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. The ARFIMA(*p*,*d*,*q*) 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 Mititelu 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

Manuscript received December 03, 2015; revised December 22, 2015. This work was supported in part by graduate college, King Mongkut's University of Technology North Bangkok, Bangkok, 10800, Thailand.

W. Peerajit is with Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, 10800, Thailand., (e-mail: p.wilasinee@hotmail.com).

Y. Areepong is with Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, 10800, Thailand., (e-mail: yupaporna@kmutnb.ac.th).

S. Sukparungsee is with Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, 10800, Thailand., (e-mail: swns@kmutnb.ac.th).

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 ARL_0 and ARL_1 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]

$$\Phi(B)(1-B)^d X_t = \mu + \Theta(B)\xi_t,\ \ \xi_t \sim Exp(\alpha). \tag{1}$$

where $\Phi(\mathrm{B}) = (1-\phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$ and $\Theta(\mathrm{B}) = (\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)$ are autoregressive and moving-average operators, respectively; B is the backward shift operator i.e., $B^p\left(X_t\right) = X_{t-p}$ and μ is a constant.

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

$$(1-B)^d = \sum_{k=0}^{\infty}(-B)^k = 1 - dB - \frac{1}{2!}d(1-d)B^2 - \ldots \tag{2}$$

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.

$$\begin{aligned} X_t = {} & \mu + \xi_t - \theta_1\xi_{t-1} - \ldots - \theta_q\xi_{t-q} - \left(-dX_{t-1} + \frac{d(d-1)}{2!}X_{t-2} - \ldots\right) \\ & + \left(\phi_1 X_{t-1} - d\phi_1 X_{t-2} + \ldots\right) + \left(\phi_2 X_{t-2} - d\phi_2 X_{t-3} + \ldots\right) \\ & \vdots \\ & + \left(\phi_p X_{t-p} - d\phi_p X_{t-p-1} + \ldots\right) \end{aligned} \tag{3}$$

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

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

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 (τ_b) for the CUSUM chart is described by equation (4) is defined as:

$$\tau_b = \inf\{t > 0;\ Y_t > b\},\ b > u. \tag{5}$$

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

$$C(u) = E_\infty(\tau_b). \tag{6}$$

where E_∞ 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 Fledhom 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:

$$C(u) = 1 + \mathrm{E}_Y[I\{0 < Y_1 < b\}C(Y_1)] + \mathrm{P}_Y\{Y_1 = 0\}C(0). \quad (7)$$

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

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

The integral equation (8) can be rewritten as follow

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

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

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 $0 \le a_1 \le \ldots \le a_m \le b$ and weights $w_j = b/m \ge 0$ on the interval $[0,b]$, we obtain

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

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

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) can be approximated by

$$\tilde{C}(a_i) = 1 + \tilde{C}(a_1)F(a - a_i - X_t) + \sum_{j=1}^m w_j \tilde{C}(a_j) f(a_j + a - a_i - X_t) \quad (11)$$

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

$$\tilde{C}(a_1) = 1 + \tilde{C}(a_1)\left[F(a - a_1 - X_t) + w_1 f(a - X_t)\right] + \sum_{j=2}^m w_j \tilde{C}(a_j) f(a_j + a - a_1 - X_t)$$

$$\tilde{C}(a_2) = 1 + \tilde{C}(a_1)\left[F(a - a_2 - X_t) + w_1 f(a_1 + a - a_2 - X_t)\right] + \sum_{j=2}^m w_j \tilde{C}(a_j) f(a_j + a - a_2 - X_t)$$

$$\vdots$$

$$\tilde{C}(a_m) = 1 + \tilde{C}(a_1)\left[F(a - a_m - X_t) + w_1 f(a_1 + a - a_m - X_t)\right] + \sum_{j=2}^m w_j \tilde{C}(a_j) f(a_j + a - a_m - X_t)$$

It can be written for the matrix form as

$$\mathbf{C}_{m\times1} = \mathbf{1}_{m\times1} + \mathbf{R}_{m\times m}\mathbf{C}_{m\times1} \quad (12)$$

where $\mathbf{C}_{m\times1} = \begin{bmatrix} \tilde{C}(a_1) \\ \tilde{C}(a_2) \\ \vdots \\ \tilde{C}(a_m) \end{bmatrix}$, $\mathbf{1}_{m\times1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

$$\mathbf{R}_{m\times m} = \begin{bmatrix} F(a-a_1-X_t) + w_1 f(a-X_t) & \ldots & w_m f(a_m + a - a_1 - X_t) \\ F(a-a_1-X_t) + w_1 f(a_1+a-a_2-X_t) & \ldots & w_m f(a_m + a - a_2 - X_t) \\ \vdots & & \vdots \\ F(a-a_m-X_t) + w_1 f(a_1+a-a_m-X_t) & \ldots & w_m f(a_m + a - a_m - X_t) \end{bmatrix}$$

and $I_m = diag(1,1,\ldots,1)$ is the unit matrix order m.

Therefore, $\mathbf{C}_{m\times1} = \mathbf{1}_{m\times1} + \mathbf{R}_{m\times m}\mathbf{C}_{m\times1}$, or equivalently $(\mathbf{1}_m - \mathbf{R}_{m\times m})\mathbf{C}_{m\times1} = \mathbf{1}_{m\times1}$. If it exists $(1 - \mathbf{R}_{m\times m})^{-1}$, then the solution of matrix equation in equation (12) is as follow

$$\mathbf{C}_{m\times1} = (\mathbf{1} - \mathbf{R}_{m\times m})^{-1}\mathbf{1}_{m\times1} \quad (13)$$

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

$$\tilde{C}(u) = 1 + \tilde{C}(a_1)F(a - u - X_t) + \sum_{j=1}^m w_j \tilde{C}(a_j) f(a_j + a - u - X_t) \quad (14)$$

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

V. Comparison of Analytical Results with Numerical Integral Equation (NIE) Method

This section compares the NIE and explicit formulas values 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:

$$Diff(\%) = \frac{|\tilde{C}(u) - C(u)|}{\tilde{C}(u)} \times 100\% \quad (15)$$

TABLE I

Comparison of ARL_0 values for ARFIMA(1, 0.3, 2) process using NIE against explicit formulas when given u = 1, θ_1 = 0.10 and θ_2 = 0.20 for ARL_0 = 370.

ϕ_1	*a*	*b*	NIE	Explicit Formulas	Diff (%)
0.10	3	3.29192	369.2882	370.0002	0.1928
	3.5	2.705049	369.3982	370.0004	0.1630
-0.10	3	3.159773	369.3113	370.0003	0.1866
	3.5	2.5868	369.4221	370.0003	0.1565

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 II

Comparison of ARL_1 values for ARFIMA(1, 0.3, 2) process using NIE against explicit formulas when given u = 1, ϕ_1 = 0.10, θ_1 = 0.10, θ_2 = 0.20 for ARL_0 = 370.

a	*b*	δ	NIE	Explicit	*Diff* (%)
3	3.29192	0.01	346.8406	347.5009	0.1904
		0.03	307.0583	307.6282	0.1856
		0.10	207.4936	207.8471	0.1704
		0.30	85.3814	85.4672	0.1005
		050	44.6340	44.6830	0.1098
3.5	2.705049	0.01	347.6116	348.1718	0.1612
		0.03	308.8895	309.3761	0.1575
		0.10	211.2173	211.5251	0.1457
		0.30	89.0993	89.2049	0.1185
		050	47.1599	47.3028	0.3030

TABLE III

Comparison of ARL_1 values for ARFIMA(1, 0.3, 2) process using NIE against explicit formulas when given u = 1, ϕ_1 = - 0.10, θ_1 = 0.10, θ_2 = 0.20 for ARL_0 = 370.

a	*b*	δ	NIE	Explicit	*Diff* (%)
3	3.159773	0.01	347.0446	347.6839	0.1842
		0.03	307.5510	308.1039	0.1798
		0.10	208.4944	208.8391	0.1653
		0.30	86.3609	86.4753	0.1325
		050	45.3127	45.3617	0.1081
3.5	2.5868	0.01	347.7297	348.2677	0.1547
		0.03	309.1599	309.6276	0.1513
		0.10	211.7664	212.0632	0.1402
		0.30	89.6664	89.7690	0.1144
		050	47.6685	47.7139	0.0952

From table I and II, the comparison of performance between the NIE and explicit formulas of ARL are shown, which found fixed ARL_0 = 370. The control chart was given a shift size (δ) where δ = 0.01, 0.03, 0.10, 0.30 and 0.50. For NIE we used m = 800 nods and fixed parameters a = 3, 3.5, ϕ_1 = 0.10, - 0.10 and θ_1 = 0.10 and θ_2 = 0.20 for Long memory process with ARFIMA(1, 0.3, 2). The results were in good 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

[1] E. S. Page, “Continuous inspection schemes,” *Biometrika*, vol. 41, pp. 100-114, 1954.

[2] E. S. Page, “Cumulative Sum Charts,” *Technometrics*, vol. 3, no. 1, pp. 1-9, Feb. 1961.

[3] W. D. Ewan, “When and How to Use CUSUM Charts,” *Technometrics*, vol. 5, no. 1, pp. 1-22, Feb. 1963.

[4] F. F. Gan, “An optimal design of CUSUM quality control charts,” *Journal of Quality Technology*, vol. 23, no.4, pp. 279-286, 1991.

[5] J. M. Lucas, “The design and use of quality control schemes,” *Journal of Quality Technology*, vol. 8, pp. 1-12, 1976.

[6] D. M. Hawkins, “A CUSUM for Scale Parameter,” *Journal of Quality Technology*, vol. 13, no. 4, pp. 228-231, Oct. 1981.

[7] D. M. Hawkins, “Cumulative Sum Control Charting: An Underutilized SPC Tool,” *Quality Engineering*, vol. 5, no. 3, pp. 463-477, 1993.

[8] W. H. Woodall and B. M. Adams, "The Statistical designs of CUSUM Charts", *Quality Engineering*, vol. 5, pp. 559-570, 1993.

[9] S. W. Roberts, “Control Chart Test Based on Geometric Moving Average,” *Technometrics*, vol. 1, pp. 239-250, 1959.

[10] P. B. Robinson and T. Y. Ho, “Average run lengths of Geometric moving average charts by numerical methods,” *Technometrics*, vol. 20, pp. 85-93, 1978.

[11] S. V. Crowder, “A simple method for studying run length distributions of exponentially weighted moving average charts,” *Technometrics*, vol. 29, pp. 401-407, 1987.

[12] J.M. Lucas, and M.S. Saccucci, “Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements,” *Technometrics*, vol. 32, no.1, pp. 1-29, Feb. 1990.

[13] S. Sukparungsee and A. A. Novikov, “On EWMA procedure for detection of a change in observations via martingale approach,” *An International Journal of Science and Applied Science*, vol. 6, pp. 373-380, 2006.

[14] Y. Areepong and A. A. Novikov, “An integral eq uation approach for analysis of control charts,” Ph.D. dissertation, University of Technology, Australia, 2009.

[15] G. Mititelu, Y. Areepong, S. Sukparungsee, and A. Novikov, “Explicit Analytical Solutions for the Average Run Length of CUSUM and EWMA Charts,” *East-West J. Math.*, vol. 1, pp. 253-265, 2010.

[16] C.W. Granger and R. Joyeux, “An introduction to long-memory time series models and fractional differencing,” *Journal of Time Series Analysis*, vol. 1, pp. 15-29, 1980.

[17] J.R.M., Hosking, “Fractional differencing”, *Biometrika*, vol. 68, pp. 165-176, 1981.

[18] J. Beran, *Statistics for long-memory processes.* London: Chapman & Hall, 1994

[19] W. Palma, *Long-memory time series: theory and methods.* New York: Wiley, 2007.

[20] J. Beran, Y. Feng, S. Ghosh and R. Kulik, *Long-memory processes—probabilistic properties and statistical methods*. New York: Springer, 2013

[21] R. Caballero, S. Jewson and A. Brix, “Long memory in surface air temperature: detection, modeling, and application to weather derivative valuation,” *Climate Research*, vol. 21, pp. 127–140, June. 2002.

[22] J. N. Pan and S. T. Chen, “Monitoring long-memory air quality data using ARFIMA model,” *Environmetrics*, vol. 19, pp. 209–219, 2008.

[23] P.A. Jacob, and P.A.W. Lewis, “A Mixed Autoregressive-Moving Average Exponential Sequence and Point Process (EARMA 1, 1),” *Adv.Appl. Probab.*, vol. 9, no. 1, pp. 87-104, Mar. 1977.

[24] I. Mohamed, and F. Hocine, “Bayesian Estimation of an AR(1) Process with Exponential White Noise,” *A Journal of Theoretical and Applied Statistics*, vol. 37, no. 5, pp. 365-372, 2010.

[25] W. Suriyakat, Y. Areepong, S. Sukparungsee, and G. Mititelu, “On EWMA Procedure for an AR(1) Observations with Exponential White Noise.” *International Journal of Pure and Applied Mathematic*, vol. 77, pp. 73-83, 2012.

[26] J. Busaba, S. Sukparungsee, Y. Areepong, and G. Mititelu, “Numerical approximations of average run length for AR(1) on exponential CUSUM,” presented at the International Muti Conference of Engineers and Computer Scientists, Hong Kong, March. 7-10, 2012.

[27] K. Petcharat, Y. Areepong, S. Sukparungsri, and G. Mititelu, “Numerical Approximation to the Performance of CUSUM Charts for EMA (1) Process,” *World Academy of Science, Engineering and Technology*, vol. 6, pp. 12-21, 2012.

[28] S. Phanyaem, Y. Areepong and S. Sukparungsee, “Numerical Integration of Average Run Length of CUSUM Control Chart for ARMA Process,” *International Journal of Applied Physics and Mathematics*, vol. 4, no. 4, pp.232-235, July 2014.