Numerical Integral Equation Method of Average Run Length of Cumulative Sum Control Chart for Long Memory Process with ARFIMA Model

Wilasinee Peerajit, Yupaporn Areepong, and Saowanit Sukparungsee

Abstract—In this paper, approximation of the average run length (ARL) for long memory process by numerical integral equation (NIE) method on CUSUM control chart is presented. The Gauss-Legendre quadrature rule was used to approximate the NIE. In addition, we compared the efficiency of the ARL between NIE and explicit formulas for autoregressive fractionally integrated moving average (ARFIMA) model used exponential distribution of white noise. The results shown that NIE method alternative to explicit formulas because ARL values by using NIE and explicit formulas in good agreement.

Index Terms—Cumulative Sum (CUSUM) Control Chart, Autoregressive fractionally integrated moving average (ARFIMA), numerical integration equation (NIE), explicit formulas, average run length (ARL).

I. INTRODUCTION

The CUSUM control chart is based on the charting of cumulative sum of previous observations which allows us to use all the information about the process to make more accurate decisions. This control chart alternatively is also an effective method for quick shift identifications to detect small changes in the process control. CUSUM control chart was first proposed by Page [1] and has been studied by many researchers; in particular, see [2]-[8].

Average Run Length (ARL) is the expected number of observation taken from an in-control process until the control chart falsely signals out-of-control. ARL, as a common characteristic, is used to compare the performance of control charts. When ARL is large enough to keep the level of false alarms at an acceptable level, ARL is defined as acceptable.

The methods to evaluate ARL including Monte Carlo simulations (MC), Markov Chain approach (MCA) and numerical integral equation approach (NIE), have been discussed in literatures. Roberts [9] was the first person who introduced the ARL for EWMA control chart by using a simulation to estimate the ARL. This model was numerically evaluated by Robinson and Ho [10] who used Edgeworth expansion for the probability density function (pdf) and cumulative distribution function (cdf) of the process. The NIE was used by Crowder [11] to find the ARL for Gaussian distribution. Lucas and Saccucci [12] evaluated the ARL by using a finite-state MCA approximation. Recently, Sukparungsee and Novikov [13] used the Martigale approach to derive approximate analytical formulas of ARL and Average Delay (AD) in the case of Gaussian distribution and some Non-Gaussian distribution. Later, Areepong and Novikov [14] derived the explicit formulas of ARL and AD for EWMA control chart with Exponential distribution. The explicit formulas of ARL was recently presented by Mittel et al. [15] who used Fredholm Integral Equation for one-sided EWMA control chart with Laplace distribution and CUSUM control chart with Hyperexponential distribution.

The model of autoregressive fractionally integrated moving average (ARFIMA) processes have fractional differencing parameter ($d$) which are used to model a long-memory. These processes were introduced by Granger and Joyeux [16] and Hosking [17], a detailed description of long memory processes can be found in; e.g. [18]-[20]. The long-memory process is involved in a number of applications including finance and economics, environmental sciences and engineering. Control chart was selected to combine the long-memory process with time series. The control chart is necessary as it is a number of time series following ARFIMA models. Caballero et al. [21] performed a number of tests on the analysis of daily time series of mid-latitude near-surface air temperature by plotting long-range dependent processes. Pan and Chen [22] studied control chart for autocorrelated data using ARFIMA model to monitor the long memory air quality data in order to compare with ARIMA model. As a result, the residual control charts using ARFIMA models are more appropriate than those using ARIMA models. The exponential white noise was coordinate with time series. Jacob and Lewis [23] analyzed autoregressive moving average process order $(1, 1)$ denoted by ARMA $(1, 1)$ when observations are exponentially distributed with exponential white noise. The exponential white noise was also used to analyze the
autoregressive model as proposed by Mohamed and Hocine [24].

A number of studies using the explicit formulas or numerical integral equation method of ARL on control chart have been conducted. ARL was estimated using an integral equation technique and MCA to evaluate EWMA and CUSUM control charts in case the first order AR(1) process had additional random error. Suriyakat et al. [25] used an integral equation technique to solve the ARL for EWMA control chart for AR(1) process with exponential white noise. Busaba et al. [26] analyzed the explicit formulas of ARL for CUSUM control chart in case of, a stationary first order AR(1) process with exponential white noise. The numerical integral equations of ARL and solution to the numerically using the Gauss-Legendre numerical integration equations was derived by Petcharat et al. [27] when observations are first order of moving average process, MA(1), with exponential white noise. Recently, Phanyaem et al. [28] presented the numerical integration equation of ARL for CUSUM control chart with the p order autoregressive and the q order moving average, ARMA(p,q) process with exponential distribution white noise.

In this paper, we approximated ARL₀ and ARL₁ using the numerical integral equation (NIE) method with Gauss-Legendre quadrature rule on CUSUM control chart for long memory process with ARFIMA model. In section II, the characteristic of ARFIMA model with exponential distribution white noise is derived and proposed in the general form of ARFIMA model. Section III describes the characteristics of ARL for long memory process on CUSUM chart. The NIE for ARL will be described in section IV. The comparisons of ARL between NIE and explicit formulas were presented in section V.

II. ARFIMA MODEL WITH EXPONENTIAL WHITE NOISE

The model of an autoregressive fractionally integrated moving average process of a time series denoted by ARFIMA(p,d,q), with the p order autoregressive, the d order fractional difference and the q order moving average processes, are represented in the operator notation form. [16] and [17]

\[ \Phi(B) (1-B)^d X_t = \mu + \Theta(B) \xi_t, \quad \xi_t \sim \text{Exp}(\alpha). \] (1)

where \( \Phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) \) and \( \Theta(B) = (\theta_0 B - \theta_1 B^2 - \cdots - \theta_q B^q) \) are autoregressive and moving-average operators, respectively; \( B \) is the backward shift operator i.e., \( B^q(X_t) = X_{t-p} \) and \( \mu \) is a constant.

The term \((1-B)^d\) is generalized this expression to fractional differences defined by the binomial expansion. [17]

\[ (1-B)^d = \sum_{k=0}^{\infty} (-1)^k \frac{(d-1)\cdots(2)(1)}{k!} B^k. \] (2)

If \( d \in (-0.5, 0.5), \) then \( X_t \) is a stationary, invertible process. The ARFIMA(p,d,q) process characteristic of long memory when the parameter \( d \in (0,0.5), \) intermediate memory when \( d \in (-0.5, 0) \) and short memory when \( d = 0 \) corresponding to a standard ARMA process.

Then \( X_t \) is so-called the general form of the ARFIMA(p,d,q) process with exponential distribution white noise which is used in CUSUM control charts.

\[ X_t = \mu + \xi_t - \theta_1 \xi_{t-1} - \cdots - \theta_q \xi_{t-q} \frac{d(d-1)}{2!} X_{t-2} - \cdots \]
\[ + (\phi_1 X_{t-1} - \phi_2 X_{t-2} + \cdots + (\phi_q X_{t-q} - \phi_{q+1} X_{t-q+1} + \cdots) \]
\[ + (\phi_{q+1} X_{t-(q+1)} - \phi_{q+2} X_{t-(q+2)} + \cdots) \]

where \( \xi_t \) is a white noise process assumed with exponential distribution, Assume the initial value of ARFIMA(p,d,q) process \( \xi_{t-1}, \xi_{t-2}, \ldots, \xi_{t-q} \) and \( X_{t-1}, \ldots, X_{t-p}, X_{t-p-1}, \ldots = 1, \phi_i \) is an autoregressive coefficient, \( i = 1, 2, \ldots, p; |\phi_i| < 1 \) and \( \theta_i \) is moving average coefficient \( i = 1, 2, \ldots, q; |\theta_i| < 1 \).

III. AVERAGE RUNG LENGTH OF ARFIMA MODEL WITH EXPONENTIAL WHITE NOISE ON CUSUM CONTROL CHART

The CUSUM chart was first introduced by Page [1] to detect a small shifting in the mean of a process and widely implemented in statistical process control.

The recursive equation of CUSUM control chart based on ARFIMA(p,d,q) process and defined as:

\[ Y_t = \max(0, Y_{t-1} + X_t - a), \quad t = 1, 2, \ldots, \quad Y_0 = u. \] (4)

where \( Y_t \) is the CUSUM value of a statistic, \( X_t \) is a sequence of ARFIMA(p,d,q) process, \( Y_0 \) is an initial value, \( a \) is a reference value of CUSUM chart.

The corresponding stopping time (\( \tau_\alpha \)) for the CUSUM chart is described by equation (4) is defined as:

\[ \tau_\alpha = \inf \left\{ t > 0 ; \ Y_t > b \right\}, \quad b > u. \] (5)

where \( b \) is a constant parameter known as the upper control limit (UCL).

The ARL of ARFIMA(p,d,q) process on upper-sided CUSUM chart denoted by \( C(u) \) can be written in the form

\[ C(u) = E_u(\tau_\alpha). \] (6)

where \( E_u \) be the expectation corresponding to an initial value \( u \).

IV. NUMERICAL INTEGRAL EQUATION (NIE) METHOD OF ARL ON CUSUM CONTROL CHART

The ARL of ARFIMA(p,d,q) process by using the Fledhoom integral equation of second kind [15].

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)
This section presents the numerical integral equation (NIE) method to compute the solutions equation (6) for ARL of ARFIMA($p,d,q$) process on CUSUM chart. Let $P_c$ and $E_c$ are the probability measurement and the expectation, respectively, which correspond to an initial value $Y_0 = u$.

The solution of integral equations is as follows:

$$C(u) = 1 + E_c[I(0 < Y_t < b|C(Y_t))] + P_e[Y_t = 0|C(0)].$$ (7)

Therefore, the integral equation of ARFIMA($p,d,q$) process can be written in the form

$$C(u) = 1 + \alpha e^{a(u-a_0-X_t)} \int_0^b C(y) e^{-ay} dy + (1-e^{a(b-a_0-X_t)})C(0).$$ (8)

The integral equation (8) can be rewritten as follow

$$C(u) = 1 + C(0) F(a - u - X_t) + \int_0^b C(y) f(y + a - u - X_t) dy$$ (9)

where $F(u) = 1 - e^{-\lambda u}$ and $f(u) = \frac{dF(u)}{du} = \lambda e^{-\lambda u}$

According to the elementary quadrature rule, the integral

$$\int_0^b f(y)dy$$ can be approximated from a sum of areas of rectangles where the integral ($f$) value is chosen by base $b/m$ with heights at the midpoints of intervals of length $b/m$ beginning at zero. Then, with the division points $0 \leq a_1 \leq \ldots \leq a_n \leq b$ and weights $w_j = b/m \geq 0$ on the interval $[0,b]$, we obtain

$$\int_0^b W(y) f(y) dy = \sum_{j=1}^n w_j f(a_j)$$ (10)

with $a_j = \frac{b}{m} \left(j - \frac{1}{2}\right) ; j = 1,2,\ldots,m$,

where $W(y)$ is a weight function, $a_j$ is a set of point and $w_j$ is a weight define different quadrature rules.

Let $\tilde{C}(u)$ denotes the NIE of $C(u)$ then the integral equation in equation (8) is approximated by

$$\tilde{C}(a_j) = 1 + \tilde{C}(a_j) F(a - a_j - X_t) + \sum_{j=1}^m w_j \tilde{C}(a_j) f(a_j + a - a_j - X_t)$$ (11)

The above equation is a system of $m$ linear equations in the $m$ unknowns $\tilde{C}(a_1), \tilde{C}(a_2), \ldots, \tilde{C}(a_m)$, which can be rearranged as

$$\tilde{C}(a_1) = 1 + \tilde{C}(a_1) \left[F(a - a_1 - X_t) + w_i f(a - X_t)\right] + \sum_{j=2}^m w_j \tilde{C}(a_j) f(a_j + a - a_1 - X_t)$$ (12)

It can be written for the matrix form as

$$C_n = I_n + R_n \cdot C_n$$

where

$$C_{n1} = \begin{bmatrix} \tilde{C}(a_1) \\ \tilde{C}(a_2) \\ \vdots \\ \tilde{C}(a_n) \end{bmatrix}$$

The solution of matrix equation in equation (12) is as follow

$$C_{n1} = (I_n - R_n)^{-1} I_n$$ (13)

Solving set of equations for the approximate values of $\tilde{C}(a_1), \tilde{C}(a_2), \ldots, \tilde{C}(a_n)$, the NIE for function $C(u)$ is

$$\tilde{C}(u) = 1 + \tilde{C}(u) F(a - u - X_t) + \sum_{j=1}^n w_j \tilde{C}(a_j) f(a_j + a - u - X_t)$$ (14)

with $w_j = \frac{b}{m}$ and $a_j = \frac{b}{m} \left(j - \frac{1}{2}\right) ; j = 1,2,\ldots,m$.

V. COMPARISON OF ANALYTICAL RESULTS WITH NUMERICAL INTEGRAL EQUATION (NIE) METHOD

This section compares the NIE and explicit formulas for ARL-0 and ARL-1 on CUSUM control chart. The ARL obtained from the NIE and explicit formulas denoted by $\tilde{C}(u)$ and $C(u)$, respectively. $\tilde{C}(u)$ and $C(u)$ are used for ARL from two methods which define the percentage of absolute different as:

$$\text{Diff} \% = \left(\frac{\tilde{C}(u) - C(u)}{C(u)}\right) \times 100\%$$ (15)
The comparison of performance between the NIE and explicit formulas of ARL is shown in Table I. Process is in-control which was a fixed $ARL_0 = 370$. The results show that $ARL_0$ of NIE close to explicit formulas and approach to 370. The percentage of absolute different of NIE and the explicit formula was less than 0.2%.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$\theta$</th>
<th>NIE</th>
<th>Explicit</th>
<th>Diff (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>3</td>
<td>3.29192</td>
<td>346.8406</td>
<td>347.5009</td>
</tr>
<tr>
<td>0.10</td>
<td>0.10</td>
<td>3.070583</td>
<td>307.6282</td>
<td>307.6282</td>
</tr>
<tr>
<td>0.10</td>
<td>0.20</td>
<td>3.207403</td>
<td>207.8471</td>
<td>207.8471</td>
</tr>
<tr>
<td>0.10</td>
<td>3</td>
<td>3.853184</td>
<td>85.4672</td>
<td>85.4672</td>
</tr>
<tr>
<td>0.10</td>
<td>0.10</td>
<td>44.6340</td>
<td>44.6340</td>
<td>44.6340</td>
</tr>
<tr>
<td>0.10</td>
<td>0.20</td>
<td>50.0000</td>
<td>50.0000</td>
<td>50.0000</td>
</tr>
<tr>
<td>3.5</td>
<td>0.50</td>
<td>347.6116</td>
<td>347.6116</td>
<td>347.6116</td>
</tr>
<tr>
<td>3.5</td>
<td>0.10</td>
<td>308.8895</td>
<td>308.8895</td>
<td>308.8895</td>
</tr>
<tr>
<td>3.5</td>
<td>0.30</td>
<td>211.2173</td>
<td>211.2173</td>
<td>211.2173</td>
</tr>
<tr>
<td>3.5</td>
<td>0.50</td>
<td>47.1599</td>
<td>47.1599</td>
<td>47.1599</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$\theta$</th>
<th>NIE</th>
<th>Explicit</th>
<th>Diff (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1.00</td>
<td>347.0446</td>
<td>347.0446</td>
<td>347.0446</td>
</tr>
<tr>
<td>3</td>
<td>0.10</td>
<td>307.5510</td>
<td>307.5510</td>
<td>307.5510</td>
</tr>
<tr>
<td>3</td>
<td>0.20</td>
<td>287.5198</td>
<td>287.5198</td>
<td>287.5198</td>
</tr>
<tr>
<td>3</td>
<td>0.30</td>
<td>208.4944</td>
<td>208.4944</td>
<td>208.4944</td>
</tr>
<tr>
<td>3</td>
<td>0.50</td>
<td>309.1599</td>
<td>309.1599</td>
<td>309.1599</td>
</tr>
<tr>
<td>3.5</td>
<td>0.10</td>
<td>347.6839</td>
<td>347.6839</td>
<td>347.6839</td>
</tr>
<tr>
<td>3.5</td>
<td>0.20</td>
<td>307.0583</td>
<td>307.0583</td>
<td>307.0583</td>
</tr>
<tr>
<td>3.5</td>
<td>0.30</td>
<td>211.5251</td>
<td>211.5251</td>
<td>211.5251</td>
</tr>
<tr>
<td>3.5</td>
<td>0.50</td>
<td>44.6340</td>
<td>44.6340</td>
<td>44.6340</td>
</tr>
</tbody>
</table>

From Table I and II, the comparison of performance between the NIE and explicit formulas of ARL is shown, which found fixed $ARL_0 = 370$. The control chart was given a shift size ($\delta$) where $\delta = 0.01$, 0.03, 0.10, 0.30 and 0.50. For NIE we used $m = 800$ and fixed parameters $a = 3, 3.5$, $\phi_1 = 0.10, -0.10$ and $\theta_1 = 0.10$ and $\theta_2 = 0.20$ for Long memory process with ARFIMA(1, 0.3, 2). The results were in agreement with the numerical approximation with percentage of absolute difference less than 0.2%.

VI. CONCLUSION

In conclusion, from the above results, one can see that the numerical integral equation (NIE) method of ARFIMA(p,d,q) process with exponential white noise on CUSUM control chart can be successfully applied in real applications for different processes of data, for example in economics, agriculture.

REFERENCES


