Explicit Expressions of Average Run Length of Moving Average Control Chart for Poisson Integer Valued Autoregressive Model

S. Phantu, S. Sukparungsee and Y. Areepong

Abstract— An aim of this paper is to propose the explicit formulas for evaluating performance characteristics of Moving Average control chart (MA) for the first order of autoregressive of serial dependence Poisson process. The characteristics of the control chart is frequently measured as Average Run Length (ARL) which means that the average of observations are taken before a system is signaled to be out-of-control. These proposed explicit formulas of ARL are simple and easy to implement for practitioner. The numerical results show that MA chart performs better than others when the magnitudes of shift are moderate and large.

***Keywords*— Average Run Length, Moving Average control chart, Integer-valued Autoregressive.**

I. Introduction

STATISTICAL Quality Control (SQC) is widely used in industries to monitor the quality of Poisson counting process. In manufacturing industry, for instance, the number of nonconformities of a unit of a product process is of interest, while in service industry the number of complaints of customers within certain period of time is an important quality characteristic. The marginal distribution of count processes can often be modelled by Poisson distribution, where μ denotes the average of the Poisson distribution. Let N_t be a process of count data, which is assumed to be stationary with Poisson distributed marginals in the state of statistical control. Most prominent are two charts of Shewhart type, namely c chart and u chart, which both monitor the marginal distribution of the process (N_t). For a detailed description, consult Montgomery (2009) [1].

Count data are often occurred in manufacturing industries because of ease of data collection [2-3]. One particular area where these counts can be useful is in process monitoring to detect shift of a process from an in-control state to various out-of-control states.

Manuscript received December 17, 2015; revised Janurary 28, 2016. This work was supported grant in financial by the Graduate School of King Mongkut's University of Technology North Bangkok.

S. Phantu is with the King Mongkut's University of Technology North Bangkok, Faculty of Applied Science, Department of Applied Statistics, Bangkok, Thailand, e-mail: phantu.s@hotmail.com.

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

Y. Areepong is with the King Mongkut's University of Technology North Bangkok, Faculty of Applied Science, Department of Applied Statistics, Bangkok, Thailand, e-mail: yupaporn.a@sci.kmutnb.ac.th.

Hence, quality losses can be reduced and prevented through corrective actions to put the process back in a normal state. Often, the c chart has been for monitoring count data.

Although originally developed for independent count data have also been discussed in literatures. Shewhart control chart use only the information in the last sample and ignore information given by the entire data sequence. Thus, the c chart are known to have poor performance in detecting small shift in process mean. In the past few decades, Exponentially Weighted Moving Average control chart (EWMA) was introduced by Robert [4]. It is effective chart for detecting small and moderate shifts. Cumulative Sum control chart (CUSUM) was introduced by Page [5]. It is sensitive to small shift for process mean. Recently, Khoo [6] studied MA chart for monitoring the fraction of non-conforming observations and showed that MA chart is more efficient than p chart. Mostly, the count process assumed that the count data are independent and identically distributed (i.i.d.). However, observations could be serially autocorrelated and it may adversely affect the performance of the control charts under this assumption of independence. Generally, the Integer-valued time series occur in many situations, for example counts of events in consecutive points of time, the number of births at a hospital in successive months, the number of road accidents in a city in successive months and big numbers even for frequently traded stocks. Because of the broad field of potential applications, a number of time series models for counts have been proposed in many literatures. McKenzie [7] introduced the first order integer-valued autoregressive (INAR(1)) model. The statistical properties of INAR(1) are discussed in McKenzie [8], Al-Osh and Alzaid [9].

Mostly, the performance of control chart when the process is in-control is usually characterized by in-control Average Run Length (ARL_0) - an average of times before the control chart gives a false alarm as the in-control process has gone to out-of-control process. Conversely, the performance under out-of-control situation is Average of Delay Times (ADT) –average times between process goes out-of-control and control chart giving an alarm that the process has gone out-of-control. Ideally, the value of ARL_0 of an acceptable chart should be sufficient large and the value of ADT should be minimum. Most researches for evaluating the ARL_0 and ADT for control charts have been studied in the literature. A basic approach that is often used to test other methods is Monte Carlo (MC) simulation. Roberts [5] studied the ARL for EWMA charts by using simulations for processes following a normal distribution that can be used to find the ARL for a variety of parameter

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

values. Crowder [10] studied numerical quadrature methods to solve the exact Integral Equations (IE) for the ARL for the normal distribution. Brook and Evans used an approximate formula for the ARL of EWMA chart by using a finite-state Markov Chain Approach (MCA). Areepong and Novikov [11] derived explicit formulas for ARL of Exponentially Weighted Moving Average control charts. Recently, Areepong [12] studied explicit formulas for ARL of MA chart for monitoring the number of defective products. In the literature one can find at least four numerical procedure for evaluating average run length. Monte Carlo (MC) is simple to program and based on a large number of sample trajectories so it is very time consuming to run. Moreover, it is difficult to use for optimization though it is convenient to control accuracy of analytical approximations. Integral Equation (IE) is the most advanced method currently available but it requires intensive programming to implement even for the case of Gaussian distribution and also for the continuous observations. Markov Chain Approach (MCA) is considered a popular technique. It is based on approximation of matrix inversions. In addition there no theoretical results on accuracy of this procedure in terms of rate of convergence. Martingale approach is simple and convenient used to approximate but it could be implemented for the case of light-tailed distributions or the moment generating function exits. In this paper we suggest to explicit formula for evaluate ARL of moving average control chart when observation are Poisson count process. The results show that the performance of MA chart is good when the magnitudes of shift are moderate and large.

II. Control Charts

1. Integer Valued Autoregressive for Poisson Count Data

In case of measurement data, it is common to directly model the autocorrelation structure of the process, especially with models of the ARMA family. There are several different approach to model count processes, which have an integer structure, in the literature. Because the conventional autoregressive moving average (ARMA) recursion includes a scalar multiplication, it cannot applied to the count data. To adapt the ARMA model to the integer-valued case. Therefore, Steutel and Van Harn [13] proposed a probabilistic operation, so called binomial thinning, as an adequate alternative to scalar multiplication. Weiβ [14-15] reviews different thinning operations and summarizes essential properties of the specific thinning operations. It was successfully applied to define the integer-valued ARMA models. Let N be a discrete random variable with range $\{0,...,n\}$ and $\alpha \in [0,1]$ be a parameter. Then, the binomial thinning operation is

$$\alpha \circ N = \sum_{i=1}^{N} X_i \tag{1}$$

where X_i are i.i.d. Bernoulli random variables, which are independent of N and $P(X_i = 1) = \alpha$. We say $\alpha \circ N$ arise from N by binomial thinning, and $\circ$ is the binomial operator.

The first integer-valued ARMA model, the INAR(1) model, was introduced by McKenzie [8]. Alzaid and Al-Osh [9] derived a number of important statistical properties of these models, which are the discrete analogue of the usual AR(1) model. The INAR(1) process is defined by the recursion.

$$N_t = \alpha \circ N_{t-1} + \varepsilon_t. \tag{2}$$

where N_t is the observable count at time t and the innovations ε_t are i.i.d. count data. The INAR(1) model is the best fitting model for Poisson marginal. If ε_t follows the Poisson distribution with mean $\mu(1-\alpha)$ that is $Poi(\mu(1-\alpha))$ and if the initial count N_0 is distributed as $Poi(\mu)$, then N_t is stationary and distributed as $Poi(\mu)$. According to the above situation, it can be modelled as Poisson INAR(1) model. The expectation and variance of INAR(1) model are as follows:

$$E[N_t] = V[N_t] = \frac{\mu}{1-\alpha}. \tag{3}$$

2. The Moving Average Control Chart

A moving average control chart is a type of memory control chart based on unweighted moving average. Suppose individual observations, $N_1, N_2, ...,$ are collected moving average of width w at time i is defined as (Montgomery, 2009, [1])

$$MA_i = \frac{N_i + N_{i-1} + ... + N_{i-w+1}}{w}. \tag{4}$$

For period $i \ge w,$ for period $i < w,$ we do not have w observations to calculate a moving average of width $w.$ For these periods the average of all observations up to period i defines the moving average. When the process is in-control, the mean and variance (for $i \ge w$) of MA_i are

$$E[MA_i] = \frac{\mu}{1-\alpha} \tag{5}$$

and

$$V[MA_i] = \begin{cases} \dfrac{\mu}{i^2(1-\alpha)} & ;\ i < w \\ \dfrac{\mu}{w^2(1-\alpha)} & ;\ i \ge w. \end{cases} \tag{6}$$

If a target value for the process mean is $\frac{\mu_0}{1-\alpha_0},$ then, for periods $i \ge w,$ centerline and 3σ control limit are given by

$$UCL / LCL = \frac{\mu_0}{1-\alpha_0} \pm H\sqrt{\frac{\mu_0}{w(1-\alpha_0)}} \tag{7}$$

For periods $i < w,$ the limits of the MA chart are

$$UCL / LCL = \frac{\mu_0}{1-\alpha_0} \pm H\sqrt{\frac{\mu_0}{i(1-\alpha_0)}}. \tag{8}$$

The characteristics for statistical process control charts is Average Run Length (ARL). It is the average of observation s are taken before the first point signal out of control and the expectation of an alarm time about a possible change. Ideally, an acceptable ARL_0 of in-control process should be enough large and a small ADT when the process is out-of-control, so-called Average of Delay Time (ADT).

3. The Explicit Formula for Evaluate Average Run Length for Moving Average Control Chart

The ARL values of Moving Average control chart can be derived. Let $ARL = n,$ then

$$\frac{1}{ARL} = \frac{1}{n} P(o.o.c\ \ signal\ \ at\, time\, i < w)$$

$$+[\frac{n-(2w-1)}{n}]P(o.o.c\ \ signal\ \ at\, time\, i \geq w)$$

$$= \frac{1}{n}\left\{\sum_{i=1}^{w}[P(\frac{\sum_{j=1}^{i} N_j}{i} > UCL_i) + P(\frac{\sum_{j=1}^{i} N_j}{i} < LCL_i)]\right\}$$

$$+[\frac{n-(2w-1)}{n}]\{P(\frac{\sum_{j=i-w+1}^{i} N_j}{w} > UCL_w)$$

$$+P(\frac{\sum_{j=i-w+1}^{i} N_j}{w} < LCL_w)\}.$$

Let

$$Z_1 = \frac{\sum_{j=1}^{i} N_j - \frac{\mu}{1-\alpha}}{\sqrt{\frac{\mu}{i(1-\alpha)}}} \quad \text{and} \quad Z_2 = \frac{\sum_{j=i-w+1}^{i} N_j - \frac{\mu}{1-\alpha}}{\sqrt{\frac{\mu}{w(1-\alpha)}}}.$$

The solution can be obtained by central limit theorem, then the explicit formula of ARL_0 for MA chart

$$ARL_0 = \left[1 - \sum_{i=1}^{w-1}\left[P\left(Z_1 > \frac{\left(\frac{\mu_0}{1-\alpha_0} + H\sqrt{\frac{\mu_0}{i(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_0}{1-\alpha_0}}{\sqrt{\frac{\mu_0}{i(1-\alpha_0)}}}\right) + P\left(Z_1 > \frac{\left(\frac{\mu_0}{1-\alpha_0} - H\sqrt{\frac{\mu_0}{i(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_0}{1-\alpha_0}}{\sqrt{\frac{\mu_0}{i(1-\alpha_0)}}}\right)\right]\right]$$
$$\times\left[P\left(Z_2 > \frac{\left(\frac{\mu_0}{1-\alpha_0} + H\sqrt{\frac{\mu_0}{w(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_0}{1-\alpha_0}}{\sqrt{\frac{\mu_0}{w(1-\alpha_0)}}}\right) + P\left(Z_2 > \frac{\left(\frac{\mu_0}{1-\alpha_0} - H\sqrt{\frac{\mu_0}{w(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_0}{1-\alpha_0}}{\sqrt{\frac{\mu_0}{w(1-\alpha_0)}}}\right)\right]^{-1} + (w-1). \tag{9}$$

The explicit formula of ADT can be written as follows:

$$ADT = \left[1 - \sum_{i=1}^{w-1}\left[P\left(Z_1 > \frac{\left(\frac{\mu_0}{1-\alpha_0} + H\sqrt{\frac{\mu_0}{i(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_1}{1-\alpha_0}}{\sqrt{\frac{\mu_1}{i(1-\alpha_0)}}}\right) + P\left(Z_1 > \frac{\left(\frac{\mu_0}{1-\alpha_0} - H\sqrt{\frac{\mu_0}{i(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_1}{1-\alpha_0}}{\sqrt{\frac{\mu_1}{i(1-\alpha_0)}}}\right)\right]\right]$$
$$\times\left[P\left(Z_2 > \frac{\left(\frac{\mu_0}{1-\alpha_0} + H\sqrt{\frac{\mu_0}{w(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_1}{1-\alpha_0}}{\sqrt{\frac{\mu_1}{w(1-\alpha_0)}}}\right) + P\left(Z_2 > \frac{\left(\frac{\mu_0}{1-\alpha_0} - H\sqrt{\frac{\mu_0}{w(1-\alpha_0)}}\right) - 0.5 - \frac{\mu_1}{1-\alpha_0}}{\sqrt{\frac{\mu_1}{w(1-\alpha_0)}}}\right)\right]^{-1} + (w-1). \tag{10}$$

III. Numerical Results

The numerical results for ARL_0 and ADT for MA chart was calculated from Equation (9) and Equation (10). The parameter values for MA chart following moving average control charts consider w = 2, 3, 4, 5, 10 and 15. The in-control parameter are $\mu_0 = 1,\ 3$ and $\alpha_0 = 0.2$ and out-of-control parameter values are μ_1 and shift parameters (δ) = 0.1, 0.2, 0.3, …,1.5). The results shows that the MA chart proposed are sensitive to only some of the out-of-control situations considered. Table I shown average run length for $\mu_0 = 1$ and ARL_0 = 370 of MA chart, the results shows that when small shifts are δ ≤ 0.5 the MA chart have the performance when w = 15. For moderate shifts (0.6 ≤ δ ≤ 0.8), the performance of MA chart with w=10 is superior to others. For the shift sizes (0.9 ≤ δ ≤ 1.0), the performance of MA chart with w = 5 is the best control chart. For large shifts (δ > 1.5) the MA chart have the performance when w = 4. Table II show the average run length for $\mu_0 = 3$ and ARL_0 = 370 of MA chart, the results show that when small shifts, δ ≤ 0.3 the MA chart have the better performance when w = 15. For moderate shifts (0.4 ≤ δ ≤ 0.5) the MA chart have good performance when w = 10. For magnitude of shifts (0.6 ≤ δ ≤ 0.7), the MA chart have good performance when w = 5. For parameter shift (0.8 ≤ δ ≤ 0.9) the MA chart have the performance when w = 4. For parameter shifts (δ ≤ 1.0) the MA chart have good performance when w = 3. For large shifts (δ > 1.5) the MA chart have good performance when w = 2. The average run length for MA chart was shown that for the shift increasing MA performs better as the value of (w) decreases. So the average run length for MA chart show when shifts increase MA performs better as the values of w decrease. Table III and IV show that the numerical results for ARL_0 and ADT for MA chart present in a similar manner. For desired ARL_0=500, the parameter of H = 3 and 3.0905 of MA chart are used μ_0 = 1 and 3, respectively. The results when ARL_0 = 370 are in good agreement with the results when ARL_0 = 500. Therefore, use of the proposed formulas for ARL_0 and ADT for MA chart are simply to calculate, time saving and easy to implement which greatly reduce computation times, and useful to practitioners.

IV. Conclusions

We derived the explicit formulas for average run length of moving average control chart (MA) for Poisson counting process. The INAR(1) model is a simple but well

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I

ARL OF MA CHART FOR GIVEN $\mu_0 = 1$, $\alpha_0 = 0.2$ AND $ARL_0 = 370$ CONSIDERING A CHANGE IN μ.

δ	w=2	w=3	w=4	w=5	w=10	w=15
0.0	370.4	370.4	370.4	370.4	370.4	370.4
0.1	217.63	211.06	204.84	198.94	173.84	**154.56**
0.2	125.08	114.63	105.69	97.98	71.89	**57.64**
0.3	74.32	64.70	57.21	51.27	34.49	**27.74**
0.4	46.59	39.05	33.62	29.58	19.87	**17.28**
0.5	30.89	25.25	21.46	18.80	13.41	**12.94**
0.6	21.57	17.39	14.75	13.01	**10.19**	10.81
0.7	15.77	12.65	10.78	9.66	**8.41**	9.60
0.8	11.99	9.65	8.33	7.60	**7.32**	8.78
0.9	9.43	7.65	6.72	**6.26**	6.59	8.14
1.0	7.64	6.28	5.62	**5.35**	6.06	7.60
1.5	3.66	3.27	**3.23**	3.33	4.44	5.34

TABLE II

ARL OF MA CHART FOR GIVEN $\mu_0 = 3$, $\alpha_0 = 0.2$ AND $ARL_0 = 370$ CONSIDERING A CHANGE IN μ.

δ	w=2	w=3	w=4	w=5	w=10	w=15
0.0	370.40	370.40	370.40	370.40	370.40	370.40
0.1	187.27	172.89	160.38	149.42	110.86	**88.38**
0.2	83.96	69.85	59.56	51.81	31.99	**25.11**
0.3	40.72	31.82	26.08	22.19	14.51	**13.58**
0.4	22.14	16.85	13.77	11.88	**9.35**	10.33
0.5	13.38	10.20	8.53	7.63	**7.32**	8.85
0.6	8.85	6.89	5.98	**5.60**	6.25	7.79
0.7	6.30	5.07	4.60	**4.49**	5.53	6.84
0.8	4.77	3.99	**3.78**	3.82	4.94	5.94
0.9	3.79	3.31	**3.25**	3.36	4.41	5.10
1.0	3.14	**2.85**	2.88	3.03	3.93	4.36
1.5	**1.77**	1.81	1.92	2.02	2.22	2.24

TABLE III

ARL OF MA CHART FOR GIVEN $\mu_0 = 1$, $\alpha_0 = 0.2$ AND $ARL_0 = 500$ CONSIDERING A CHANGE IN μ.

δ	w=2	w=3	w=4	w=5	w=10	w=15
0.0	500.11	500.11	500.11	500.11	500.11	500.11
0.1	285.21	276.13	267.55	259.43	224.94	**198.52**
0.2	159.07	145.14	133.29	123.09	88.71	**69.89**
0.3	92.00	79.58	69.95	62.32	40.78	**31.90**
0.4	56.35	46.84	40.02	34.96	22.63	**19.02**
0.5	36.62	29.65	24.96	21.67	**14.80**	15.09
0.6	25.13	20.04	16.81	14.67	**10.98**	11.35
0.7	18.09	14.34	12.09	10.69	**8.91**	9.97
0.8	13.58	10.78	9.20	8.28	**7.67**	9.07
0.9	10.56	8.44	7.32	**6.74**	6.86	8.40
1.0	8.46	6.85	6.06	**5.60**	6.28	7.84
1.5	3.90	3.45	**3.36**	3.46	4.57	5.53

TABLE IV

ARL OF MA CHART FOR GIVEN $\mu_0 = 3$, $\alpha_0 = 0.2$ AND $ARL_0 = 500$ CONSIDERING A CHANGE IN μ.

δ	w=2	w=3	w=4	w=5	w=10	w=15
0.0	500.11	500.11	500.11	500.11	500.11	500.11
0.1	243.55	223.94	206.97	192.15	140.33	**110.22**
0.2	104.94	86.60	73.30	63.32	37.78	**28.66**
0.3	49.22	38.03	30.85	25.98	16.14	**14.55**
0.4	26.03	19.53	15.74	13.40	**9.99**	10.76
0.5	15.36	11.51	9.47	8.35	**7.65**	9.14
0.6	9.95	7.60	6.49	**5.99**	6.48	8.04
0.7	6.97	5.50	4.91	**4.73**	5.71	7.08
0.8	5.19	4.27	**3.97**	3.98	5.11	6.16
0.9	4.08	3.50	**3.39**	3.49	4.57	5.31
1.0	3.34	**2.97**	2.98	3.13	4.08	4.54
1.5	**1.82**	1.85	1.96	2.07	2.30	2.32

interpretable model for correlated process of Poisson counts data. We derived explicit formula of moving average control chart for evaluate average run length. The result show that when out - of - control for the shift increasing MA performs better as the value of (w) decreases. Consequently, calculations with explicit formulas is simple and very fast with computational times of less than 1 second.

V. DISCUSSION

The good properties of MA chart are memory control chart then it is good for small and moderate shifts. Without loss of generality, this chart can be relax due to its feasibility with width (w) of control limit. The performance of MA chart will be better as the values of w increases for small shift, however, the number of observations must be sufficient large.

ACKNOWLEDGMENT

The author would like to express my gratitude to the graduate school of King Mongkut's University of Technology, North Bangkok, Thailand. We also thanks to Department of Applied Statistics for supporting materials and high performance computer.

REFERENCES

[1] D.C. Montgomery, *Statistical Quality Control* (6th ed.). New York: John Wiley & Sons. 2009.

[2] T.E. Harris, *The Theory of Branching Processes*. Berlin: Springer. 1963.

[3] P.J. Brockwell and R.A. Davis, *Time Series: Data Analysis and theory* (2nd ed.). New York: Springer. 2009.

[4] E.S. page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100-144, 1954.

[5] S.W. Roberts, "Control chart tests based on geometric moving average," *Technometrics*, vol. 1, pp. 239-250, 1959.

[6] M.B.C. Khoo, "A moving average control chart for monitoring the fraction non-conforming," *Journal of Quality and Reliability Engineering International*, vol. 20, pp. 617-635, 2004.

[7] E. McKenzie, "Some simple models for discrete variate time series," *Water Resources Bulletin*, vol. 21, pp. 645-650, 1985.

[8] E. McKenzie, *Discrete Variate Time Series in Handbook of statistics eds*. D.N. Shanbhag, and C.R. Rao, Amsterdam: Elsevier, vol. 20, pp. 573-606, 2003.

[9] M.A. Al.Osh and A.A. Alzaid, "First-order integer-valued autoregressive (INAR(1)) process," *Journal of Time Series Analysis*, vol. 8, pp. 261-275, 1987.

[10] S.V. Crowder," A simple method for studying run length distributions of exponentially weighted moving average charts," *Technometrics*, vol. 29, pp. 401-407, 1979.

[11] Y. Areepong and A.A. Novikov, "Martingale approach to EWMA control chart for changes in Exponential distribution," *Journal of Quality Measurement and Analysis*, vol. 4, pp. 197-203, 2008.

[12] Y. Areepong, "Explicit formulas of average run length for a moving average control chart for monitoring the number of defective product," *International Journal of Pure and Applied Mathematics*, vol. 3, pp. 331-343, 2012.

[13] F.W. Steutel and K.V. Harn, "Discrete analogues of self decomposability and stability," *Annals of Probability*, vol. 7, pp. 839-899. 1979.

[14] C.H. Weiβ, "Thinning operations for modelling time series of counts a survey," *Advances in Statistical Analysis* (*AStA*), vol.92, pp. 319-341, 2008.

[15] C.H. Weiβ, "Controlling correlated processes of Poisson counts," *Quality Reliability Engineering International*, vol. 23, pp. 741-754, 2007.