 Volume-5, Issue-11, November - 2016 • ISSN No 2277 - 8160

Original Research Paper | Statistics

Use of Difference Equations in Time Series: A Study on Traffic Accidents Data

Senol Celik | Bingol University, Agricultural Faculty, Department of Animal Science, Bingol, Turkey

Zahide Kocabas | Ankara University, Agricultural Faculty, Department of Animal Science, Ankara, Turkey

ABSTRACT *The objective of this study is to estimate difference equations containing time series models. Solutions of difference equations are closely related to conditions of stationary time series models. A time series was created with the data from Turkish Statistical Institute related to the numbers of the traffic accidents happened in Turkey 1955-2012 period. It was seen in the study that the difference of the series itself and the series became stable after the second difference was taken. The most appropriate estimated model defined for the road traffic accidents is the one called ARIMA (0, 2, 3) which is an integrated moving average model with a third degree mobility. Difference homogeneous linear equation solving of ARIMA (0, 2, 3) model could be expressed as the trigonometric functions.*

KEYWORDS : Difference equation, ARIMA model, traffic accidents

Introduction

Linear difference equations play an important role in the time series models. The properties of these models often depend on the characteristics of the roots of these difference equations (Wei, 2006). Higher order difference equations arise quite naturally in economic analysis (Enders, 2010).

The study of the behavior of ARMA processes and their ACFs is greatly enhanced by a knowledge of difference equations, simply because they are difference equations. This topic is a useful in the study of the domain models and stochastic processes in general (Shumway and Stoffer, 2006). The objective of evaluation of time series models is based on analysis of difference equations containing stochastic (random) components, in order to forecast the observed phenomena in the future (Arneric and Kordic, 2010).

Difference equation is parametric stochastic process model that express the value of a variable as a function of its own lagged value and other variable. Arneric and Kordic (2010), estimated difference equation it will be examined whether Croatia is implementing a stable policy of exchange rates.

The purpose of this study, the time series model establish using linear difference equations and show the application.

Material and Methods

The study material consists of data regarding traffic accident between 1955 and 2012. The data were obtained from Statistical Indicators Book published by the Turkish Statistical Institute (2014). The data were used Dr. Senol Celik's (2013) PhD thesis in order to implement time series analysis.

In time series models, difference equations are quite typical to estimate second and higher order equations. The second order equation is as equality (1).

$$y_t - a_1 y_{t-1} - a_2 y_{t-2} = 0 \qquad (1)$$

Given the findings in the first order case, it should be suspected that the homogeneous solution has the form $y_t^h = A\alpha^t$. Substitution of this trial solution (1) provides

$$A\alpha^t - a_1 A\alpha^{t-1} - a_2 A\alpha^{t-2} = 0 \qquad (2)$$

Obviously, any arbitrary value of A is satisfactory. If you divide (2) by $A\alpha^{t-2}$, the problem is to find the values of α that respond

$$\alpha^2 - a_1\alpha - a_1 = 0 \qquad (3)$$

Solving this quadratic equation called the characteristic equation yields two values of α thecharacteristic roots. Using the quadratic formula, it found that the two characteristic roots are

$$\alpha_1, \alpha_2 = \frac{a_1 \pm \sqrt{a_1^2 + 4a_2}}{2} = (a_1 \pm \sqrt{d})/2 \qquad (4)$$

here d is the discriminant $(a_1^2 + 4a_2)$ (Enders, 2010).

If $a_1^2 + 4a_2 > 0$, d is real number and there are two different real characteristic roots. There are two separate solutions to the homogenous equation represented by $(\alpha_1)^t$ and $(\alpha_2)^t$.

$$y_t^h = A_1(\alpha_1)^t + A_2(\alpha_2)^t$$

If $a_1^2 + 4a_2 = 0$, d=0 and $\alpha_1 = \alpha_2 = a_1/2$. A homogenous solution is $a_1/2$. When d=0, there is a second homogenous solutions given by $t(a_1/2)^t$. To expression that

$$y_t^h = t(a_1/2)^t$$

is a homogenous solution.

If $a_1^2 + 4a_2 < 0$, d is negative so that the characteristic roots are imaginary. These characteristic roots

$$\alpha_1 = (a_1 + i\sqrt{-\text{d}})/2 \qquad \alpha_2 = (a_1 - i\sqrt{-\text{d}})/2$$

Here $i = \sqrt{-1}$. Homogenous solution as

$$y_t^h = \beta_1 r^t + \cos(\theta t + \beta_2)$$

Here β_1 and β_2 are arbitrary constant, $r = (-a_2)^{1/2}$ and the value of θ is chosen so as to satisfy $\cos(\theta) = a_1/\left[2(-a_2)^{1/2}\right]$

The trigonometric functions impart a wave like pattern to the path of the homogeneous solution; note that the frequency of the oscillations is determined θ (Enders, 2010).

Results

Time series relating to traffic accident data was analyzed for the period 1955 to 2012. For infer the trend clearer, autocorrelation (ACF) and partial autocorrelation functions (PACF) of the time series are investigated. ACF and PACF graphs of difference of the second series the traffic accident number were given in Figures 1 and 2, respectively.

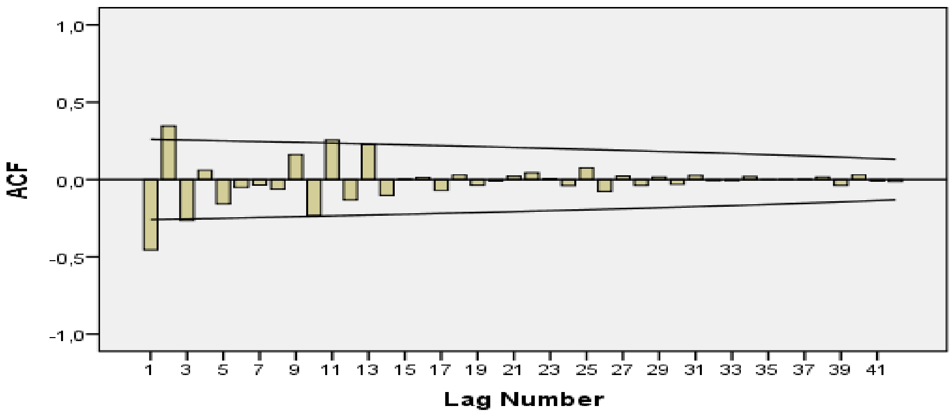

Figure 1. ACF chart of the second difference traffic accident number time series

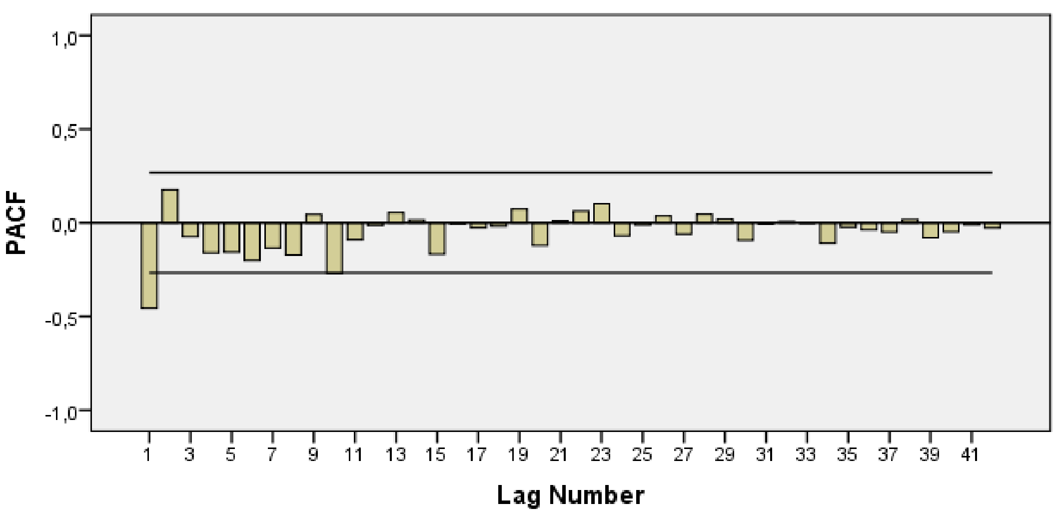

Figure 2. PACF chart of the second difference traffic accident number time series

The appropriate model was determined by examining ACF and PACF graphs of the second difference time series under the time series study. After the third lag in ACF graph (Table 1), magnitude of relationships decreased rapidly and thereafter was near-zero. The first lag obtained for PACF graph in Figure 2 was significant, and next values decreased slower. This means that moving average model was the fitting model for second difference time series. In the ACF graph, q=3 because the relationship concerning the third lag was significant. For taking the second difference of the time series, d=1 and p=0 were considered. In this case, the ARIMA (0,2,3) model is appropriate. Time series analysis results of the traffic accident data are given in Table 1.

Table 1. The significance of the appropriate model parameter estimation

Parameters	Coefficient	Standard error	t	Significant (p<0.05)
Constant	1661.968	831.205	1.999	0.051
Difference	2			
MA (1) (θ_1)	0.432	0.114	3.798	0.000
MA (2) (θ_2)	-0.360	0.127	-2.846	0.006
MA (3) (θ_3)	0.693	0.126	5.512	0.000

Model of equation according to series ARIMA (0,2,3),

$$(1-B)^2 X_t = (1-\theta_1 B - \theta_2 B^2 - \theta_3 B^3)e_t$$

is shaped. More specifically,

$$(1-B)^2 X_t = (1-0.432B + 0.36B^2 - 0.693B^3)e_t$$

is shaped.

$$X_t = (1-B)^{-2}(1-0.432B + 0.36B^2 - 0.693B^3)e_t$$

$(1-B)^{-2}$ is statement made by the Taylor series expansion (Arslan, 2015) and by multiply $(1-0.432B + 0.36B^2 - 0.693B^3)$. Equation (5) is obtained.

$$X_t = (1 + 1.568B + 2.496B^2 + \cdots)e_t \tag{5}$$

Equation (5) using Backshift operator,

$$X_t = e_t + 1.568e_{t-1} + 2.496e_{t-2} + \cdots \tag{6}$$

equation (6) is obtained.

Backshift operator, B, takes as input a time series and produces as output the series shifted backwards in time by j time unit in given equation (7) (Cooray, 2008).

$$B^j Y_t = Y_{t-j} \tag{7}$$

$e_t = m^t$ transformation is performed to Equation (6).

$$m^t + 1.568m^{t-1} + 2.496m^{t-2} = 0$$

$$m^{t-2}(m^2 + 1.568m + 2.496) = 0$$

$$(m^2 + 1.568m + 2.496) = 0 \tag{8}$$

when the equation (8) is solved,

$$\mathrm{d} = 1.568^2 - 4 * 1 * 2.496 < 0 \tag{9}$$

the roots are imaginary because of $d < 0$ in Equation (9). The obtained roots

$$m_1 = -0.784 + 1.3716i$$

$$m_2 = -0.784 - 1.3716i$$

are shaped.
This roots are found according to the equation (10).

$$Y_t = r^t(c_1 \cos\theta + c_2 \sin\theta) \tag{10}$$

where,

$$r = \sqrt{a^2 + b^2} = \sqrt{(0.784^2 + 1.3716^2)} = 1.58$$

$$tan\theta = \frac{1.3716}{0.784} = 1.749 \Rightarrow \theta = 60 = \frac{\pi}{3}$$

$$Y_t = c_1\,(-0.784 + \mathrm{i}\ 1.3716)^{\mathrm{t}} + c_2(-0.784 - \mathrm{i}\ 1.3716)^{\mathrm{t}}$$

$$Y_t = 1.58^t\left(\cos(\frac{\pi}{3})t - 0.91\sin(\frac{\pi}{3})t\right)$$

can be written.
For t=0 and Y_0=1; $c_1 = 1$,
For t=1 and Y_1=-0.455; $c_2 = -0.91$,
In that case,

$$Y_t = 1.58^t\left(\cos(\frac{\mu}{3})t - 0.91\sin(\frac{\mu}{3})t\right)$$

as linear difference equation is obtained.
Therefore, these solutions are not unique. In fact, for any two arbitrary constants c_1 and c_2 the linear combination of homogeneous solutions also solves second order difference equation

$$Y_t = A_1\,(-0.784 + \mathrm{i}\ 1.3716)^{\mathrm{t}} + A_2(-0.784 - \mathrm{i}\ 1.3716)^{\mathrm{t}}$$

Because these roots are less then unity in absolute value, the homogeneous solution is convergent.
From Figure 3 it can be seen that two homogeneous solutions converges, because of $0 < \theta_1$, θ_2, $\theta_3 < 1$. Figure 3 shows the time path of this solution for the case in which the arbitrary constants equal unity for 30 time units.

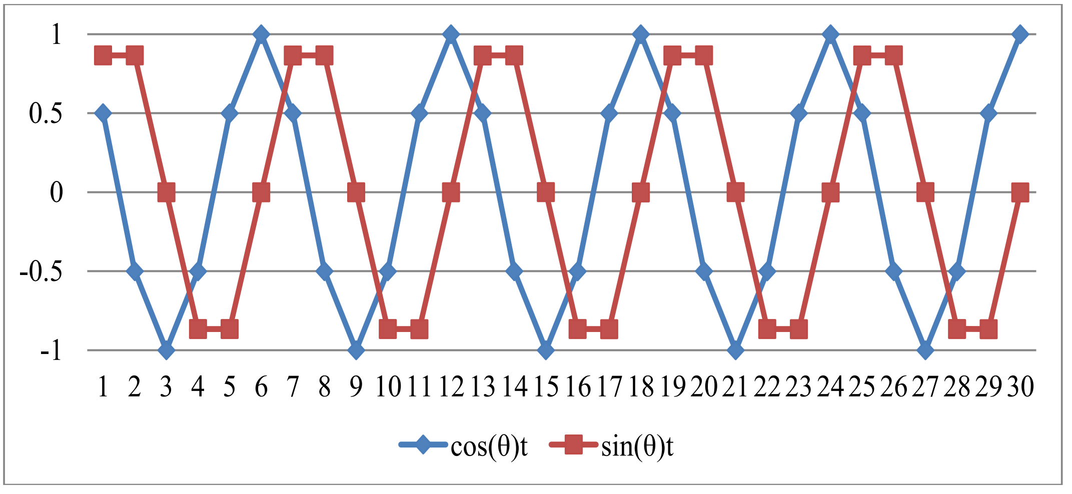

Figure 3. Convergence of the homogeneous solutions meanwhile $A_1 = A_2 = 1$

Conclusion

In this study, from estimated ARIMA(0,2,3) model of traffic accidents data it can be seen that homogeneous solution converges. It can also conclude that deterministic part of particular solution in terms of statistics corresponds to the expected value of the process. This model,

$$Y_t = 1.58^t(\cos(\pi/3)t - 0.91\sin(\pi/3)t)$$

is written as linear difference equation.

Volume-5, Issue-11, November - 2016 • ISSN No 2277 - 8160

References

1. Arnerić, J., Kordić, K. 2010. Econometric Approach to Difference Equations modeling of exchange rates changes. Croatian Operational Research Review (CRORR), 1:234-243.
2. Arslan, 2015. Matematiksel Analiz. Nobel Akademik Yayıncılık Eğitim Danışmanlık Tic. Ltd. Şti., Ankara, Turkey.
3. Celik, S. 2013. Time Series Analysis and its Applications to Data on Traffic Accidents. Ankara University Graduate School of Natural and Applied Sciences Department of Animal Science. Ph. D. Thesis.
4. Cooray, T. M. J. A. 2008. Applied Time Series Analysis and Forecasting. Narosa Publishing House PVT LTD., India
5. Enders, W. 2010. Applied Econometric Time Series.3rd Edition. John Wiley Sons, Inc, USA.
6. Sevüktekin, M.,Nargeleçekenler, M. 2010. Ekonometrik Zaman Serileri Analizi. EViews Uygulamalı. 3. Baskı. Nobel Yayın Dağıtım Tic. Ltd. Şti., Ankara.
7. Shumway, R. H., Stoffer, D. S. 2006. Time Series Analysis and Its Applications with R Examples. Springer Science Business Media, LLC, USA.
8. TÜİK, 2012. Türkiye İstatistik Yıllığı 2011. Yayın No: 3665, Türkiye İstatistik Kurumu Matbaası, Ankara.
9. TUIK, 2014. Statistical Indicators 1923-2013. Turkish Statistical Institute, Ankara, Turkey.
10. Wei, W. W. S. 2006. Time Series Analysis. Univariate and Multivariate Methods. Second Edition. Pearson Addison Wesley, USA.