A Study of the Performance of EWMA Chart with Transformed Weibull Observations

Saowanit Sukparungsee *[a] and Yupaporn Areepong [b]

[a] Department of Applied Statistics, Faculty of Applied Science, King Mongkut's
   University of Technology North Bangkok, Bangkok 10800, Thailand.

[b] Department of Mathematical Sciences, Faculty of Science, University of Technology,
   Sydney NSW, 2007, Australia.

*Author for correspondence; e-mail: saowanits@hotmail.com, swns@kmutnb.ac.th

Abstract

Statistical Process Control (SPC) chart for detecting small shifts in parameter of distributions are widely used in quality control and other area of applications is an Exponentially Weighted Moving Average (EWMA) procedure. The objective of this paper is to implement an explicit formula for characteristics of EWMA as Average Run Length (ARL) – the expectation of false alarm times and Average Delay time (AD) – the expectation of delay of true alarm times in case of Weibull distribution. Using the simple transformation technique, we obtain the explicit expressions for evaluating ARL and AD when observations are Weibull by taking power of such observations. The accuracy of results is compared with Monte Carlo simulations. In addition, we present the table of optimal parameter values for Weibull EWMA designs and the comparisons of performance of EWMA versus CUSUM charts are considered.

_____

Keywords: averaged run length and average delay, cumulative sum, exponentially weighted moving average chart, stopping times.

1.   Introduction

   Statistical Process Control (SPC) charts are widely used for monitoring, measuring, controlling and improving quality of production in many areas of applications, for in stance industry and manufacturing, finance and economics, epidemiology and health care, environmental sciences and in other fields. Typically, SPC charts are studied under assumption that observations are Normal distributed. In many real applications, however, observations are frequently non-Normal, e.g. with Bernoulli, Poisson, Exponential, and Weibull distributions.

   One of the popularly important characteristics for SPC charts is *Average Run Length (ARL)* – the expectation of an *alarm time* (τ) is taken to signal (wrongly) about a possible change. Ideally, an acceptable ARL of in-control process should be enough large and a small ARL when the process is out-of-control, so-called *Average Delay* (AD) - the expectation of delay for true alarm time.

   In literature one can find many methods for evaluating ARL and AD for EWMA procedure have been studied. Robert [1], the first who introduced the ARL of EWMA by using simulations derived nomograms for the ARL in case of Gaussian distribution. Brook and Evans [2] approximated the run length of EWMA by using a finite-state *Markov Chain Approach* (MCA). Crowder [3] used a system of *Integral Equations* (IE) to find both ARL and AD. Later, Lucas and Saccucci [4] evaluated ARL by using a finite-state MCA similar to Brook and Evans [2] and also studied the optimal EWMA designs. Borror [5] examined the ARL performance of EWMA chart for both skewed and heavy-tailed symmetric non-normal distributions using MCA. EWMA control charts for Exponential distribution are introduced by Gan [6] who calculated the ARL by using the differential equations. Sukparungsee and Novikov [7, 8] derived an analytical closed-form formula for determining the characteristics of EWMA charts for the cases of Gaussian and some non-Gaussian distributions by use of a martingale-based technique. Recently, the explicit formulas of ARL and AD for Exponential EWMA charts have been found by Areepong and Novikov [9].

   MC is simple to program but usually it is related to a large number of sample trajectories (very time consuming). Moreover, it is difficult to study for optimal designs though it is convenient to control accuracy of analytical approximations.

   MCA is considered as a most popular technique (Lucas and Saccucci [4]). It is based on use of matrix inversions for approximating Markov Chain. As far as we know there are no theoretical results on accuracy of this procedure besides just direct comparisons with MC.

IE is the most advanced method but it requires intensive programming or special software to implement even for the case of Gaussian distribution [3, 10, 11].

Martingale-based technique is innovation tool, effective alternatives to traditional approaches, fast and easily to implement but it can be derived the chart characteristics for the case of light-tailed distributions.

In this paper we derived analytical formulas for ARL and AD of EWMA charts when observations are Weibull distributed. Originally, our derivations are based on results of Novikov [12] who found explicit formulae for the expectations of alarm times for first-order autoregressive processes and Areepong and Novikov [9] who implemented to evaluating ARL and AD for Exponential EWMA charts from the former. We suggest implementing for derivative chart characteristics of Weibull EWMA charts by transforming from Weibull to Exponential distribution. In addition, we compare our analytical results for ARL and AD with results from simulations and the optimal parameter values for Weibull EWMA designs are addressed.

2. EWMA and CUSUM Procedures and Theirs Properties.

In this paper we consider SPC charts under the assumption that sequential observations $\xi_1, \xi_2, \ldots,$ are independent random variables with a distribution function $F(x, \alpha)$, the parameter $\alpha = \alpha_0$ before a change-point time $\theta \leq \infty$ ("in-control" state; $\theta = \infty$ means that there are no change at all) and $\alpha > \alpha_0$ after the change-point time θ ("out-of-control" state).

All popular charts like Shewhart, Cumulative Sum (CUSUM) and EWMA charts (see e.g. [13-15] are based on use of stopping times τ. The typical condition on choice of the stopping times τ is the following:

$$E_\infty(\tau) = T, \tag{1}$$

where T is given (usually large). Let $E_\infty(.)$ denote that the expectation under distribution $F(x, \alpha_0)$ (in-control) that the change-point occurs at point θ (where $\theta \leq \infty$). In literature on quality control the quantity $E_\infty(\tau)$ is called as Average Run Length (ARL) of the algorithm. Then, by definition, $ARL = E_\infty(\tau)$ and the typical practical constraint is

$$ARL = T.$$

Another typical constraint consists in minimizing of the quantity

$$Q(\alpha) = \sup_\theta E_\theta(\tau - \theta + 1 | \tau \ge \theta), \qquad (2)$$

where $E_\theta(.)$ is the expectation under distribution $F(x, \alpha)$ (out-of-control) and α is the value of parameter after the change-point. We restrict on the special case, usually $\theta = 1$. The quantity $E_1(\tau)$ is called as average delay time (AD) and one could expect that a sequential chart has a near optimal performance if its AD is close to a minimal value.

They are many other criteria for optimality of SPC (see e.g. [16-18]); however, in practice, ARL and AD remain the most popular characteristics which are convenient to use for comparisons of different charts.

EWMA chart is defined as a recursive form

$$X_t = (1 - \lambda)X_{t-1} + \lambda\eta_t, \quad t = 1, 2, \dots. \qquad (3)$$

Typically, smoothing parameter $\lambda \in (0, 1)$, $\eta_t = g(\xi_t)$ and X_t is the weighted average between current and previous observations. The target mean is supposed to be steady and the initial value X_0 is usually chosen to be the process mean α_0. The alarm time for this type of procedure is the following:

$$\tau_h = \inf\{t > 0 : X_t > h\}.$$

The ARL of the EWMA chart depends on the control limit (h) and the smoothing parameter (λ). Moreover, the optimal design parameters (λ, h) are given by minimizing AD under constraint (1). In this paper we compare our results with CUSUM procedure, thus we shall briefly overview the CUSUM chart.

The standard **CUSUM** chart is defined by the following statistics

$$Y_t = \max(Y_{t-1} + q(\xi_t), 0), \quad t = 1, 2, \dots \quad Y_0 = y, \qquad (4)$$

where

$$q(\xi_t) = \log \frac{dF(x, \alpha)}{dF(x, \alpha_0)} = \log \frac{f(x, \alpha)}{f(x, \alpha_0)}$$

and Y_0 is an initial value.

Under the assumptions that the $F(x,\alpha)$ is absolute continuous with respect to $F(x,\alpha_0)$. In literature one could find some modification of algorithms (see in [13]). The alarm time for this type of procedure is typically

$$\tau_A = \inf\{t > 0 : Y_t > A\},$$

where A is a control limit.

CUSUM is usually considered as a candidate for the optimal algorithm. Its performance could be below the performance of EWMA [4,6-8] for moderate values of ARL and detecting small changes.

3. Explicit Formulas for Evaluating ARL and AD for Weibull EWMA Chart.

The first-order autoregressive *(AR(1))* process is defined as a solution of equation

$$Z_t = \rho Z_{t-1} + \eta_t; \quad t = 1, 2, \dots \text{ and } Z_0 = z, \tag{5}$$

where we assume that independent random variables η_t have the standard Exponential distribution $(\eta_t \sim Exp(1))$, z and ρ are given constants and Z_0 is initial value.

Set $\tau_b = \inf\{t > 0 : Z_t > b\}$. For this case Novikov [12] obtained the following formula

$$E(\tau_b) = Q(b/\rho) + 1 - Q(z), \tag{6}$$

where

$$Q(z) = \sum_{m=1}^{\infty} \frac{(\rho z)^m (\rho,\rho)_{m-1}}{m!}. \tag{7}$$

The function $(\rho,\rho)_{m-1}$ in Equation (3) is q-Pochhamer symbols from the theory of q-series (see e.g. [Andrews et al. [19]):

$$\prod_{j=1}^{m}(1 - \gamma\rho^{j-1}) := (\gamma,\rho)_m, \quad (\gamma,\rho)_0 = 1.$$

In Equation (5), setting $\rho = 1 - \lambda$ and $z = \dfrac{1}{\lambda}$ one can check that $\tau_b = \gamma_{h/\lambda}$. This implies that ARL of the one-sided EWMA procedure for the case of Exponential distribution with $\alpha_0 = 1$ is the following

$$ARL = E_\infty(\tau_h) = Q\left(\frac{h}{\lambda(1-\lambda)}\right) + 1 - Q\left(\frac{1}{\lambda}\right), \qquad (8)$$

where

$$Q(z) = \sum_{m=1}^{\infty} \frac{\left((1-\lambda)z\right)^m (1-\lambda, 1-\lambda)_{m-1}}{m!}. \qquad (9)$$

Similar, the closed-form formula for AD (that is when observations are exponentially distributed with parameter α is

$$AD = E_1(\tau_h) = Q\left(\frac{h}{\alpha\lambda(1-\lambda)}\right) + 1 - Q\left(\frac{1}{\alpha\lambda}\right), \qquad (10)$$

where $Q(z)$ is defined in Equation (9). Some of these results have already been published in Areepong and Novikov [9].

Let $\xi_t, t = 1, 2, ...$ be sequentially observed independent random variables. The change-point models are the following:

$$\xi_t \sim \begin{cases} Weibull(r, \alpha_0) & ; t = 1, 2, ..., \theta - 1 \\ Weibull(r, \alpha) & ; t = \theta, \theta + 1, ..., \alpha > \alpha_0. \end{cases}$$

A Weibull distribution is defined by the following function:

$$P(\xi_t > x) = e^{-\left(\frac{x}{\alpha}\right)^r} ; x \geq 0.$$

A common method used to transform the Weibull distribution to the Exponential distribution is to take power r of the Weibull observations. For example, if the Weibull distribution has the form:

$$P(\xi_t > x) = e^{-\left(\frac{x}{\alpha}\right)^r},$$

then taking the power r we get

$$P\left(\xi_t^r > x\right) = P\left(\xi_t > x^{\frac{1}{r}}\right)$$

$$= e^{-\left(\frac{x^{\frac{1}{r}}}{\alpha}\right)^r}$$

$$= e^{-\left(\frac{x}{\alpha^r}\right)}$$

and therefore $P(\xi_t^r > x)$ has the Exponential distribution with $E(\xi_t^r) = \alpha^r$.

Therefore, if $\xi_t \sim Weibull(r, \alpha)$ then $\xi_t^r \sim Exp(\alpha^r)$ and $\eta_t = \dfrac{\xi_t}{\alpha^r} \sim Exp(1)$.

For the Weibull distribution, we define the EWMA recurrence relation by:

$$X_t = (1 - \lambda)X_{t-1} + \lambda\alpha^r\eta_t, \quad X_0 = 1 \tag{11}$$

with the stopping time

$$\tau_h = \inf\{t > 0 : Z_t > h\}.$$

Equation (11) can be transformed to the *AR(1)* form of Equation (5) with $\eta_t = \dfrac{\xi_t}{\alpha^r} \sim Exp(1)$ by the substitutions

$$\tilde{X}_t = \frac{X_t}{\lambda\alpha^r}, \qquad \tilde{X}_0 = \frac{1}{\lambda\alpha^r}, \qquad \rho = (1 - \lambda), \qquad \eta_t = \frac{\xi_t}{\alpha^r}.$$

The EWMA equation is then:

$$\tilde{X}_t = (1 - \lambda)\tilde{X}_{t-1} + \eta_t, \quad \tilde{X}_0 = \frac{1}{\lambda\alpha^r}$$

with the stopping time

$$\tau_H = \inf\left\{t > 0 : \tilde{X}_t > \frac{h}{\lambda\alpha^r}\right\}.$$

Therefore, it can be seen that the ARL and AD for the case of a Weibull distribution can also be calculated by using the closed-form formulas for the Exponential distribution. The explicit formulas are as follows:

$$ARL = E_\infty(\tau_h) = Q\left(\frac{h}{\lambda(1-\lambda)}\right) + 1 - Q\left(\frac{1}{\lambda}\right), \tag{12}$$

$$AD = E_1(\tau_h) = Q\left(\frac{h}{\alpha^r \lambda(1-\lambda)}\right) + 1 - Q\left(\frac{1}{\alpha^r \lambda}\right), \tag{13}$$

where $Q(x)$ is defined in Equation (9).

4. Comparison of Formulas with Simulations

In Table 1, we compare the numerical results for ARL and AD with Equation (12) and (13) with Monte Carlo simulation both for EWMA with $h = 1.76672, \lambda = 0.09206$ and CUSUM with $A = 4.495$. We always simulate 10^6 sample trajectories by using the package R.

Obviously, the numerical results from suggested formulas are very close to Monte Carlo simulation. Besides, EWMA and CUSUM show a similar performance when α is greater than 1.4. However, when α relatively close to 1 (in-control value), EWMA shows a better performance than the corresponding CUSUM scheme. Note that, calculations with exact formulas (12) and (13) are much faster. For example, when $\alpha = 1$, computing time based on our technique takes less than 1 second while CPU time required for simulations for both EWMA and CUSUM run 23843.1 and 41912.3 seconds respectively as shown in Table 2.

The numerical procedure for obtaining optimal parameters for EWMA designs

1. Select an acceptable in-control value of ARL and decide on the change parameter value α for an out-of-control state.

2. For given α and T, find optimal values of λ^* and h^* to minimise the AD values given by Equation (13) subject to the constraint that ARL=T in Equation (12), i.e. λ^* and h^* are solutions of the optimality problem:

$$\min_{\lambda,h} AD = Q\left(\frac{h}{\alpha^r \lambda(1-\lambda)}\right) + 1 - Q\left(\frac{1}{\alpha^r \lambda}\right).$$

Subject to:

$$ARL = T = Q\left(\frac{h}{\lambda(1-\lambda)}\right) + 1 - Q\left(\frac{1}{\lambda}\right),$$

where $Q(x)$ is defined in Equation (9).

In Table 3 we present AD as a function of λ for ARL=500, 1000, 3000 and 5000. For example, given ARL=500 and $\alpha = 1.5$, calculations with formula (13) give AD^*=9.333 for the optimal set of parameters $\lambda^* = 0.10250, h^* = 1.72788$. This optimal AD value agrees with the value of AD^*=9.345 that we computed from MC for the same parameter values.

Table 1. Comparison ARL and AD by suggested formulas with Monte Carlo simulations

(r, α^2)	ARL and AD		
	Formula (12),(13)	Monte Carlo simulations	
		EWMA	CUSUM
$(2,1.0^2)$	999.861	999.610 [0.996][a]	1000.930 [1.008]
$(2,1.1^2)$	138.679	138.720 [0.066]	158.420 [0.158]
$(2,1.2^2)$	45.731	45.710 [0.023]	49.028 [0.045]
$(2,1.3^2)$	23.496	23.495 [0.013]	23.753 [0.027]
$(2,1.4^2)$	15.074	15.070 [0.008]	14.296 [0.011]
$(2,1.5^2)$	*10.915*	*10.926 [0.006]*	*10.289 [0.008]*
$(2,1.6^2)$	8.500	8.507 [0.005]	7.718 [0.006]
$(2,1.7^2)$	6.945	6.942 [0.004]	6.312 [0.004]
$(2,1.8^2)$	5.869	5.864 [0.004]	5.291 [0.004]
$(2,1.9^2)$	5.085	5.087 [0.003]	4.561 [0.003]
$(2,2.0^2)$	4.491	4.491 [0.003]	3.999 [0.003]
$(2,2.5^2)$	2.897	2.897 [0.002]	2.652 [0.002]
$(2,3.0^2)$	2.217	2.217 [0.002]	2.025 [0.001]
$(2,5.0^2)$	1.394	1.395 [0.001]	1.328 [0.001]

[a]standard deviation.

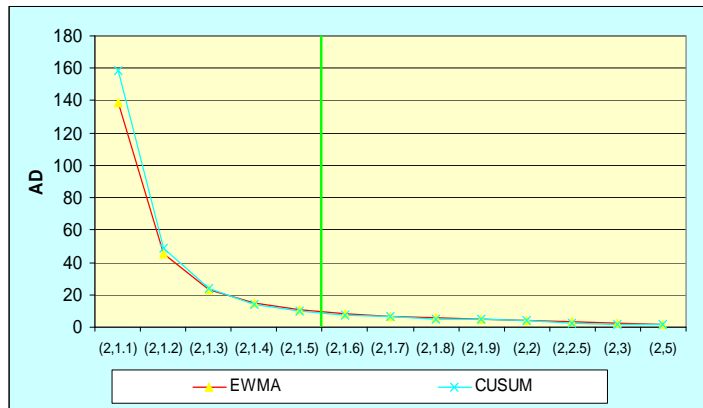**Figure1:** Comparison of AD calculated by explicit formula for EWMA and simulated by CUSUM for Weibull observations

Table 2. Comparison CPU Times by suggested formulas with Monte Carlo simulations

(r, α^2)	CPU Times (seconds)		
	Formula (12),(13)	Monte Carlo simulations	
		EWMA	CUSUM
$(2,1.0^2)$	0.13	23843.1	41912.3
$(2,1.1^2)$	0.14	3125.8	6626.7
$(2,1.2^2)$	0.13	1217.5	2163.1
$(2,1.3^2)$	0.14	577.4	1075.4
$(2,1.4^2)$	0.13	377.1	681.1
$(2,1.5^2)$	*0.14*	*266.3*	*486.2*
$(2,1.6^2)$	0.14	211.7	370.0
$(2,1.7^2)$	0.14	181.2	308.4
$(2,1.8^2)$	0.13	153.3	264.8
$(2,1.9^2)$	0.14	135.8	230.0
$(2,2.0^2)$	0.13	115.2	196.5
$(2,2.5^2)$	0.13	86.2	131.3
$(2,3.0^2)$	0.13	64.2	109.3
$(2,5.0^2)$	0.13	46.6	68.5

Table 3. Optimal parameters and AD of one-sided Weibull EWMA by ==suggested== formula.

T	(r,α^2)	λ^*	h^*	AD^*	AD by simulations
T=500	$(2,1.5^2)$	0.10250	1.72788	9.333	9.345 [0.021][b]
	$(2,1.7^2)$	0.15406	2.00271	5.997	5.998 [0.004]
	$(2,2.0^2)$	0.22673	2.36935	3.853	3.852 [0.009]
	$(2,3.0^2)$	0.42078	3.31407	1.916	1.974 [0.001]
T=1000	$(2,1.5^2)$	0.09206	1.76672	10.915	10.901 [0.023]
	$(2,1.7^2)$	0.13805	2.04556	6.849	6.855 [0.004]
	$(2,2.0^2)$	0.20423	2.42482	4.294	4.292 [0.003]
	$(2,3.0^2)$	0.38662	3.43084	2.039	2.039 [0.001]
T=3000	$(2,1.5^2)$	0.07632	1.79212	13.514	13.485 [0.028]
	$(2,1.7^2)$	0.11551	2.07748	8.238	8.234 [0.005]
	$(2,2.0^2)$	0.17311	2.47226	5.010	5.012 [0.003]
	$(2,3.0^2)$	0.33784	3.55300	2.236	2.234 [0.001]
T=5000	$(2,1.5^2)$	0.07010	1.79671	14.751	14.747 [0.029]
	$(2,1.7^2)$	0.10668	2.08446	8.896	8.895 [0.006]
	$(2,2.0^2)$	0.16087	2.48471	5.347	5.346 [0.004]
	$(2,3.0^2)$	0.31796	3.59266	2.329	2.331 [0.002]

[b]standard deviation.

5. Conclusion

We have presented that the explicit formulas for ARL and AD of one-sided EWMA charts for the case of an Exponential distribution can be applied to Weibull distribution. We have shown that suggested formulas are very accurate, and are easy to calculate and program. The ==suggested== formulas obviously take the computational ==times== much less than Monte Carlo (MC) simulation. Using the formulas, we have been able to provide tables for the optimal weights, boundaries and approximations for ARL and AD for one-sided EWMA charts for the Weibull distribution. The performance comparison of the control charts has been based on Average Run Length (ARL) and Average Delay (AD) criteria. For Weibull distribution when given ARL=1000, we have shown that the performance of EWMA chart is superior to CUSUM for small changes. On the contrary, the performance of EWMA is inferior to CUSUM chart for moderate to large changes.

References

[1] Roberts, S.W., Control chart tests based on geometric moving average. *Technometrics*, 1959, 1: 239–250.

[2] Brook, D. and Evans, D.A., An approach to the probability distribution of Cusum run length. *Biometrika*, 1972, 59: 539-548.

[3] Crowder, S.V., A simple method for studying run length distributions of exponentially weighted moving average charts. *Technometrics*, 1978, 29: 401-407.

[4] Lucas, J.M. and Saccucci, M.S., Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, 1990, 32: 1-29.

[5] Borror, C.M., Montgomery D.C. and Runger G.C., Robustness of the EWMA Control Chart to Non-normality. *Journal of Quality Technology*, 1999, 31: 309-316.

[6] Gan, F.F., Designs of one- and two-sided exponential EWMA chart. *Journal of Quality Technology*, 1998, 30: 55-69.

[7] Sukparungsee, S. and Novikov, A.A., On EWMA procedure for detection of a change in observations via martingale approach. *KMITL Science Journal; An International Journal of Science and Applied Science*, 2006, 6: 373-380.

[8] Sukparungsee, S., and Novikov, A.A., Analytical approximations for detection of a change-point in case of light-tailed distributions. *Journal of Quality Measurement and Analysis*, 2008, 4: 373-380.

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

[10] Knoth, S., Accurate ARL calculation for EWMA control charts monitoring normal mean and variance simultaneously. *Sequential Analysis*, 2006, 26: 251–263.

[11] Srivastava, M.S. and Wu, Y., Evaluation of optimum weights and average run lengths in EWMA control schemes. *Communications in Statistics: Theory and Methods*, 1997, 26: 1253-1267.

[12] Novikov, A.A., *Levy-driven Ornstein-Uhlenbeck processes: survey of results on first passage times.* Lecture notes presented at the conference "Stochastic Calculus with Jumps". University of Angers, May 3-9, 2006.

[13] Hawkins, D.M. and Olwell, D.H., *Cumulative Sum Charts and Charting for Quality Improvement*. New York, Springer Verlag, 1998.

[14] Page, E.S., Continuous inspection schemes. *Biometrika*, 1954, 41: 100–114.

[15] Woodall, W.H. and Adams, B.M., The statistical design of CUSUM charts. *Quality Engineering*, 1993, 4: 559–570.

[16] Lorden, G., Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, 1971, 42: 1897-1908.

[17] Shiryaev, A.N., The problem of the most rapid detection of a distribution in a stationary process. *Soviet Mathematics*, 1961, 2: 795-799.

[18] Shiryaev, A.N., *Optimal Stopping Rules*. Springer-Verlag, 1978.

[19] Andrews, G.E., Askey, R. and Roy, R., *Special Functions*. Cambridge, University Press, 1999.