Evaluation of Optimum Weights and Average Run Lengths in EWMA Control Schemes

by

M.S. Srivastava
Department of Statistics
University of Toronto

and

Yanhong Wu
Department of Statistics
University of Alberta

Technical Report No. 9503, April 20, (1995)

TECHNICAL REPORT SERIES

University of Toronto

Department of Statistics

Evaluation of Optimum Weights and Average Run Lengths in EWMA Control Schemes *

M.S. Srivastava
University of Toronto

Yanhong Wu
University of Alberta

Abstract

In this paper, a two sided exponentially weighted moving average (EWMA) control scheme of Roberts (1959) is studied. By using a continuous time model, explicit formulae for the average run lengths are obtained. Using these results, a formula for the optimal weight in the EWMA is obtained by minimizing the out of control average run length for a given reference value and in control average run length (ARL_0), which is assumed to be large. These optimum weights are found to be very close to the optimum weights obtained by Lucas and Saccucci (1990) for the discrete time model using the Markov chain method. It is shown that the EWMA scheme performs almost as well as the CUSUM for moderate ARL_0. Another important finding that emerges from this study is that the EWMA scheme is less sensitive to the choice of a reference value as compared to the CUSUM. We also provide very accurate formulae for the ARL_0 and ARL_μ of the two-sided CUSUM procedure.

KEY WORDS: Average run lengths, CUSUM procedure, exponentially weighted moving average, Shift in the mean, Two-sided procedures.

*Research supported by NSERC of Canada

1 Introduction

In a variety of situations, such as in quality control, detecting a shift in the mean of a process as soon as it occurs is an important problem. Among the many procedures, the CUSUM procedure proposed by Page (1954) is the most popular procedure and its properties have been thoroughly investigated [see, for example Van Dobben de Bryn (1968), Lucas (1976), Pollak and Siegmund (1985), and Srivastava and Wu (1993)]. Another procedure, proposed by Roberts (1959), and known as the EWMA (exponentially weighted moving average) control scheme has recently been receiving considerable attention in the literature [Crowder (1987); Hunter (1986); Lucas and Saccucci (1987,1990); Montgomery, Gardiner and Pizzano (1987); Robinson and Ho (1978); Waldmann (1986); Srivastava and Wu (1993)]. It seems to us that the main reason for this is that the EWMA control scheme is not only as simple to apply as CUSUM, but it can be used to monitor a shift in either direction at the same time as compared to the CUSUM scheme which requires monitoring two schemes at the same time. In addition, it can be used to estimate the current mean, a feature appreciated by engineers. Any doubt that may have remained about the EWMA procedure regarding its performance was layed to rest by an extensive comparison with the CUSUM by Lucas and Saccucci (1990), where it was shown that the EWMA performs as good as the CUSUM procedure. It was further confirmed by Srivastava and Wu (1993) in the one-sided EWMA scheme for moderate ARL_0.

While Lucas and Saccucci (1990) have provided extensive tables for the decision boundary, average run lengths, and optimum weights, it may be of interest to have simple formulae for these quantities, which may make EWMA easier to use. However, the evaluation of these quantities is difficult for the discrete time model. Thus, we consider instead a continuous time version of the EWMA scheme and obtain simple asymptotic formulae for the decision boundary and optimum weights. A minor empirical correction is applied to some of these formulae to obtain results for the discrete time model. The results are very close to the numerical values given in Lucas and Saccucci (1990).

The optimum weights are obtained by minimizing the 'out-of-control' average run length, denoted by ARL_δ, in which a shift in the amount of δ occurs on the first observation, subject to a fixed 'in-control' average run length, denoted by ARL_0, in which there is no shift in the process. The amount of shift δ is assumed known and is called a reference value. It is selected in such a way that any shift larger than this value will not be acceptable. We show that the optimal weight is less sensitive to the choice of a reference value.

The organization of the paper is as follows. In Section 2, we describe the EWMA control scheme in the two-sided case and evaluate the average run lengths. The results for the CUSUM procedure are stated in Section 3, where it is shown that the simple formulae for the average run lengths provide very accurate approximations. The comparison of EWMA with CUSUM is carried out in Secion 4. The paper concludes in Section 5, with some technical details given in an appendix.

2 Two-sided EWMA control scheme

Suppose $\{Z_k\}$, $k = 1, 2, \ldots$ is the observation process which is a normal sequence of independent observations with unit variance. Let

$$\tilde{y}_k = (1-\lambda)\tilde{y}_{k-1} + \lambda Z_k, \quad k \geq 1, \qquad \tilde{y}_0 = 0, \quad 0 \leq \lambda \leq 1. \tag{2.1}$$

Then it can be shown that the variance of $\tilde{y}_k$ converges to $\lambda/(2-\lambda)$. Let

$$y_k = \frac{\tilde{y}_k}{[\lambda/(2-\lambda)]^{\frac{1}{2}}}$$

Then the EWMA process is given by

$$y_k = (1-\lambda)y_{k-1} + [\lambda(2-\lambda)]^{1/2} Z_k, \qquad k \geq 1, \quad y_0 = 0, \quad 0 \leq \lambda \leq 1$$

The process is stopped and checked (for any shift in the mean) at the smallest value of n for which $|y_n| \geq L$. That is, our stopping rule for detecting a shift in either direction is given by

$$N = \min\{n \geq 1 : |y_n| \geq L\},$$

where L is the control limit or the decision boundary. Thus given L and λ, we can determine the 'in-control' average run length, ARL_0, which is the average number of observations taken before the process goes out of the control limit, even when there has been no shift. Usually, this number is chosen large, and for a given λ, provides the decision boundary L. A measure of performance is the average number of observations required to detect the shift if the shift has occurred on the first observation. This average run length will be denoted by ARL_μ, where μ is the amount of shift which has occurred on the first observation.

In order to evaluate ARL_0, we consider the continuous time analog, as in Srivastava and Wu (1993).

We find that for the continuous model, the 'in-control' average run length is given by

$$ARL_0 = (2\lambda)^{-1} \int_0^L [\phi(z)]^{-1}[2\Phi(z) - 1]dz, \tag{2.2}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal probability density and cumulative distribution functions respectively. While this is an exact result for the continuous time model, some adjustments are needed to correct for the overshoot of the boundary in the discrete time model, since unlike the continuous time model, the boundary cannot be reached exactly in this case. Expanding the integral as shown in the Appendix and empirically correcting for the overshoot, we find that for large $T = ARL_0$, the following simple formula provides a very good approximation for the decision boundary in the discrete time model.

$$L \approx [a - \log(a-1)]^{\frac{1}{2}} + \frac{1}{2}(1-\lambda) \tag{2.3}$$

where

$$a = 2\log\left[\left(\frac{2}{\pi}\right)^{\frac{1}{2}} T\lambda\right], \quad T = ARL_0$$

A comparison of this approximation with the values as reported in L-S is presented in Table 1. Here, and in the table below, L-S stands for Lucas and Saccucci (1990).

$\lambda =$	1.00	.75	.50	.40	.30	.25	.20	.10	.05
L =	3.095 (3.090)	3.135 (3.087)	3.136 (3.071)	3.116 (3.054)	3.074 (3.023)	3.040 (2.998)	2.990 (2.962)	2.800 (2.814)	2.568 (2.615)
μ	ARL_μ	ARL_μ	ARL_μ	ARL_μ	ARL_μ	ARL_μ	ARL_μ	ARL_μ	ARL_μ
0.5	154 – – – (201)	127 109^b (140)	89 74^b (89)	73 59^b (71)	57 45^b (55)	49 39^b (48)	42 32^b (42)	31 34^b (31)	28 30 (29)
1.0	60 – – – (55)	39 31^b (31)	22 17^b (18)	18 13^b (14)	14 12^b (12)	13 14^b (11)	12 11 (11)	11 10 (10)	11 11 (11)
2.0	11 – – – (7)	7 4 (5)	5 4 (4)	5 4 (4)	4 4 (4)	4 4 (4)	4 4 (4)	5 4 (4)	5 5 (5)
3.0	4 – – – (2)	3 2 (2)	3 2 (2)	3 2 (2)	3 2 (2)	3 3 (2)	2 3 (2)	3 3 (3)	4 4 (4)

Table 1: Approximations to decision boundary and ARL_μ with $ARL_0 = 500$ for various μ and λ. Values on the first two rows of each cell are from (2.4) and (2.4a,b) respectively. The entries with b are from (2.4b). The bracketed values are from L-S.

2.1 Out of Control Average Run Length, ARL_μ

To compute ARL_μ, when μ is the amount of shift taking place on the first observation, we again appeal to the continuous time model. Using the empirical correction for the overshoot, we find that

$$ARL_\mu = (\lambda)^{-1} \left[h(\eta)f(-\eta) + h(-\eta)f(\eta)\right], \tag{2.4}$$

where

$$\alpha(\eta) = \int_0^{L_\mu} [\phi(x+\eta)]^{-1}\, dx, \qquad h(\eta) = \alpha(\eta)/[\alpha(\eta) + \alpha(-\eta)],$$

$$f(\eta) = \int_0^{L_\mu} [\phi(z+\eta)]^{-1}[\Phi(z+\eta) - \Phi(\eta)]dz.$$

$$\eta = \mu(2/\lambda)^{1/2},$$

$$L_\mu = L + 1.16(\lambda\mu)^{\frac{1}{2}},$$

and L is as in (2.2) for a given ARL_0 and λ. Clearly, when $\mu = 0$, L_μ reduces to L. Thus, the additional term in L is an empirical adjustment for the overshoot when the process is out of control. The numerical integration in (2.4) can easily be done by using, for instance, the Mathematica language. Values of ARL_μ obtained from (2.4) are reported in Table 1, where the corresponding values from Lucas and Saccucci (1990) are also given.

The integral in (2.4) can be expanded (with a minor empirical adjustment of adding 3/4 in (2.4a)), and an approximation for ARL_μ can be obtained. Let

$$b = L\left(\frac{\lambda}{2\mu^2}\right)^{\frac{1}{2}} \text{ and } w = L + 1.166(\mu\lambda)^{\frac{1}{2}} - \left(\frac{2\mu^2}{\lambda}\right)^{\frac{1}{2}}.$$

Then,

(a) For $b < 1$ and $0 < \lambda \leq .75$,

$$ARL_\mu \approx -\left(\frac{1}{\lambda}\right)\log(1-b) - \frac{b}{4(1-b)\mu^2} + \frac{3}{4}, \tag{2.4a}$$

(b) For $b > 1$, $0 < \lambda \leq .75$, and $\mu \geq 1$,

$$ARL_\mu \approx \frac{1}{\lambda w}\left[\phi(w)\right]^{-1}\Phi(w), \tag{2.4b}$$

These approximations are compared with the results of Lucas and Saccucci (1990) and the one obtained from (2.4). Clearly, the simple formula in (2.4a) provides a very good approximation as compared to the one in (2.4b), which is very good only when $\mu \geq 1$. However, in most cases we would like to detect a shift of one standard deviation in which case $\mu = 1$. In addition, λ will be chosen close to its optimum value. Thus, for all practical purposes, we may use (2.4a,b) in place of (2.4), which requires a numerical integration.

2.2 Optimal Weights and Control Limit

In the previous section, we assumed that both λ and the 'in-control' average run length ARL_0 are given. We determined the decision boundary L for these given values and studied the performance of the EWMA procedure through ARL_μ. From Table 1, it is clear that smaller values of λ will be preferable for detecting small shifts. In many situations, it is often known that a product may not be acceptable if the shift in the mean exceeds a certain given amount, say δ; δ is usually called a 'reference value' in the literature. In such situations, we may wish to determine an optimum value of λ which minimizes ARL_δ for given ARL_0. For sufficiently large $T = ARL_0$, the optimum λ, denoted by λ^*, is given by

$$\lambda^* \approx \frac{2c^*\delta^2}{b - \log(b)}, \quad c^* = 0.5117, \quad b = 2\log\left[2\left(\frac{2}{\pi}\right)^{\frac{1}{2}} c^*\delta^2 T\right] \tag{2.5}$$

The optimal control limit L, denoted by L^*, is approximately

$$L^* \approx [b - \log(b)]^{\frac{1}{2}} - \lambda^*. \tag{2.6}$$

The corresponding minimum ARL_δ, denoted by ARL_1^* is approximately given by

$$ARL_1^* \approx \frac{1}{\delta^2}\left(1.2277L^{*2} - 2.835 + 9.740L^{*-2}\right) + \frac{1}{2}(1 - \lambda^*) \tag{2.7}$$

In Table 2, we give some comparisons with the optimal values given by L-S.

Shift (δ)	ARL_0 =	100	300	500	1000	2000	5000
	λ^* =	0.06 (0.06)	0.04 (0.05)	0.04 (0.05)	0.03 (0.04)	0.03 (0.03)	0.02 (0.03)
0.5	L^* =	1.998 (1.985)	2.432 (2.431)	2.614 (2.616)	2.845 (2.817)	3.061 (3.029)	3.328 (3.299)
	ARL_1^* =	18.5 (17.3)	24.8 (24.9)	28.4 (28.7)	33.7 (34.3)	39.3 (40.1)	47.1 (47.7)
	λ^* =	0.15 (0.17)	0.12 (0.14)	0.11 (0.13)	0.09 (0.11)	0.08 (0.11)	0.07 (0.09)
1.0	L^* =	2.420 (2.322)	2.814 (2.715)	2.981 (2.883)	3.194 (3.086)	3.394 (3.300)	3.643 (3.538)
	ARL_1^* =	6.40 (6.97)	8.60 (9.14)	9.60 (10.2)	11.1 (11.7)	12.6 (13.2)	14.7 (15.2)
	λ^* =	0.45 (0.49)	0.37 (0.40)	0.34 (0.36)	0.31 (0.33)	0.28 (0.30)	0.25 (0.27)
2.0	L^* =	2.573 (2.532)	2.972 (2.890)	3.141 (3.046)	3.355 (3.247)	3.556 (3.439)	3.805 (3.682)
	ARL_1^* =	2.00 (2.62)	2.60 (3.23)	2.90 (3.51)	3.30 (3.90)	3.70 (4.29)	4.30 (4.81)
	λ^* =	0.87 (0.79)	0.73 (0.72)	0.68 (0.68)	0.62 (0.62)	0.57 (0.57)	0.51 (0.50)
3.0	L^* =	2.391 (2.571)	2.831 (2.931)	3.014 (3.085)	3.246 (3.285)	3.462 (3.475)	3.728 (3.713)
	ARL_1^* =	0.70 (1.45)	1.00 (1.72)	1.20 (1.86)	1.40 (2.06)	1.60 (2.26)	1.90 (2.51)

Table 2 : Comparison of approximations from (2.5) - (2.7) with values from L-S in brackets.

3 Average Run Lengths for the Two-Sided Symmetric CUSUM Procedure

Let $\{Z_n\}$ be a sequence of independent observations distributed normally with variance one. We shall assume without any loss of generality that the target value (the deviation from which we wish to detect) is zero. Define

$$S_0^+ = 0, \quad S_n^+ = \max\left\{S_{n-1}^+ + \left(Z_n - \frac{1}{2}\delta\right), 0\right\}$$

and

$$S_0^- = 0, \quad S_n^- = \max\left\{S_{n-1}^- + \left(-Z_n - \frac{1}{2}\delta\right), 0\right\}$$

Then in the two-sided symmetric CUSUM procedure, a signal is triggered as soon as $S_n^+ > d$ or $S_n^- > d$. That is, if

$$\tau_1 = \min\left\{n \geq 1 : S_n^+ > d\right\},$$

and

$$\tau_2 = \min\left\{n \geq 1 : S_n^- > d\right\},$$

then an alarm is sounded at $\tau = \min\{\tau_1, \tau_2\}$. It should be mentioned that in this paper, we shall call δ the 'reference value'. It will be assumed that δ is known and is chosen by the experimenter. Usually, it will be the amount of shift in the mean beyond which the product will not be acceptable.

After adjusting for the discrete time model as in Siegmund (1985), we obtain

$$ARL_\mu = \left[\frac{2(\mu - \delta/2)^2}{e^{-2(\mu-\delta/2)d^*} - 1 + 2(\mu - \delta/2)d^*} + \frac{2(\mu + \delta/2)^2}{e^{2(\mu+\delta/2)d^*} - 1 - 2(\mu + \delta/2)d^*}\right]^{-1}, \quad (3.1)$$

where $d^* = d + 2\rho$, $\rho = 0.583$, and μ is the amount of shift in the mean that takes place on the first observation. When $\mu = 0$, we have

$$ARL_0 = \frac{e^{\delta d^*} - 1 - \delta d^*}{\delta^2}, \quad (3.2)$$

and when $\mu = \delta$,

$$ARL_\delta = \left(\frac{\delta^2/2}{e^{-\delta d^*} - 1 + \delta d^*} + \frac{(3\delta)^2/2}{e^{3\delta d^*} - 1 - 3\delta d^*}\right)^{-1}. \quad (3.3)$$

The following table compares the values of ARL_μ for $\delta = 1.00$ with the corresponding values given in Lucas (1976). From the results of this table, it is clear that the above formulae provide very accurate approximations for the ARL's of the two-sided symmetric CUSUM procedure. Thus, there is no need to resort to numerical methods in this case either.

μ	*Lucas* ($d = 5$)	*Approx.* ($d^* = 6.17$)	*Lucas* ($d = 6$)	*Approx.* ($d^* = 7.17$)
0	465	469	1280	1286
0.5	38.0	38.0	51.3	51.3
1.0	10.4	10.3	12.4	12.3
1.5	5.75	5.67	6.74	6.67
2.0	4.01	3.89	4.61	4.56
2.5	3.11	2.96	3.62	3.46

Table 3 : Comparison of exact ARL_μ with the approximation (3.1), with $\delta = 1.00$.

4 Comparison of EWMA with Two-Sided CUSUM Procedure

To compare the two procedures, we first consider the case in which the true shift value is equal to the reference value. Table 4 gives some results for ARL_δ^*'s with $T = 100$ and 500 respectively. From the table, we find that the EWMA procedure appears as good as the CUSUM procedure.

		EWMA			CUSUM	
T	δ	L^*	λ^*	$ARL_\delta(L^*, \lambda^*)$	d^*	$ARL_\delta(d^*)$
	0.5	1.998	0.06	15.3	6.761	19.3
	1.0	2.420	0.15	6.13	4.660	7.34
100	1.5	2.558	0.28	3.16	3.630	3.95
	2.0	2.573	0.45	1.86	3.004	2.51
	2.5	2.511	0.64	1.19	2.580	1.74
	0.5	2.614	0.04	27.1	9.748	31.1
	1.0	2.981	0.11	9.42	6.229	10.5
500	1.5	3.113	0.21	4.71	4.688	5.36
	2.0	3.141	0.34	2.76	3.803	3.30
	2.5	3.102	0.49	1.78	3.220	2.26

Table 4: Comparison of ARL_δ with known shift δ.

When the true shift μ is not equal to the reference value δ, the situation is more complicated. For simplicity, we consider only the two most common cases, i.e. $\delta = 0.5$ and $\delta = 1.0$. In Table 5, we give the values of ARL_μ for several typical values of μ. The parameter λ and the control limit L are taken from Table 4 for the corresponding δ's. We see that even in this case, the two procedures are equally good.

T	μ	$\delta = 0.5$		$\delta = 1.0$	
		EWMA	CUSUM	EWMA	CUSUM
	0.5	15.3	19.3	15.5	21.6
	1.0	6.78	8.13	6.13	7.39
100	1.5	4.30	5.09	3.68	4.16
	2.0	3.33	3.70	2.61	2.88
	2.5	2.53	2.91	2.02	2.21
	0.5	27.1	31.1	28.1	38.8
	1.0	11.2	12.1	9.42	10.5
500	1.5	6.18	7.48	5.46	5.73
	2.0	5.21	5.41	3.83	3.93
	2.5	4.15	4.23	2.61	2.99

Table 5: Comparison of ARL_μ with μ unknown.

5 Conclusion

For the case of a one-sided shift, it has been shown by Lorden (1971) and further by Moustakides (1986) that the CUSUM procedure is the best procedure in the sense that it minimizes the supremum of the conditional average delay time, which coincidently turns out to be ARL_δ, δ known. A study by Srivastava and Wu (1993) shows that in the one-sided case, the EWMA procedure is very competitive if we base the comparison upon the stationary average delay time. In the two-sided case, although there is no clear cut choice, we have shown that the EWMA procedure is at least as good as the CUSUM procedure in terms of ARL_μ.

We have developed approximate formulae for the control limit L for a given ARL_0 and λ. We also provide an optimal choice of λ as well as the corresponding control limit L and minimum value of ARL_1. These approximations can also be used as the control parameters for the EWMA procedure for other delayed detection times as ARL_0 is assumed to be large. These approximations are also comparable with the results of the discrete time model given by Lucas and Saccucci (1990). Finally, we have also obtained very simple and accurate formulae for ARL_0 and ARL_μ for the two-sided CUSUM procedure.

Appendix

(1) Derivation of (2.3)

In what follows, we will make use of the following well-known result for x in the range 1 to L (see Feller, 1957, p179):

$$\frac{1-\Phi(x)}{\phi(x)} = \frac{1}{x} - \frac{1}{x^3} + \frac{3}{x^5} - \frac{15}{x^7} + O(x^{-9}), \quad \text{as } x \to \infty. \qquad (A.1)$$

From (2.2), with $ARL_0 = T$, we get

$$I = \int_0^L [\phi(z)]^{-1} [2\Phi(z) - 1]\, dz = 2\lambda T.$$

Since $2\Phi(x) - 1 = 1 - 2[1 - \Phi(x)]$, by using Taylor's expansion in the range 0 to 1 and the above expansion (A.1) in the range (1,L), we get

$$\begin{aligned} I &= \int_0^L [\phi(x)]^{-1}\, dx - 2\int_0^L [\phi(x)]^{-1}[1-\Phi(x)]\, dx \\ &\approx \int_1^L [\phi(x)]^{-1}\, dx - 2\int_1^L \left(\frac{1}{x} - \frac{1}{x^3}\right) dx + 1 \\ &= \int_1^L [\phi(x)]^{-1}\, dx - 2\log(L) - L^{-2} + 2 \end{aligned}$$

Thus,

$$\frac{1}{2}L^2 + \log(L-1) \approx \log\left(2\lambda T/\sqrt{2\pi}\right) = \log\left[\left(\frac{2}{\pi}\right)^{\frac{1}{2}} \lambda T\right] \equiv \frac{1}{2}a.$$

Hence,

$$L \approx [a - \log(a-1)]^{\frac{1}{2}}, \quad a = 2\log\left[(2/\pi)^{\frac{1}{2}} \lambda T\right].$$

We have added an empirically chosen correction term $\frac{1}{2}(1-\lambda)$ which appears to correct for the overshoot from the boundary of the continuous model, since in the discrete time model, the process will not lie on the boundary but above it.

(2) Derivation of optimal λ

The optimization problem is to minimize ARL_1, evaluated for a specific reference value δ, subject to $ARL_0 = T$, a specified constant. Let $T_0 = \delta^2 T$ and $\lambda_0 = \frac{\lambda}{\delta^2}$. Following as in Srivastava and Wu (1993), it can be shown that

$$0 < \lambda_0 L^2/2 = c < 1,$$

for some positive c. In the following, we shall repeatedly use (A.1). For the first term of ARL_δ we note that

$$\int_{-(2/\lambda_0)^{1/2}}^{L-(2/\lambda_0)^{1/2}} [\phi(z)]^{-1}\,\Phi(z)\, dz = \int_{-(2/\lambda_0)^{1/2}}^{L-(2/\lambda_0)^{1/2}} \left[-\frac{1}{z} + O\left(z^{-3}\right)\right] dz$$

$$= \log\left[\frac{(2/\lambda_0)^{1/2}}{(2/\lambda_0)^{1/2}-L}\right] = -\log\left[1-L(2/\lambda_0)^{1/2}\right].$$

Thus, the first term of ARL_δ in (2.4) is approximately

$$-\frac{1}{\delta^2\lambda_0}\log[1-L(\lambda_0/2)^{1/2}] = -\frac{L^2}{2c\delta^2}\log[1-\sqrt{c}]$$

The second term can be checked to be $o(1)$. Therefore, the optimal c^* which minimizes ARL_1 is approximately 0.5117. Thus, the optimal weight is about

$$\lambda^* = \frac{2c^*\delta^2}{L^{*2}},$$

where L^* is obtained from

$$\int_0^{L^*}[\phi(x)]^{-1}[2\Phi(x)-1]dx = \frac{4c^*T_0}{L^{*2}}.$$

By using the results of (A.1), we can obtain L^*. Since the optimum value of c is obtained by taking only one term of the expansion (A.1), it tends to give somewhat larger values for the optimum value of L. To correct for this, we have subtracted λ^* from it to keep the formula simple. In fact λ^* and L^* should be obtained by solving the two equations simultaneously.

In order to get the approximation for ARL_δ, we again use (A.1) and find that ARL_δ becomes

$$\frac{L^2}{2c\delta^2}\int_{L(1/\sqrt{c}-1)}^{L\sqrt{c}}\left[\frac{1}{z}-\frac{1}{z^3}+\frac{3}{z^5}+O(z^{-7})\right]dz$$
$$= \frac{1}{2c\delta^2}\left[-L^2\log(1-\sqrt{c})-\frac{c}{2(1-\sqrt{c})^2}+\frac{c}{2}+\frac{3c^2}{4L^2}\left(\frac{1}{(1-\sqrt{c})^4}-1\right)\right]+O(L^{-3}).$$

The second term goes to zero exponentially fast. Thus, by substituting L^* and c^* into the above equation, we obtain

$$ARL_1^* = \frac{1}{\delta^2}\left[1.2277L^{*2} - 2.835 + \frac{29.221}{L^{*2}}\right] + O(L^{*-3})$$

In this approximation, ARL_1^* is underestimated if we take only the first two terms and overestimated if we include the third term. As an alternative to taking many more terms for greater accuracy (as the series does not converge fast), we have empirically taken only $\frac{1}{3}$ of the last term.

References

1. Crowder,S.V. (1987) "A simple method for studying run length distributions of exponentially weighted moving average control charts." *Technometrics*, 29, 401-407.

2. Feller,W. (1957) *An Introduction to Probability Theory and Its Applications.* Wiley, New York.

3. Hunter,J.S. (1986) The exponentially weighted moving average." *J. Quality Technology.* 18, 203-210.

4. Karlin,S. and Taylor,H.M. (1981) *A Second Course In Stochastic Processes.* Academic Press.

5. Lorden,G. (1971) "Procedures for reacting to a change in distribution." *Ann. Math. Statist.*. 42, 1897-1908.

6. Lucas,J.M. (1976) "The design and use of V-mask control schemes" *J. Qual. Tech.* 8,1-12.

7. Lucas,J.M. and Saccucci,M.S. (1987) "Exponentially weighted moving average control schemes: Properties and Enhancements," *Faculty Working Series Paper 87-5, Drexel University Dept. of Quantitative Methods.*

8. Lucas,J.M. and Saccucci,M.S. (1990) "Exponentially weighted moving average control chart schemes: properties and enhancements." *Technometrics*, 32, 1-30.

9. Montgomery,D.C., Gardiner,J.S. and Pizzaro,B.A. (1987) "Statistical process control methods for detecting small process shifts," in *Frontiers in Statistical Quality Control*, eds. H.-J. Lenz, G.B. Wetherill and P.T. Wilrich, Heidelberg, West Germany: Physica-Verlag, pp. 161-178.

10. Moustakides,G.V. (1986) "Optimal stopping times for detecting change in distribution." *Ann. Statist.* 14, 1379-1387.

11. Robinson,P.B. and Ho,T.Y. (1978) "Average run lengths of geometric moving average charts by numerical methods," Technometrics, 20, 85-93.

12. Roberts,S.W. (1959) "Control chart tests based on geometric moving average." *Technometrics*, 1, 239-250.

13. Siegmund, D. (1985) *Sequential Analysis.* Springer, Berlin.

14. Srivastava,M.S. and Wu,Yanhong (1993) "Comparison of EWMA, CUSUM and Shiryayev-Roberts procedures for detecting a shift in the mean." *Ann. Statist.* 21, 645-670.

15. Van Dobben de Bruyn,C.S. (1968) *Cumulative Sum Tests.* Griffin, London.

16. Waldmann,K.H. (1986) "Bounds for the distribution of the run length of geometric moving average charts," *Applied Statistics*, 35, 151-158.