Numerical Approximation to the Performance of CUSUM Charts for EMA (1) Process

K. Petcharat, Y. Areepong, S. Sukparungsri, and G. Mititelu

***Abstract*—**These paper, we approximate the average run length (ARL) for CUSUM chart when observation are an exponential first order moving average sequence (EMA1). We used Gauss-Legendre numerical scheme for integral equations (IE) method for approximate ARL_0 and ARL_1, where ARL in control and out of control, respectively. We compared the results from IE method and exact solution such that the two methods perform good agreement.

***Keywords*—**Cumulative Sum Chart, Moving Average Observation, Average Run Length, Numerical Approximations.

I. Introduction

THE Cumulative Sum (CUSUM) chart is a simple and very effective graphical procedure for monitoring the quality control in manufacturing industry. CUSUM chart was first introduced by Page [1] to detect a change in observed parameters, and widely implemented in statistical process control. Some recent reviews are given in the paper of Mazalov and Zhuravlev [2], who implemented CUSUM chart to identified the changing point in a traffic network. Bakhodir [3] employed CUSUM charts in economics and finance to detected turning point in the stock price indices. CUSUM charts were intensively used by Ben et.al [4] in environmental science to detect mean changes in air pollution, Kennedy [5] in queuing process computed the distribution of the first passage times for a M/M/l queue and stopping times associated with sequential cumulative sum tests. In addition, there are many applications of CUSUM chart in health care and public health see Lim et al [6], Sibanda and Sibanda [7], Noyez,[8].

The common characteristic of any control chart is the Average Run Lengths (ARL), defined as the expectation of an alarm time taken to trigger a signal about a possible change in parameters distribution. Ideally, an acceptable ARL of an in-control process should be large enough to detect a small change in parameters distribution. In this paper we adopt the following notation $ARL_0 = \mathrm{E}_\infty(\tau) = T$ where $\mathrm{E}_\infty(.)$ is the expectation corresponding to the target value and is assumed to be large enough. The ARL when the process is out-of-control is called the Average Delay time denoted by (ARL_1), defined as the expectation of delay for true alarm time. This time should minimize the quantity

$$ARL_1 = \hat{\mathrm{E}}_\nu\left(\tau - \nu + 1 \mid \tau \geq \nu\right)$$

where $\hat{E}_\nu(.)$ is the expectation under the assumption that a change-point occurs at a given time.

In literature several methods for evaluating ARL_0 and ARL_1 for CUSUM and EWMA procedure have been studied. These methods are: the Monte Carlo simulations, the Integral Equations (IE) approach [9]-[11], the Markov Chain Approximation (MCA) [12]-[13]. Recently, Areepong [14] proposed analytical derivation to find explicit formulas for ARL of EWMA chart when observations are exponential distributed. Mititelu et al. [15]-[16], presented analytical expressions to determine the ARL of EWMA and CUSUM chart when observations have hyperexponential distribution via Fredholm integral equations approach. Petcharat. K, et al.[17],[18] derive closed form expressions for the ARL of CUSUM chart when observations are Pareto and Weibull distributed by approximating these distributions with a hyperexponential distribution. Traditionally, CUSUM control charts have been designed when observations are independent and identically distributed (i.i.d). However, in real life problems, correlated observations may be presented in some process [19]-[21], which the correlation may affect the properties of CUSUM chart [22]. Atieza et.al [23], applied CUSUM chart on residuals of a time series model with process observations described by a normal distribution. Jacob and Lewis [24] analyzes autoregressive –moving average process order (1,1) denoted by ARMA(1,1), when observations are exponentially distributed with exponential white noise.

The work of Lawrance and Lewis [25] presented exponential moving average of order 1. Such models are important in queuing and network process. Mohamed and Hocine[26] proposed a Bayesian analysis of the autoregressive model with exponential white noise.

In this paper, we derive integral equations for ARL_0 and ARL_1 and then solve the numerically using the Gauss-Legendre numerical integration equations when observations

K.Petcharat is with Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand.,(E-mail : yha99@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).

G.Mititelu is with Department of Mathematical Sciences, Faculty of Science, University of Technology, Sydney NSW, 2007, Australia.(E-mail: Gabreil.Mititelu@uts.edu.au).

are first order of moving average process, MA(1), with exponential white noise. In section II, we describe characteristics of ARL for CUSUM chart. In section III and section IV, we describe the numerical integral equation approach and exact solution. Section V, we show the numerical results and compare the results obtained from the numerical integration method with the results from [27].

II. The Average Run Length (ARL) for CUSUM Chart of First Order Moving Average, MA (1), Process with Exponential White Noise

The CUSUM chart is often implement in monitoring and detecting small changed in parameters of a given distributions. Let ξ_n be sequence of independent and identically distribution (i.i.d.) nonnegative random variables defined by the recurrence

$$X_t = \max(X_{t-1} + \xi_n - a, 0)\ , \qquad n = 1, 2, ... \qquad (1)$$

where ξ_n are random variables and a is non-zero constant. The corresponding stopping time for the CUSUM scheme described by (1) is defined as

$$\tau_b = \inf\{t > 0; X_t > b\} \qquad (2)$$

where b is a constant parameter known as the control limit.

In this paper ξ_n are continuous distributed i.i.d. random variables, with exponential distribution was described in [15]. The case of a stationary first order autoregressive process with exponential white noise process was analyzed by Busaba et al. [28]. In this paper, we focuses on a stationary first order moving average process, MA(1) with exponential white noise ξ_n define as follow

$$X_t = X_{t-1} + Z_t - a\ ,\ n = 1, 2, ..., X_0 = x$$

where

$$Z_t = \xi_t - \theta\xi_{t-1} \text{ where } -1 < \theta < 1 \quad \text{and} \quad \xi \sim \boldsymbol{exp(\lambda)}$$

III. Numerical Solution for the ARL Integral Equation

The ARL of Gaussian process was approximated by Fledhom integral equation of second kind [16]. In this paper, we apply the approach to the CUSUM chart for MA(1) process. We assume the process is in-control at time t if X_t is in the range $b_L < X_t < b_U$ and out-of-control if $X_t > b_U$ or $X_t < b_L$, where b_L is constant lower bound $(b_L = 0)$ and b_U is constant upper bound $(b_U = b)$. The process is in-control state x that is $X_0 = x$ and $0 \le x \le b$. Now, we define function $j(x)$ as follow $j(x) = \mathrm{E}_x\tau_b < \infty$,

$$j(x) = 1 + \mathbf{E}_X\left[I\{0 < X_1 < b\}\, j(X_1)\right] + \mathbf{P}_X\{X_1 = 0\}\, j(0),\ \ b > x.$$

$$= 1 + \int_0^b j(y) f(a - x + y)\,dy + F(a - x) j(0) \qquad (3)$$

where τ_b is the first exit time defined in(1). Then $j(x)$ is ARL for initial value x.

In a MA(1) process with exponential white noise (3) can be written as:

$$j(x) = 1 + \lambda e^{\lambda(x - a - \theta\xi_0)} \int_0^b j(y) e^{-\lambda y} dy$$

$$+ \left(1 - e^{-\lambda(a - x + \theta\xi_0)}\right) j(0), x \in [0, a)\ . \qquad (4)$$

It can be shown that, ARL of CUSUM chart, $j(x) = \mathrm{E}_x\tau_b$, is a solutions of (4). Rearrange (4) as:

$$j(x) = 1 + j(0) F(a - x + \theta\xi_0) + \int_0^b j(y) f(a - x + \theta\xi_0 + y)\,dy\ , \qquad (5)$$

where $F(x) = 1 - e^{-\lambda x}$ and $f(x) = \dfrac{dF(x)}{dx} = \lambda e^{-\lambda x}$.

Now, via Gauss-Legendre rule, we can approximate the integral $j(x)$ as:

$$j(a_i) \approx 1 + j(a_1) F(a - a_i + \theta\xi_0)$$

$$+ \sum_{k=1}^m w_k j(a_k) f(a_k + a - a_i + \theta\xi_0)\ , \qquad (6)$$

with the weights $w_k = \dfrac{b}{m} \ge 0$ and $a_k = \dfrac{b}{m}\left(k - \dfrac{1}{2}\right)$, $; k = 1, 2, ..., m$

In a MA(1) process with exponential white noise, the numerical solution for ARL integral equation can be written as follow

$$j(a_i) = 1 + j(0) F(a - x + \theta z_0)$$

$$+ \sum_{k=1}^m w_k j(a_k) f(a_k + a - a_i + \theta Z_0)\,dy \qquad (7)$$

We approximate the integral by a sum of areas of rectangles with bases $\frac{b}{m}$ with heights chosen as the value of $f(a_k)$ at the midpoints of intervals of length $\frac{b}{m}$ beginning at zero. Then, on the interval $[0, b]$ with the division points $0 \le a_1 \le a_2 \le ... \le a_m < b$ and weights $w_k = \frac{b}{m} \ge 0$ we can writing as

$$\int_0^b j(y)dy \approx \sum_{k=1}^m w_k f(a_k)$$

where

$$a_k = \frac{b}{m}\left(k - \frac{1}{2}\right) \quad ; k = 1, 2, ..., m, \tag{8}$$

The integral in (8) becomes a system of m linear equations in the m unknowns $j(a_1), j(a_2), ..., j(a_m)$ written as

$$\begin{cases} j(a_1) = 1 + j\left(a_1\right) F(a - a_1 + \theta\xi_0) + w_1 f(a + \theta\xi_0) + \sum\limits_{k=2}^{m} w_k j(a_k) f\left(a_k + a - a_i + \theta\xi_0\right) \\ j(a_2) = 1 + j\left(a_1\right) F(a - a_2 + \theta\xi_0) + w_1 f(a_1 + a - a_2 + \theta\xi_0) + \sum\limits_{k=2}^{m} w_k j(a_k) f\left(a_k + a - a_i + \theta\xi_0\right) \\ \vdots \\ j(a_m) = 1 + j\left(a_1\right) F(a - a_m + \theta\xi_0) + w_1 f(a_1 + a - a_m + \theta\xi_0) + \sum\limits_{k=2}^{m} w_k j(a_k) f\left(a_k + a - a_i + \theta\xi_0\right) \end{cases} \tag{9}$$

For numerical implementation is preferable to writing the linear system in (9) is matrix form as follow

$$\mathrm{J}_{m\times 1} = 1_{m\times 1} + \mathrm{R}_{m\times m}\mathrm{J}_{m\times 1}$$

or

$$(\mathrm{I}_m \text{-} \mathrm{R}_{m\times m})\mathrm{J}_{m\times 1} = 1_{m\times 1} \tag{10}$$

where

$$\mathrm{J}_{m\times 1} = \begin{pmatrix} j(a_1) \\ j(a_2) \\ . \\ . \\ . \\ j(a_m) \end{pmatrix}, \qquad 1_{m\times 1} = \begin{pmatrix} 1 \\ 1 \\ . \\ . \\ . \\ 1 \end{pmatrix},$$

$$\mathrm{R}_{m\times m} = \begin{pmatrix} F(a - a_1 + \theta\xi_0) + w_1 f(a) & w_2 f(a_2 + a - a_1 + \theta\xi_0) & \ldots & w_m f(a_2 + a - a_1 + \theta\xi_0) \\ F(a - a_1 + \theta\xi_0) + w_1 f(a_1 + a - a_2 + \theta\xi_0) & w_2 f(a + \theta\xi_0) & \ldots & w_m f(a_m + a - a_2 + \theta\xi_0) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ F(a - a_m + \theta\xi_0) + w_1 f(a_1 + a - a_m + \theta\xi_0) & w_2 f(a_2 + a - a_m + \theta\xi_0) & \ldots & w_m f(a + \theta\xi_0) \end{pmatrix}, \tag{11}$$

and $\mathrm{I}_m = diag(1,1,...1)$ is the unit matrix order m. If it exists $(\mathrm{I}_m - \mathrm{R}_{m\times m})^{-1}$, then the solution of $\mathrm{R}_{m\times m}$ is

$$\mathrm{J}_{m\times 1} = (\mathrm{I}_m \text{-} \mathrm{R}_{m\times m})^{-1} 1_{m\times 1}.$$

Solving the set of (11) for approximate values of $j(a_1), j(a_2), ..., j(a_m)$ we may approximate the function $j(x)$ as

$$j(x) \approx 1 + j(a_1) F(a - x + \theta\xi_0) + \sum_{k=1}^{m} w_k j(a_k) f(a_k - a - a_i + \theta\xi_0), \tag{12}$$

with $w = \frac{b}{m}$ and $a_k = \frac{b}{m}\left(k - \frac{1}{2}\right)$.

IV. The Exact Solution for ARL

Petcharat et al [27] derived exact solution for ARL of CUSUM Chart for first order moving average process with exponential white noise. We used integral equation method

and derived the exact solution via Fredholm integral equation of the second type for ARL_0 and ARL_1 as follow:

$$ARL_0 = j_0(x) = e^b\left(1 + e^{(a+\theta z_0)} - \lambda b\right) - e^x, \quad x \geq 0 \tag{13}$$

and

$$ARL_1 = j_1(x) = e^{\lambda b}\left(1 + e^{\lambda(a+\theta z_0)} - \lambda b\right) - e^{\lambda x}, \quad x \geq 0 \tag{14}$$

where λ is parameter of exponential distribution, θ is smoothing parameter, Z_0 is initial value of MA(1), b is boundary value and a id reference value.

V. Numerical Results

In this section, we will compare the ARL from two solutions as approximated solution $j(x)$ and explicit solution. We use "IE" and "Explicit" for ARL from two methods and define the absolute percentage difference:

$$Diff(\%) = \frac{|IE - Explicit|}{IE} \times 100 \tag{15}$$

TABLE I
COMPARISON OF ARL VALUES COMPUTED USING NUMERICAL APPROXIMATION (IE) FOR λ=1 AND *M=500* AGAINST EXPLICIT FORMULA (EXPLICIT)

θ	b	ARL	$a=3.5$		$a=4$	
			$x=0$	$x=2$	$x=0$	$x=2$
0.23	0.38	IE	60.831	54.444	100.353	93.967
		Explicit	60.853	54.464	100.391	94.002
		Diff (%)	0.036	0.037	0.038	0.037
	1.7	IE	222.947	216.569	370.701	364.322
		Explicit	223.317	216.928	371.323	364.943
		Diff (%)	0.166	0.165	0.168	0.170
	2.0	IE	298.995	292.619	498.381	492.005
		Explicit	299.580	293.191	499.366	492.977
		Diff (%)	0.196	0.195	0.198	0.198
0.53	0.38	IE	82.145	75.759	135.495	129.108
		Explicit	82.176	75.787	135.546	129.157
		Diff (%)	0.037	0.037	0.038	0.038
	1.7	IE	302.631	296.253	502.078	495.699
		Explicit	303.138	296.745	502.924	496.535
		Diff (%)	0.168	0.166	0.168	0.169
	2.0	IE	406.525	400.149	675.668	669.292
		Explicit	407.326	400.937	677.009	670.620
		Diff (%)	0.197	0.197	0.198	0.198
0.83	0.38	IE	110.917	104.531	182.932	176.545
		Explicit	110.959	104.570	183.001	176.612
		Diff (%)	0.038	0.037	0.038	0.038
	1.7	IE	410.194	403.816	679.418	673.040
		Explicit	410.883	404.494	680.566	674.177
		Diff (%)	0.168	0.168	0.169	0.169
	2.0	IE	551.676	545.299	914.981	908.605
		Explicit	552.768	546.278	916.802	910.413
		Diff (%)	0.198	0.180	0.199	0.199

Table I shows absolute percentage difference less than 0.2% between the analytical expression the Gauss-Legendre numerical scheme for integral equation with $m=500$ nodes and the explicit formula. The two methods are good agreement with the results of ARL.

TABLE II
COMPARISON OF ARL VALUES COMPUTED USING NUMERICAL APPROXIMATION (IE) FOR $a=4$, $b=1.7$ AND $m=500$ AGAINST EXPLICIT FORMULA (EXPLICIT)

λ	$\theta=0.23$		ε_r
	IE	Explicit	
1.0	370.701	371.323	0.168
1.1	215.518	215.845	0.152
1.2	137.097	137.285	0.137
1.3	93.4754	93.5929	0.126
1.4	67.3116	67.3893	0.115
1.5	50.6407	50.6946	0.106

TABLE III
COMPARISON OF ARL VALUES COMPUTED USING NUMERICAL APPROXIMATION (IE) FOR $a=4$, $b=2$ AND *M=500* AGAINST EXPLICIT FORMULA (EXPLICIT)

λ	$\theta=0.23$		ε_r
	IE	Explicit	
1.0	498.381	499.366	0.198
1.1	282.154	281.652	0.178
1.2	175.238	174.955	0.161
1.3	117.071	116.898	0.148
1.4	82.8386	82.7262	0.136
1.5	61.3812	61.3045	0.125

In Tables III and III, the columns IE and Explicit shows comparisons between the numerical and explicit values of the ARL. For a fixed ARL=370 and 500, $a=4$, $b=1.7, 2.0$, and fixed parameter $\theta=0.23$ for the number of division points $m=500$. Notice that $\lambda=1$ is the value assumed for the in-control parameter, so the first row gives the values of the ARL_0. Rows for $\lambda>1$ corresponds to values of out-of-control parameters, therefore these rows give the values for ARL_1. The results are good agreement with the numerical approximation with absolute percentage difference less than 0.2%.

VI. CONCLUSION

We have presented numerical methods for evaluate ARL_0 and ARL_1 of CUSUM chart, when observation are MA(1) process with exponential white noise distribution. The accuracy for numerical integration approach was compare with explicit formula. We have shown that the results of two methods are good agreement.

REFERENCES

[1] E.S. Page, "Continuous Inspection Schemes," *Biometrika*, Vol. 41, No. 1, 1954, pp. 100-115.
[2] V.V. Mazalov, and D.N. Zhuravlev, "A Method of Cumulative Sums in the Problem of Detection of Traffic in Computer Networks," *Program. Comput. Soft.*, Vol. 28 No.6, 2002, pp. 342-348.
[3] E.A. Bakhodir, "Sequential Detection of US Business Cycle Turning Points: Performances of Shiryayev-Roberts, CUSUM and EWMA Procedures," *Econ. Papers*, Submitted for publication.
[4] M. Ben, and J. Antony "Statistical Process Control an Essential Ingredient for Improving Service and Manufacturing Quality" *Managing Service Quality,* Vol.10, No.4, 2000, pp. 233-238.
[5] P.D. Kennedy, "Some Martingales Related to Cumulative Sum Test and Single-Server Queues *Stoch. Proc. Appl.,* Vol. 4, 1975, pp. 261-269.
[6] T.O. Lim, A. Soraya, L.M. Ding, and Z. Morad, "Assessing Doctors' Competence: Application of CUSUM Technique in Monitoring Doctors' Performance," *Int. J. Qual. Health C.,* Vol. 14, No.3, 2002, pp. 251-258.
[7] T. Sibanda, and N. Sibanda, "The CUSUM Chart method as a Tool for Monitoring of Clinical Outcomes Using Routinely Collected Data," *BMC Medical Research Methodology*, Vo.7, No. 46, 2007.

[8] L. Noyez, "Control Charts, CUSUM Techniques and Funnel Plots. A Review of Methods for Monitoring Performance in Healthcare," *Interactive Cardio Vascular and Thoracic Surgery,* Vol. 9, 2009, pp. 494–499.

[9] S. Vardeman, and D. Ray, "Average Run Lengths for CUSUM Schemes When Observations are Exponentially Distributed," *Technometrics*, Vol. 27, No. 2 ,May. 1985, pp. 145-150.

[10] S.V. Crowder, "A Simple Method for Studying Run Length Distributions of Exponentially Weighted Moving Average Charts," *Technometrics,* Vol.29. No. 4, Nov. 1987, pp.401-407

[11] M.S. Srivastava, and Y. Wu, "Evaluation of Optimum Weights and Average Run Lengths in EWMA Control Schemes," *Commun. Stat. Theory.,* Vol. 26, 1997, pp. 1253-1267.

[12] D. Brook, and D.A. Evans, "An Approach to the Probability Distribution of CUSUM Run Length," *Biometrika*, Vol. 59, 1972, pp. 539-548.

[13] J.M. Lucas, and M.S. Saccucci, "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," *Technometrics*, Vol. 32, No.1, Feb. 1990, pp. 1-29.

[14] Y. Areepong, "An Integral Equation Approach for Analysis of Control Charts," *Ph.D.Thesis in Mathematical Sciences*, Faculty of Science, University of Technology, Sydney, 2009.

[15] G. Mititelu, V.V. Mazalov, and A. Novikov, "On CUSUM Procedure for Hyperexponential Distribution," *Stat. Probabil. Lett.*, Submitted for publication.

[16] 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, 2010, pp. 253-265.

[17] K. Petcharat, Y. Areepong, S. Sukparungsee, and G. Mititelu, "Fitting Weibull Distributions with Hyperexponential to Evaluate the Average Run Length for Cumulative Sum Chart," In *Proc. 14th Conf. the ASMDA Int. Soc. Appl. Stoch. Models and Data Anal.*.

[18] K. Petcharat, Y. Areepong, S. Sukparungsee, and G. Mititelu, "Fitting Pareto Distributions with Hyperexponential to Evaluate the Average Run Length for Cumulative Sum Chart," *Int. J. Pure Appl. Math*., Vol.77 No.1, 2012, pp. 233-244.

[19] D.G. Wardell, H. Moskovitz, and R.D. Plante, "Control Charts in the Presence of Data Correlation," *Management Science*, Vol. 38, No. 8, Aug.1992, pp. 1084-1105.

[20] M. Yashchin, "Performance of CUSUM Control Schemes for Serially Correlated Observations," *Technometrics*, Vol. 35 No. 1, Feb. 1993, pp. 37-52.

[21] N.F. Zhang, "A Statistical Control Chart for Stationary Process Data. *Technometrics*, Vol. 40. No. 1, Feb. 1998, pp. 24-38.

[22] C.W. Lu, and M.R. Reynolds, "CUSUM Charts for Monitoring an Autocorrelated Process," *J. Qual. Technol.*, Vol. 33, Jul. 2001, pp. 316–334.

[23] O.O. Atienza, L.C. Tang, and B.W. Ang, "A CUSUM Scheme for SPC Procedures for Monitoring Autocorrelated Processes Autocorrelated Observations," *J. Qual. Technol.*, Vol. 34 No. 2, Apr. 2002, pp. 187-199.

[24] P.A. Jacob, and P.A.W. Lewis, P, "A Mixed Autoregressive-Moving Average Exponential Sequence and Point Process (EARMA 1, 1)," *Adv. Appl. Probab.*, Vol. 9, No. 1, Mar. 1977; pp. 87-104.

[25] J.A. Lawrance, and P.A.W. Lewis, "An Exponential Moving-Average Sequence and Point Process (EMA1)," *J. Appl. Probab*., Vol. 14 No. 1, Mar. 1977;pp. 98-113.

[26] 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, 2010, pp. 365-372.

[27] K. Petcharat, Y. Areepong, S. Sukparungsee, and G. Mititelu, "Exact Solution for Average Run Length of CUSUM Charts for MA(1) Process" Chiang Mai J. Sci. submitted for publication.

[28] J. Busaba, S. Sukparungsee, Y. Areepong, and G. Mititelu, "Numerical Approximations of Average Run Length for AR(1) on Exponential CUSUM," *In proc. Int. multiconference of engineers and computer scientists*, March 7-10, Hong Kong(IMECS 2012).