.AR(1)} = \frac{E(y_k - \mu)^4}{[E(y_k - \mu)^2]^2} = \frac{\mu_{4.AR(1).trad}}{\mu^2_{2.AR(1)}}, \quad (19)$$

By substituting equation (12) and (16) in equation (19), we get:

$$\alpha_{4.AR(1)} = \frac{\left[a_k\sigma^4 + 6a_k\sigma^2\rho^2\mu_{2.AR(1)}\right]}{(1-\rho^4)\left[\mu_{2.AR(1)}\right]^2} \quad (20)$$

$$\alpha_{4.AR(1)} = \frac{\left[a_k\sigma^4 + 6a_k\sigma^2\rho^2(\frac{a_k\sigma^2}{1-\rho^2})\right]}{(1-\rho^4)\left[\frac{a_k\sigma^2}{1-\rho^2}\right]^2} \quad (21)$$

$$\alpha_{4.AR(1)} = \frac{\left[a_k\sigma^4(1-\rho^2)+6a_k\sigma^4\rho^2\right]}{(1-\rho^2)(1-\rho^4)\frac{a_k\sigma^4}{\left[1-\rho^2\right]^2}} \quad (22)$$

$$\alpha_{4.AR(1)} = \frac{1+5\rho^2}{1+\rho^2} \quad (23)$$

By using equations (3, 4,8 and 9) in equation (23). We can get the kurtosis of AR (1) model by using the estimator of OLS, YW, WLS.I and WLSII.company.

4 Simulation study

In this section, we analyze how the Monte Carlo simulations-based kurtosis for OLS, YW, WLS.I and WLS.II behave in finite samples. A Comparison between OLS, YW, WLS.I and WLS.II methods for AR (1) is presented using Bias and Relative Bias of Kurtosis [The setting of model and the results of the simulation study are discussed] To perform the simulation, the model is needed to construct as follows:

1- AR (1) model without constant term is generated. The errors are generated ~ IIDN (0, 1), and the autoregressive parameter ρ is chosen to be (0.1,-0.1,0.3, -0.3, 0.5, -0.5) and γ is chosen to be (0.2).based on (Issa 2022)

2- The values of sample size n are equal to 20, 50, 100 and 250 to represent small, medium and large samples of time series.

3- To verify that the data follows a normal distribution, we randomly generate different percentages of missing values equals to [(5 to 10), (15 to 20) and (25 to 30)]

4- All replications in Monte Carlo attempts involved 10000.

We compare the p value of Kurtosis for ordinary least squares estimator (OLS) defined in equation (4), with the Yule Walker estimator (YW) which is defined in equation (3) and weighted least squares estimators (WLSI and WLSII) defined in equations (7 and 9). By using (0.05) Level of significance to test the claim Kurtosis=3 and the Relative Bias (RB) based on ordinary least squares estimator (OLS) is comparison standard for as methods. The results of simulation study explained in different cases of percentages of missing data: case (I) (5 to 10) missing data, case (II) (15 to 20) missing data and case (III) (25 to 30) missing data [When $\gamma = 0.2$ (with different values of γ do not affect in statistically significant for p-value)].

Case (I): (5 to 10) missing data

- From table (1) for small, medium and large sample size, A p-value less than 0.05 is statistically significant. It indicates strong evidence against the null hypothesis. Therefore, we reject the null hypothesis, and accept the alternative hypothesis, except the case ($n = 20$ and $\rho = -0.5$) A p-value more than 0.05 is not statistically significant by using YW, WLS.I and WLS.II methods. Then the missing values affect the shape of the distribution in coefficient of kurtosis for all methods used when presenting missing from (5 to 10) percentage and compare the RB of these estimators as follows:

In case small sample sizes, and different values of (ρ), the (OLS) estimator has the best RB, but when ρ=- 0.3, the (OLS) has the worst RB. In case medium sample sizes, and different values of (ρ), the (OLS) estimator has the best RB, but when ρ= 0.1, the (OLS) has the worst RB and when ρ=-0.1and -0.3, the (YW and WLSI) has the best RB. In case large sample sizes, and different values of (ρ), the (OLS) estimator has the worst RB, but when ρ= 0.1, the (OLS) has the worst RB and when $\rho = \pm 0.5$, the (YW and WLSI) has the best RB.

Case (II): (15 to 20) missing data

- From table (2) for small, medium and large sample size, A p-value less than 0.05 is statistically significant. Then the missing values affect the shape of the distribution in coefficient of kurtosis for all methods used when presenting missing from (15 to 20) percentage. And compare the RB of these estimators as follows:

In case small and medium sample sizes, and different values of (ρ), the (OLS) estimator has the best RB. In case large sample sizes, and different values of (ρ), the (WLSI and WLSII) estimator has the best RB where n=250 and the (YW and WLSI) has the best RB where n =100 but when ρ= -0.5 and n=100, the (OLS) has the best RB.

Case (III): (25 to 30) missing data

- From table (3) for small sample size, A p-value less than 0.05 is statistically significant, except the case ($\rho = -0.3$) A p-value more than 0.05 is not statistically significant by using WLS.II method. and compare the RB of these estimators as follows:

In case small sample sizes, and different values of (ρ), the (OLS) estimator has the best RB where $\rho = 0.1$ and 0.3 but when $\rho = \pm 0.5$ the (OLS) estimator has the worst RB and n=100, the (YW) has the best RB where$\rho = -0.1 \; and - 0.3$.

- From table (3) for medium sample size, A p-value less than 0.05 is statistically significant, except the case ($\rho = -0.5$) A p-value more than 0.05 is not statistically significant by using OLS, YW, WLS.I and WLS.II methods. and compare the RB of these estimators as follows:

In case medium sample sizes, and different values of (ρ), the (OLS) estimator has the best RB where $\rho = 0.1, 0.3 \; and - 0.5$ and n=50 but when $\rho = 0.5 \; and - 0.3$ the (YW) estimator has the best RB and when $\rho = -0.1$ the (WLSI) estimator has the best RB. Where n=100 the (YW) has the best RB where $\rho = \pm 0.5$, the (WLSI) has the best RB where $\rho = \pm 0.3$ and the (OLS) has the worst RB where $\rho = \pm 0.1$.

- From table (3) for large sample size, A p-value less than 0.05 is statistically significant, except the case ($\rho = -0.5$) A p-value more than 0.05 is not statistically significant by using YW method. And compare the RB of these estimators as follows:

In case large sample sizes, and different values of (ρ), the (WLSII) estimator has the best RB in $\rho = \pm 0.1 \; and = \pm 0.3$ but the (OLS) estimator has the best RB in $\rho = \pm 0.5$.

5 Applications of Real Data

To clarify the extent of the impact of missing data in time series model as the data follow a normal distribution. Using time series data for the US unemployment rate recorded obtained from Bureau of Labor Statistics, US Department of Labor, accessed June 5, 2020, choosing Unemployment Rate icon for Historical Data by year from 1960 to 2019.

The unemployment rate (UR) data series can plotted and shows the results of the stationarity test (ADF test) on the data by using EViews software

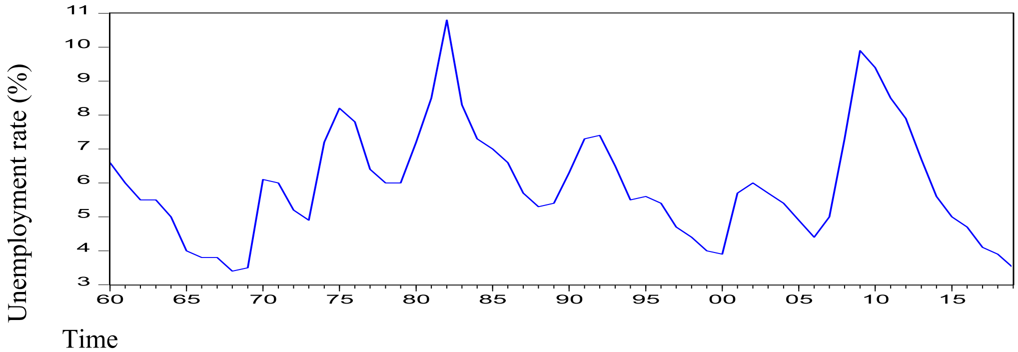

Fig.1: US unemployment rate from 1960 to 2019.

In addition, augmented Dickey-Fuller unit root test on US unemployment rate, the results are shown in the table (1)

Table (I) augmented Dickey-Fuller unit root test

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.583332	0.0091
Test critical values:	1% level	-3.548208	
	5% level	-2.912631	
	10% level	-2.594027	

*MacKinnon (1996) one-sided p-values.

The initial unemployment rate (UR) sequence is stationary when the significance level of 0.01, 0.05, and 0.1 is larger than ADF=-3.583332. The UR sequence continues to reject the null hypothesis with a low P value. The UR sequence is stationary, however, and Figure 2 shows the autocorrelation and partial autocorrelation function graphs for the UR series.

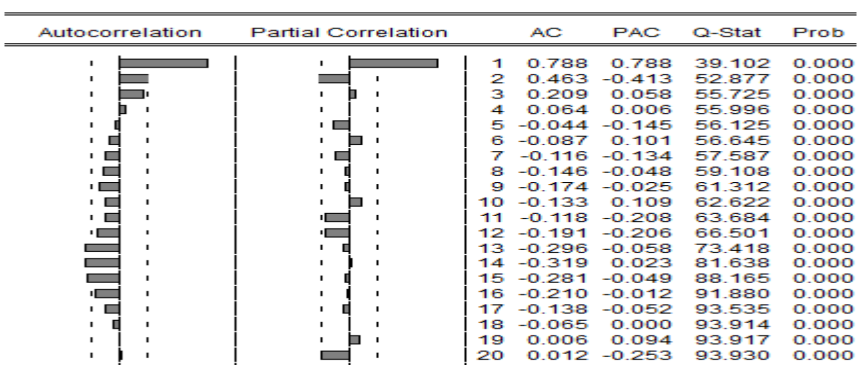

Fig. 2: Autocorrelation and partial autocorrelation function graphs of the UR series.

The autocorrelation coefficient of the UR sequence is substantially higher than zero when the lag order is 1, as seen in Figure 2, hence q can be taken as 1. The partial autocorrelation coefficient is significantly different from zero when the lag order is 1, and it is also significantly different from zero when the lag order is 1, therefore p=1 or p=2 can be taken into consideration. The range of p and q values is suitably modified in light of the subjective character of the assessment, and several ARMA (p, q) models are developed to produce a more accurate model. The order with 0, 1, and 2 in autoregressive moving average pre-estimation is applied to the processed sample data. The results of the ARMA (p, q) test for various parameters are shown in Table 3. AIC value, SC value, and regression S.E. are crucial selection factors for models. The best model is often chosen and ranked using the AIC and SC criterion. The higher the coefficient of determination, the smaller the residual variance, AIC, and SC values. It is superior to the ARIMA (P, I, Q) model that it relates to in table (2).

Table (2) Test results of ARMA (P, Q)

(P, Q)	Adjusted R-squared	AIC	SC	S.E. of regression
(0,1)	-3.466552	5.384421	5.454232	3.465767
(0,2)	-3.973685	5.517633	5.587444	3.657231
(1,0)	0.606106	2.986069	3.055880	1.029207
(1,1)*	0.654750	2.875880	2.980597	0.963563
(1,2)*	0.612302	2.987257	3.091974	1.021081
(2,0)	-0.032893	3.976129	4.045940	1.666637
(2,1)	0.599885	3.018668	3.123385	1.037303
(2,2)*	-0.032422	3.992187	4.096905	1.666257

It should be noted that although the AIC and SC values are typically used to define the appropriate ARMA model, they are insufficient for the optimum ARIMA model. The approach used in this work is to first build a model with the lowest AIC and SC values, after which the estimated data are subjected to parameter significance tests and residual randomness tests. If the test is successful, the model can be regarded as the best one; if not, the second smallest AIC and SC values are picked, and the appropriate statistical test is run. Until the ideal model is picked. A "*" was placed in this Table next to the model that failed both the residual randomness and parameter significance tests. Ultimately, it is recommended to use the ARMA (1, 0) model.

The ARIMA model's estimated results are as follows:

Estimation results of the ARMA model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.984333	0.021679	45.40432	0.0000
SIGMASQ	1.023959	0.166863	6.136532	0.0000

R-squared	0.612782	Mean dependent var	5.960000
Adjusted R-squared	0.606106	S.D. dependent var	1.639884
S.E. of regression	1.029207	Akaike info criterion	2.986069
Sum squared resid	61.43752	Schwarz criterion	3.055880
Log likelihood	-87.58207	Hannan-Quinn criter.	3.013376
Durbin-Watson stat	1.379347		
Inverted AR Roots	.98		

The final model in the LWC sequence is ARMA (1, 0), and Equation (1) demonstrates its particular shape. Under the equation, the matching estimate value's t-test statistic is shown in parentheses.

URt = 0.984333 URt-1

The t statistic of the model coefficients and its P value demonstrate that the parameter estimates of each explanatory variable are all statistically significant at the significance level of 0.01 for the model. The outcome of fitting the model to the UR data is shown in Figure 3. The upper and lower dotted lines reflect the model's fitted values and residuals, while the solid line represents the actual data.

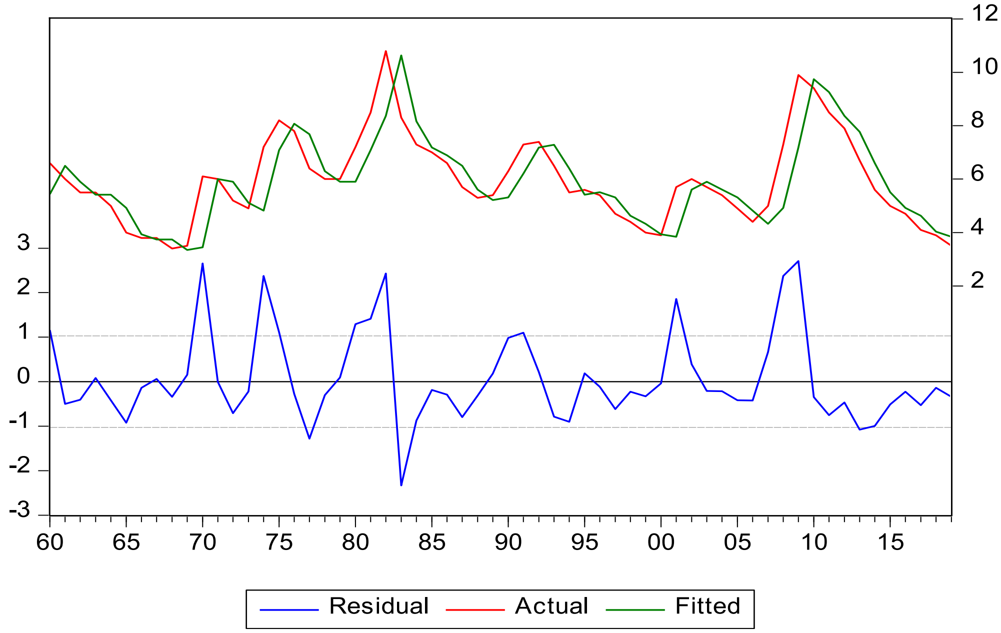

Fig. 3: Actual series, fitted series and residual series of the UR sequence.

After appropriate the ARIMA (1, 0) model, a white noise test is performed on the residual. The autocorrelation and partial autocorrelation function graphs for the residual series are shown in Figure 4. The residual is clearly white noise, showing that the model is correct.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.297	0.297	5.5479	
		2	-0.173	-0.286	7.4574	0.006
		3	-0.255	-0.125	11.708	0.003
		4	-0.087	-0.004	12.213	0.007
		5	-0.147	-0.240	13.664	0.008
		6	0.006	0.092	13.667	0.018
		7	0.051	-0.070	13.847	0.031
		8	0.017	-0.063	13.867	0.054
		9	-0.139	-0.138	15.285	0.054
		10	0.107	0.201	16.130	0.064
		11	0.118	-0.058	17.184	0.070
		12	-0.014	-0.062	17.199	0.102
		13	-0.234	-0.169	21.540	0.043
		14	-0.120	-0.052	22.698	0.045
		15	-0.060	-0.069	22.991	0.060
		16	0.008	-0.075	22.996	0.084
		17	-0.003	-0.087	22.997	0.114
		18	-0.010	-0.180	23.007	0.149
		19	0.094	0.201	23.813	0.161
		20	0.102	-0.098	24.778	0.168
		21	0.109	0.124	25.911	0.169
		22	0.047	-0.069	26.131	0.202
		23	-0.157	-0.151	28.620	0.156
		24	-0.216	-0.046	33.458	0.073
		25	-0.142	-0.168	35.593	0.060
		26	0.032	-0.016	35.706	0.076
		27	0.247	0.103	42.577	0.021
		28	0.144	-0.054	44.993	0.016

Fig.4: Autocorrelation and partial autocorrelation function graphs of the residual series.

Comparing the techniques of estimate for the parameter of the AR (1) model using (OLS), (YW), (WLSI), and (WLSII) in the presence of missing observations using the value of kurtosis and different values of γ. The results of table (3) showed that for the majority of missing observation percentages, the WLSI and WLSII technique produces kurtosis values that are higher. and different values of γ with respect to other methods.

Table (3) Kurtosis for different (methods, % of missing observations and γ)

γ	Estimators		percentages	
		5 to 10	15 to 20	25 to 30
0.1	OLS	1.198336	1.131494	1.147030
	YW	1.212363	1.166641	1.145087
	WLSI	1.207032	1.142901	1.153240
	WLSII	1.203134	1.150722	1.151579
0.2	OLS	1.291318	1.292417	1.299306
	YW	1.285654	1.290379	1.279982
	WLSI	1.342161	1.316161	1.322527
	WLSII	1.378974	1.324333	1.327920
0.3	OLS	1.281212	1.252516	1.112602
	YW	1.264566	1.201447	1.118709
	WLSI	1.279351	1.245705	1.117111
	WLSII	1.242882	1.202478	1.136324
0.4	OLS	1.217243	1.192098	1.251655
	YW	1.224087	1.185476	1.230913
	WLSI	1.318916	1.267309	1.272837
	WLSII	1.294004	1.254415	1.275301
0.5	OLS	1.196618	1.084760	1.202985
	YW	1.279259	1.112132	1.167792
	WLSI	1.199837	1.092342	1.207728
	WLSII	1.241871	1.115158	1.203658
0.6	OLS	1.168562	1.166945	1.106645
	YW	1.165875	1.182462	1.131156
	WLSI	1.147478	1.237511	1.185423
	WLSII	1.194755	1.196052	1.139742
0.7	OLS	1.200663	1.206101	1.294753
	YW	1.196806	1.218353	1.308418
	WLSI	1.202228	1.031471	1.161168
	WLSII	1.226253	1.223142	1.314782
0.8	OLS	1.156595	1.190031	1.200779

γ	Estimators		percentages	
		5 to 10	15 to 20	25 to 30
	YW	1.019129	1.072152	1.167914
	WLSI	1.154722	1.203272	1.246950
	WLSII	1.162164	1.205311	1.209746
0.9	OLS	1.166150	1.268261	1.275134
	YW	1.165650	1.274266	1.282757
	WLSI	1.725289	1.298057	1.399807
	WLSII	1.165729	1.277651	1.286213

6 Conclusions

In this article, Kurtosis and Skewness for AR (1) model with missing data without constant term has been derived by using OLS, YW and WLS methods. In addition, Monte Carlo simulation has been constructed using different methods of estimation (OLS, YW, WLSI and WLSII) Based on p-value and RB criteria. The results of simulation are divided to the presenting missing and $\gamma=2$: for different percentage a p-value less than 0.05 is statistically significant and the (WLSII) estimator has the best RB in large sample size. In small and medium sample sizes, the (OLS) estimator has the best RB followed by (YW) and (WLSI) methods.

Finally, the results of Monte Carlo simulation confirmed the missing data in time series models affect in terms of a normal distribution.

References

[1] Abdelwahab, M. M., El-Sayed, S. M., El-Sheikh, A. A. 2012. The joint moment generating function of the unit root estimator of AR (1). on JP Journal of Fundamental and Applied Statistics 2(1): 39-45.

[2] Abdelwahab, M.M. 2016. On Parameter estimation of time series models with missing observations. Unpublished PhD thesis. Cairo University Egypt.

[3] Abdelwahab, M. M. and Issa M. K. A, 2019. Forms of the moments of AR (P) model with missing observations. The Egyptian statistical journal 63(1): 39-44.

[4] Bloomfield, P. 1970. Spectral analysis with randomly missing observations. Journal of the Royal Statistical Society Series B, 32:369–380.

[5] Brewer, K. R. 2002. "Combining survey sampling inferences: weighing of Basu's elephants", Arnold: London and Oxford University Press.

[6] Dunsmuir, W. and Robinson, P. M. 1981a. Asymptotic theory for time series containing missing and amplitude modulated observations. Sankhyā: The Indian Journal of Statistics, Series A, 43(3): 260–281.

[7] Dunsmuir, W. and Robinson, P. M. 1981b. Parametric estimators for stationary time series with missing observations. Advances in Applied Probability. 13(1):129–146.

[8] El-Sayed, S. M., El-Sheikh, A. A., and Abdelwahab, M. M. 2016. Estimation of AR (1) models with missing values. Advances and Applications in Statistics. 49 (6):485-492.

[9] Enany, M.A., Issa M.K. and Gad A.A. 2023. Maximum likelihood estimation of unconditional AR (1) model with missing observations, Thailand statistic, 21(3): 607-615.

[10] H. Park and W. Fuller. 1995. Alternative estimators and unit root tests for the autoregressive process, Journal of Time Series Analysis 16: 415-429.

[11] Issa M. K. A. 2021. New Estimator for AR (1) Model with Missing Observations. Journal of University of Shanghai

for Science and Technology, 23(9): 147-159.

[12] Issa M. K. A. 2022. Weighted Least Squares Estimation for AR(1) Model With Incomplete Data, Mathematics and Statistics. 41(9):121-138.

[13] Issa M. K. A and Abdelwahab, M. M. 2020. Estimation of AR (1) panel data model with missing observations. Advances and Applications in Statistics. 63 (1):109-117.

[14] J. G. MacKinnon. 1996. Numerical distribution functions for unit root and cointegration tests, J. Appl. Econometrics 11: 601-618.

[15] Jones, R. H. 1962. Spectral analysis with regularly missed observations. The Annals of Mathematical Statistics. 33: 455–461.

[16] Parzen, E. 1963. On spectral analysis with missing observations and amplitude modulation. Sankhya Series A. 25. 383–392.

[17] Parzen, E. (ed) 1983. Proceedings of time series analysis of irregularly observed data. Lecture Notes in Statistics, New York: Springer Verlag.

[18] Saadatmand, A., Nematollahi, A. R., and Sadooghi-Alvandi, S. M. 2017. On the estimation of missing values in AR (1) model with exponential innovations. Communications in Statistics-Theory and Methods. 46(7): 3393-3400.

[19] Scheinok, P. A. 1965. Spectral analysis with randomly missed observations: The binomial case. The Annals of Mathematical Statistics. 36: 971–977.

[20] Shively T. S. 1993. Testing for autoregressive disturbances in a time series regression with missing observations, Journal of econometrics. 57(3): 233-255.

[21] Takeuchi, K. 1995. A comment on ″Recent Development of Economic Data Analysis ″at the 63^{rd} Annual Meeting of Japan Statistical Society.

[22] Vrbik.J. 2005. Moments of AR(1)-Model Estimators. Communications in Statistics - Simulation and Computation. 34 (3): 595 – 600.

[23] Youssef, A. H. 2006. A performance of alternative predictors for the unit root process. interstate journal, USA