The Expected Average Run Length of the EWMA Median Chart with Estimated Process Parameters

You Huay Woon
Universiti Kebangsaan Malaysia

Michael Khoo Boon Chong
Universiti Sains Malaysia

Chong Zhi Lin
Universiti Tunku Abdul Rahman

Teoh Wei Lin
Heriot-Watt University Malaysia

Abstract

The performance of a control chart is commonly investigated based on the assumption of known process parameters. Nevertheless, in most manufacturing and service applications, the process parameters are usually unknown to practitioners. Hence, they are estimated from an in-control Phase-I samples. As such, the performance of the control chart with estimated process parameters will behave differently from the corresponding chart with known process parameters. To study this issue, the exponentially weighted moving average (EWMA) median chart is examined in this article. The EWMA median chart is traditionally investigated based on the average run length (ARL). The limitation of the ARL is that it requires practitioners to specify the shift size in advance. This phenomenon is not ideal for practitioners who do not have background knowledge of the process. In view of this, the EWMA median chart with known and estimated process parameters is studied based on the ARL and expected average run length (EARL). The results indicate that as long as the particular shift size is within the range of shifts, the performance of the chart is almost the same, for the EWMA median chart with known and estimated process parameters.

Keywords: estimated process parameters, EWMA median chart, expected average run length.

1. Introduction

Statistical Process Control (SPC) is a popular statistical approach used to monitor and enhance the process quality. This is essential so that the final product meets consumersâĂŹ satisfaction. The control chart, which is the most effective tool in SPC, which is employed to reduce variability and to monitor the performance in manufacturing and service processes(Steland 2006).

The usual assumption in designing control chart is that the process parameters are assumed known. However, this assumption is impractical in actual application. Many times, the application of a control chart requires the estimation of the process parameters from a limited number of samples (Jensen, Jones-Farmer, Champ, and Woodall 2006; Woodall and Montgomery 2014; Saleh, Mahmoud, Jones-Farmer, Zwetsloot, and Woodall 2015). Due to the variability of the estimators, the control chart with estimated process parameters will lead to

deterioration in the performance of the chart when compared to the corresponding chart with known process parameters (Psarakis, Vyniou, and Castagliola 2014). Readers can refer to Zhang, Castagliola, Wu, and Khoo (2012), Saleh *et al.* (2015), Chong, Khoo, Teoh, You, and Castagliola (2018), to name a few, for the study on the control chart with estimated process parameters.

The exponentially weighted moving average (EWMA) chart was proposed by Roberts (1959). The EWMA chart performs well in the detection of small-to-moderate shifts in a process under normality assumption. However, Human, Kritzinger, and Chakraborti (2011) showed that this generally does not hold when the process contains contaminated normal data. An ideal alternative in such a case is the EWMA median chart. The EWMA median chart was proposed by Castagliola (2001). The main advantage of the EWMA median chart is that it is robust against outliers.

Castagliola (2001) designed the EWMA median chart based on the average run length (ARL). The assumption for the computation of the ARL is that the magnitude of process shift size is known. However, in real practice, user of the chart may not know the particular shift size beforehand (Celano 2009). From this point of view, the shift size should be considered as a range of shifts. Therefore, the expected average run length (EARL) is necessary as an alternative performance measure. Castagliola, Celano, and Psarakis (2011) and You (2017) studied the performance of control charts when the shift size is unknown. In this study, the EWMA median chart with known and estimated process parameters is evaluated, with respect to the ARL and EARL.

The rest of this article is organized as follows: The properties of the EWMA median chart with known and estimated process parameters are given in Section 2. Section 3 presents the performance comparison in terms of the ARL and EARL when the process parameters are known and estimated. Finally, some conclusions are drawn in the last section.

2. Run length properties of the EWMA median chart

The EWMA median chart monitors the process using the sample median, i.e.

$$\widetilde{Y}_i = \begin{cases} Y_{i,((n+1)/2)} & \text{if } n \text{ is odd} \\ \\ \frac{Y_{i,(n/2)}+Y_{i,((n/2)+1)}}{2} & \text{if } n \text{ is even} \end{cases} \tag{1}$$

When dealing with the sample median, it is typical to assume that the sample size, n is an odd number. This makes the sample median simpler and easier to be calculated.

The statistic for EWMA median chart is

$$Z_i = (1-\lambda)Z_{i-1} + \lambda\widetilde{Y}_i, \text{for } i\text{=1,2,...,} \tag{2}$$

where $\lambda(0 < \lambda \leq 1)$ is a smoothing constant and $Z_0 = \mu_0$. The lower control limit (LCL) and upper control limit (UCL) of the EWMA median chart are

$$LCL/UCL = \mu_0 \mp K\sigma_0, \tag{3}$$

where K is the charting constant.

Here, the Markov chain method can be used to assess the run length properties of the EWMA median chart. Let

$$\mathbf{Q} = [Q_{i,j}]_{l\times l} \tag{4}$$

be the transition probability matrix (tpm) for the transient states, where $i, j = 1,2,...l$. Note that $l = 2g+1$ subintervals, each of width $2d$, where $d = (UCL\text{-}LCL)/(2l)$. Let H_j, for $j = 1,2,...,2g+1$, is the midpoint of the jth subinterval. The $Q_{i,j}$ of the matrix $\mathbf{Q}$ are

$$Q_{i,j} = F_{\widetilde{Y}_i}\left(\frac{H_j + d - (1-\lambda)H_i}{\lambda}|n\right) + F_{\widetilde{Y}_i}\left(\frac{H_j - d - (1-\lambda)H_i}{\lambda}|n\right), \tag{5}$$

for $i, j = 1, 2, ..., 2g+1$. Note that $F_{\widetilde{Y}_i}(...|n)$ is the cumulative distribution function (cdf) of the sample median $\widetilde{Y}_i$, with $i = 1, 2, ...$.

According to Castagliola (2001), the formula to compute the ARL of the EWMA median chart is

$$\text{ARL} = v_1, \tag{6}$$

with

$$v_1 = \mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}. \tag{7}$$

Note that $\mathbf{q}$ is the vector of initial probabilities and $\mathbf{I}$ is the identity matrix. To compute the ARL, the chartâĂŹs user is required to determine the process shift size. This is too restrictive as the user of the chart may not have the background knowledge of the process to specify the process shift size. As such, it is crucial to consider the EARL for an overall range of shifts $(\delta_{min}, \delta_{max})$, where δ_{min} and δ_{max} represent the lower and upper bounds of the mean shift, respectively. The EARL of the EWMA median chart is

$$\text{EARL} = \int_{\delta_{min}}^{\delta_{max}} f_\delta(\delta)\text{ARL}d\delta. \tag{8}$$

Here, $f_\delta(\delta)$ is the probability density function of the process shift, i.e. δ.

When the process parameters, i.e. μ_0 and σ_0 are unknown, they have to be estimated from an in-control Phase-I samples, which consists of $i = 1, 2, \text{âĂę}, m$ samples, each with size n. Then, the estimator of μ_0 and σ_0 are

$$\hat{\mu}_0 = \frac{1}{m}\sum_{i=1}^{m}\widetilde{X}_i \tag{9}$$

and

$$\hat{\sigma}_0 = \frac{1}{d_{2,n}}\left(\frac{1}{m}\sum_{i=1}^{m}R_i\right) \tag{10}$$

respectively. Here, $\widetilde{X}_i$ and R_i are the sample median and sample range, respectively and $d_{2,n} = E(R_i/\sigma_0)$.

3. Performance analysis of the EWMA median chart

The performance of the EWMA median chart with known and estimated process parameters is evaluated based on the ARL and EARL criteria. Table 1 presents the optimal charting parameters (λ, K) and the corresponding out-of-control ARL (ARL_1) for $n = 3, 5, 7, 9$ with different combinations of (m, δ), when the in-control ARL (ARL_0) is 370.4. For comparison purposes, the case for $m = +\infty$, which represents the known process parameters based chart, is included in Tables 1 and 2.

For illustration, when $n = 3$ and $\delta = 0.4$, the optimal charting parameters $(\lambda, K) = (0.1000, 0.4156)$ give the lowest ARL_1 value, i.e. $\text{ARL}_1 = 21.12$, while attaining the desired $\text{ARL}_0 = 370.4$, when $m = +\infty$. Using the similar optimal charting parameters corresponding to known process parameters case, $\text{ARL}_1 = 33.56$ is obtained when $m = 25$.

From Table 1, it signifies that parameter estimation significantly affects the performance of the EWMA median chart, especially when m, n and δ are small. Nevertheless, the effects of parameter estimations become smaller and the ARL_1 values approach the case of known process parameters when m increases. For example, when $m = 50$, $n = 5$ and $\delta = 0.2$, $\text{ARL}_1 = 68.17$ and it decreases to 50.87 when $m = 200$. This indicates that the corresponding ARL_1 when $m = 200$ is closer to that of the known process parameters case, i.e. $\text{ARL}_1 = 46.50$ when $m = +\infty$.

Nevertheless, the use of ARL as the performance measure requires practitioners to determine the exact process mean shift. In practice, practitioners do not have the necessary experience to specify the process mean shift. To circumvent this problem, the EARL is proposed to evaluate the EWMA median chart when the shift size is unknown. The EARL performance measure takes into account a range of shifts. Here, the shift intervals $(\delta_{min}, \delta_{max}) = (0.2, 1.0)$ and $(\delta_{min}, \delta_{max}) = (1.0, 2.0)$ are considered so that it includes the exact shift size in Table 1. For instance, for $(\delta_{min}, \delta_{max}) = (0.2, 1.0)$, it includes $\delta = \{0.2, 0.4, 0.8\}$ which is considered in Table 1. Similarly for $(\delta_{min}, \delta_{max}) = (1.0, 2.0)$, it includes $\delta = \{1.2, 1.5, 2.0\}$.

Table 1: The ARL_1 values of the EWMA median chart based on the optimal charting parameters (λ, K) corresponding to the known process parameters case when $n = 3, 5, 7, 9$ and $ARL_0 = 370.4$ with different combinations of (m, δ).

n	δ	(λ, K)	$m=25$	$m=50$	$m=80$	$m=100$	$m=200$	$m = +\infty$
3	0.2	(0.1000, 0.4156)	137.19	106.31	91.91	86.78	76.58	67.63
	0.4	(0.1000, 0.4156)	33.56	25.34	23.41	22.87	21.93	21.12
	0.8	(0.1808, 0.6007)	8.06	7.66	7.53	7.49	7.42	7.34
	1.2	(0.3202, 0.8607)	4.17	4.07	4.03	4.02	4.00	3.97
	1.5	(0.4456, 1.0696)	2.95	2.89	2.87	2.87	2.85	2.84
	2.0	(0.6833, 1.4546)	1.88	1.84	1.83	1.83	1.82	1.81
5	0.2	(0.1000, 0.3323)	88.96	68.17	59.11	56.19	50.87	46.50
	0.4	(0.1000, 0.3323)	18.88	16.36	15.72	15.53	15.18	14.86
	0.8	(0.2489, 0.5863)	5.52	5.37	5.31	5.29	5.26	5.23
	1.2	(0.4468, 0.8564)	2.92	2.88	2.86	2.86	2.85	2.84
	1.5	(0.6293, 1.0918)	2.06	2.03	2.02	2.02	2.01	2.01
	2.0	(0.8520, 1.3941)	1.32	1.30	1.30	1.30	1.29	1.29
7	0.2	(0.1000, 0.2846)	65.14	48.99	43.07	41.31	38.22	35.70
	0.4	(0.1146, 0.3098)	13.79	12.60	12.27	12.16	11.97	11.79
	0.8	(0.3096, 0.5765)	4.30	4.21	4.18	4.17	4.15	4.13
	1.2	(0.5719, 0.8709)	2.28	2.26	2.25	2.25	2.24	2.23
	1.5	(0.7587, 1.0813)	1.60	1.58	1.57	1.57	1.57	1.57
	2.0	(0.9362, 1.2996)	1.11	1.10	1.10	1.10	1.10	1.10
9	0.2	(0.1000, 0.2528)	50.12	37.80	33.89	32.77	30.82	29.20
	0.4	(0.1365, 0.3069)	11.12	10.40	10.19	10.12	9.99	9.87
	0.8	(0.3673, 0.5717)	3.57	3.51	3.49	3.48	3.47	3.46
	1.2	(0.6749, 0.8748)	1.88	1.87	1.86	1.86	1.85	1.85
	1.5	(0.8443, 1.0503)	1.33	1.32	1.32	1.32	1.32	1.31
	2.0	(0.9761, 1.2006)	1.04	1.03	1.03	1.03	1.03	1.03

The optimal charting parameters (λ, K) and the corresponding out-of-control EARL $(EARL_1)$ based on the minimization of the $EARL_1$ with different combinations of $(m, n, \delta_{min}, \delta_{max})$ when the in-control EARL $(EARL_0)$ is 370.4, are displayed in Table 2. All the $EARL_1$ values for different m are calculated from the optimal charting parameters (λ, K) presented in Table 2. For instance, when $n = 7$, $\delta_{min} = 0.2$ and $\delta_{max} = 1.0$, the optimal charting parameters are $(\lambda, K) = (0.1057, 0.2945)$ corresponding to the known process parameters case $(m = +\infty)$. With these optimal charting parameters, the $EARL_1$ for $m = 25$ is 12.56, while $EARL_1 = 9.72$ when $m = +\infty$. The difference of the values for estimated process parameters case is again quite large compared to the known process parameters case.

The proposed EARL as the performance measure when the process shift size is unknown, is crucial. Taking $n = 5$, $\delta_{min} = 0.2$ and $\delta_{max} = 1.0$ in Table 2 as an example, the optimal charting parameters of the EWMA median chart are $(\lambda, K) = (0.1000, 0.3323)$ and the corresponding $EARL_1 = 12.19$ when $m = +\infty$. By considering $\delta = 0.8$ $\left(\text{i.e.} \delta \in (\delta_{\min}, \delta_{\max})\right)$, the $ARL_1 = 5.84$ is obtained by using the similar optimal charting parameters, $(\lambda, K) = (0.1000, 0.3323)$. For similar n and δ values, the $ARL_1 = 5.23$ is obtained using $(\lambda, K) = (0.2489, 0.5863)$ from Table 1. This shows that the performance of the EWMA median chart is almost

the same when the optimal charting parameters are obtained based on the minimization of ARL_1 and $EARL_1$.

For the estimated process parameters case, similarly, using the same n, δ_{min} and δ_{max} combination as mentioned above (i.e., $n = 5$, $\delta_{min} = 0.2$ and $\delta_{max} = 1.0$), by considering $\delta = 0.8$ $\left(\text{i.e.} \delta \in \left(\delta_{\min}, \delta_{\max} \right) \right)$, the $ARL_1 = 5.93$ is obtained when $m = 50$ by using $(\lambda, K) = (0.1000, 0.3323)$. For fixed m, n and δ values, the $ARL_1 = 5.37$ is yielded using $(\lambda, K) = (0.2489, 0.5863)$ from Table 1. This suggests that the optimal charting parameters (λ, K) obtained by minimizing $EARL_1$, is reliable to employ as long as $\delta \in \left(\delta_{\min}, \delta_{\max} \right)$.

Table 2: The $EARL_1$ values of the EMWA median chart based on the optimal charting parameters (λ, K) corresponding to the known process parameters case when $n = \{3, 5, 7, 9\}$ and $EARL_0 = 370.4$ with different combinations of $(m, \delta_{min}, \delta_{max})$.

n	δ_{min}	δ_{max}	(λ, K)	m=25	m=50	m=80	m=100	m=200	$m = +\infty$
3	0.2	1.0	(0.1000, 0.4156)	28.00	21.84	19.82	19.21	18.09	17.17
	1.0	2.0	(0.3820, 0.9652)	3.27	3.19	3.17	3.16	3.14	3.13
5	0.2	1.0	(0.1000, 0.3323)	17.08	14.24	13.33	13.07	12.59	12.19
	1.0	2.0	(0.5340, 0.9689)	2.29	2.26	2.24	2.24	2.23	2.22
7	0.2	1.0	(0.1057, 0.2945)	12.56	10.86	10.36	10.21	9.95	9.72
	1.0	2.0	(0.6394, 0.9457)	1.79	1.77	1.76	1.76	1.75	1.75
9	0.2	1.0	(0.1242, 0.2894)	10.09	8.94	8.61	8.51	8.34	8.18
	1.0	2.0	(0.7110, 0.9110)	1.50	1.49	1.48	1.48	1.48	1.48

4. Conclusion

In conclusion, this article manifests that the EARL is an ideal alternative performance measure when the process shift size cannot be determined in advance for the EWMA median chart with known and estimated process parameters. In such a case, practitioners can use the proposed optimal charting parameters based on the minimization of $EARL_1$ as long as the considered process shift size is within $\left(\delta_{min}, \delta_{max} \right)$. This can alleviate the inaccuracy following the misuse of the optimal charting parameters, when practitioners have no experience in specifying the process shift size. By looking at the large number of in-control Phase-I samples required, future research can consider to develop the optimal charting parameters for the EWMA median chart with estimated process parameters, based on the EARL.

Acknowledgement

References

Castagliola P (2001). "An $\bar{X}/R$-EWMA Control Chart for Monitoring the Process Sample Median." *International Journal of Reliability, Quality and Safety Engineering*, **8**(2), 123–135.

Castagliola P, Celano G, Psarakis S (2011). "Monitoring the Coefficient of Variation using EWMA Charts." *Journal of Quality Technology*, **43**(3), 249–265. doi:https://doi.org/10.1080/00224065.2011.11917861.

Celano G (2009). "Robust Design of Adaptive Control Charts for Manual Manufacturing/Inspection Workstations." *Journal of Applied Statistics*, **36**(2), 181–203. `doi:https://doi.org/10.1080/02664760802443947`.

Chong ZL, Khoo MBC, Teoh WL, You HW, Castagliola P (2018). "Optimal Design of the Side-sensitive Modified Group Runs (SSMGR) $\bar{X}$ Chart when Process Parameters Are Estimated." *Quality and Reliability Engineering International*, **35**(1), 246–262. `doi:https://doi.org/10.1002/qre.2395`.

Human S, Kritzinger P, Chakraborti S (2011). "Robustness of the EWMA Control Chart for Individual Observations." *Journal of Applied Statistics*, **38**(10), 2071–2087. `doi:10.1080/02664763.2010.545114`.

Jensen W, Jones-Farmer L, Champ C, Woodall W (2006). "Effects of Parameter Estimation on Control Chart Properties: A Literature Review." *Journal of Quality Technology*, **38**(4), 349–364. `doi:10.1080/00224065.2006.11918623`.

Psarakis S, Vyniou A, Castagliola P (2014). "Some Recent Developments on the Effects of Parameter Estimation on Control Charts." *Quality and Reliability Engineering International*, **30**(8), 1113–1129. `doi:https://doi.org/10.1002/qre.1556`.

Roberts S (1959). "Control Chart Tests Based on Geometric Moving Averages." *Technometrics*, **1**(3), 239–250. `doi:10.1080/00401706.1959.10489860`.

Saleh NA, Mahmoud MA, Jones-Farmer LA, Zwetsloot I, Woodall WH (2015). "Another Look at the EWMA Control Chart with Estimated Parameters." *Journal of Quality Technology*, **47**(4), 363–382. `doi:10.1080/00224065.2015.11918140`.

Steland A (2006). "A Bootstrap View on Dickey-Fuller Control Charts for AR(1) Series." *Austrian Journal of Statistics*, **35**(2 & 3), 339–346. `doi:https://doi.org/10.17713/ajs.v35i2&3.381`.

Woodall WH, Montgomery DC (2014). "Some Current Directions in the Theory and Application of Statistical Process Monitoring." *Journal of Quality Technology*, **46**(1), 78–94. `doi:10.1080/00224065.2014.11917955`.

You HW (2017). "Run Length Distribution of Synthetic Double Sampling Chart." *International Journal of Applied Engineering Research*, **12**(24), 14268–14272. `doi:https://doi.org/10.1155/2018/7583610`.

Zhang Y, Castagliola P, Wu Z, Khoo MBC (2012). "The Variable Sampling Interval $\bar{X}$ Chart with Estimated Parameters." *Quality and Reliability Engineering International*, **28**(1), 19–34. `doi:DOI:10.1002/qre.120`.

Affiliation:

You Huay Woon
Pusat GENIUS@Pintar Negara
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor, Malaysia
E-mail: `hwyou@ukm.edu.my`