Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol III,
IMECS 2010, March 17 - 19, 2010, Hong Kong

A Novel CUSUM Median Control Chart

Ling Yang*, Suzanne Pai, and Yuh-Rau Wang

***Abstract*— This work presents a novel Cumulative Sum (CUSUM) median control chart. Based on experience, some processes occasionally have outliers. The mean control charts are sensitive to outliers, and the median control charts are outliers-resistant. In this paper, several mean control charts, the proposed CUSUM median control chart, and two median control charts are used for comparison. With various shifts of the process sample mean, the average run lengths of control charts are evaluated under some contaminated normal distributions. The simulation result reveals that the outlier-resistance of the CUSUM median chart performs best, and the shifts-detecting ability of the CUSUM median chart is similar to those of the EWMA median chart and the mean charts.**

***Index Terms*— control chart, cumulative sum, median, outliers.**

I. Introduction

Based on experience, some processes occasionally have outliers. In monitoring the process mean, the mean ($\bar{X}$) control charts, namely the Shewhart- $\bar{X}$ chart, cumulative sum (CUSUM)- $\bar{X}$ chart [8], exponentially weighted moving average (EWMA)- $\bar{X}$ chart [3], and generally weighted moving average (GWMA)- $\bar{X}$ chart [13], have been investigated extensively. Because the sample average, $\bar{X}$, is sensitive to outliers, using the $\bar{X}$ chart for monitoring the process mean will lead to high level false alarms. The sample median, $\tilde{X}$, is a robust estimator of location for samples. Recently, there are some $\tilde{X}$ control charts, such as the EWMA- $\tilde{X}$ chart [2], Shewhart- $\tilde{X}$ chart [6], and GWMA- $\tilde{X}$ chart [14], were developed in succession.

In general, the desirable properties expected from control charts are fast detection of assignable causes, robustness to underlying assumptions and economical usefulness. Samples containing outliers are said to be contaminated. In [14], under various contaminated normal distributions, the average run lengths (*ARL*s) of the Shewhart- , EWMA-, and GWMA-, $\bar{X}$ and $\tilde{X}$ charts were evaluated by simulation. Their study revealed that the in-control *ARL*s (ARL_0s) of $\tilde{X}$ charts are significantly higher than that of the $\bar{X}$ charts, and the out-of-control *ARL*s (ARL_1s) of $\tilde{X}$ charts are smaller than that of the $\bar{X}$ charts. It means that the CUSUM- $\tilde{X}$ chart is outliers-resistant and the $\bar{X}$ charts can detect the process shifts faster.

The CUSUM control charts were first proposed by [10] and had been studied by many authors: such as [11], [4], [7], [5] and [15], etc. The CUSUM control charts had been demonstrated that can detect the small shifts of the process quickly when the process has no outlier. However, to the best of our knowledge, the CUSUM technique is still not applied in the design of $\tilde{X}$ control chart in the literature.

In this paper, a CUSUM- $\tilde{X}$ control chart is developed for monitoring the process sample mean with outliers. We assume that the process characteristic follows the normal distribution. To compare the statistical usefulness of the CUSUM-, Shewhart-, EWMA-, $\tilde{X}$ and $\bar{X}$ charts, the simulation [12] is used to evaluate the *ARL*s of various process sample mean shifts under various contaminated normal distributions. Finally, some conclusions are included in the last section.

II. Descriptions of some control charts

Suppose that the quality characteristic is a variable and samples were collected at each point in time (the size of rational subgroups is *n*). Let $\bar{X}_i$ and $\tilde{X}_i$ be sample average and sample median of *i*th subgroup respectively, which are composed of *n* independent normal (μ_i, σ^2) random variables $X_{i,1}, \cdots, X_{i,n}$, where μ_i is the process mean (and the process median), and σ^2 is the common process variance. That is,

$$\bar{X}_i = \sum_{j=1}^{n} X_{i,j} / n ,$$

$$\tilde{X}_i = X_{[i,(n+1)/2]}, \text{(if } n \text{ is odd)}$$

where $X_{[i,j]}$ is the *j*th order statistic for the *i*th sample. When the process is in control, let $\mu_i = \mu_0$ (the target value of process mean).

A. The CUSUM- $\bar{X}$ control chart

In Montgomery [8], the plotted statistics of the tabular

Manuscript received December 9, 2009. This work was supported in part by National Science Council of Taiwan under the contract number NSC 98-2221-E-129-007.

Ling Yang is with the Department of Marketing and Logistics Management, St. John’s University, Tamsui, Taipei 251, Taiwan, ROC. (corresponding author, phone: +886-2-28013131 ext. 6050, e-mail: lgyang@mail.sju.edu.tw).

Suzanne Pai is with the Department of Marketing and Logistics Management, St. John’s University, Tamsui, Taipei 251, Taiwan, ROC. (e-mail: sjpai@mail.sju.edu.tw).

Yuh-Rau Wang is with the Department of Computer Science and Information Engineering, St. John’s University, Tamsui, Taipei 251, Taiwan, ROC. (e-mail: yrwang@mail.sju.edu.tw)

CUSUM- $\overline{X}$ control chart, C_i^+ and C_i^-, are computed as follows:

$$C_i^+ = \max[0, \overline{X}_i - (\mu_0 + K) + C_{i-1}^+] \quad (1)$$

$$C_i^- = \max[0, (\mu_0 - K) - \overline{X}_i + C_{i-1}^-] \quad (2)$$

where the starting values are $C_0^+ = C_0^- = 0$. K is the reference value and often chosen about halfway between the target value (μ_0) and the out-of-control value of the mean (μ_1) that we are interested in detecting the process shift quickly, i.e.,

$$\begin{aligned} K &= \frac{|\mu_1 - \mu_0|}{2} = \frac{|(\mu_0 + \delta_0\sigma) - \mu_0|}{2} \\ &= \delta_0\sigma / 2 = (\frac{\delta_0\sqrt{n}}{2})\frac{\sigma}{\sqrt{n}} = k\sigma_{\overline{X}} \end{aligned} \quad (3)$$

where $\sigma_{\overline{X}}$ denotes the standard deviation of the sample mean, δ_0 denotes the magnitude of the process mean shift (multiple of σ) that we are interested, and

$$k = \delta_0\sqrt{n} / 2 \quad (4)$$

Let H be the decision interval. Define

$$H = h\sigma_{\overline{X}} = h\sigma / \sqrt{n} \quad (5)$$

If either C_i^+ or C_i^- exceeds H, the process is considered to be out of control. Once k is selected, we should choose h to achieve the desired ARL_0. For example, in [5], when n = 1, δ_0 =1/2, k = 1/2, and ARL_0 = 370, the value of h is 4.77.

B. The Shewhart- $\overline{X}$ control chart

In [8], the plotted statistic of the Shewhart- $\overline{X}$ chart is $\overline{X}_i$. The central line (CL), upper control limit (UCL) and lower control limit (LCL) of the Shewhart- $\overline{X}$ chart are represented as

$$CL = \mu_0 \quad (6)$$

$$UCL / LCL = \mu_0 \pm L\sigma / \sqrt{n} \quad (7)$$

where L determines the width of control limits. The process is considered out of control and some actions should be undertaken whenever $\overline{X}_i$ falls outside the range of control limits.

C. The EWMA- $\overline{X}$ control chart

In [3], the EWMA- $\overline{X}$ control statistic, Y_i, can be represented as

$$Y_i = \alpha\overline{X}_i + (1-\alpha)Y_{i-1}, \quad \text{for } i = 1, 2, \cdots \quad (8)$$

where $Y_0 = \mu_0$, and the smooth parameter, α ($0 < \alpha \le 1$), is determined by the practitioner. The CL, time-varying UCL and LCL of the EWMA- $\overline{X}$ control chart are represented as

$$CL = \mu_0 \quad (9)$$

$$UCL / LCL = \mu_0 \pm L\sqrt{\frac{\alpha(1-(1-\alpha)^{2i})}{2-\alpha}}\frac{\sigma}{\sqrt{n}} \quad (10)$$

where L determines the width of control limits. When α = 1, the EWMA- $\overline{X}$ control chart reduces to the Shewhart- $\overline{X}$ control chart.

D. The EWMA- $\tilde{X}$ and Shewhart- $\tilde{X}$ control charts

The distribution of sample median ($\tilde{X}_i$), derived by [1], is very close to the $(\mu_i, \sigma_{\tilde{X}}^2)$ normal distribution, where $\sigma_{\tilde{X}}^2$ is the variance of $\tilde{X}_i$. If $\tilde{\sigma}_{0,1}$ is the standard deviation of normal (0, 1) sample median, we have $\sigma_{\tilde{X}} = \sigma \times \tilde{\sigma}_{0,1}$. For the values of $\tilde{\sigma}_{0,1}$ of various sample size n refer to Castagliola (1998) for details. In [2], the EWMA- $\tilde{X}$ control statistic, Z_i is represented as

$$Z_i = \beta\tilde{X}_i + (1-\beta)Z_{i-1}, \quad \text{for } i = 1, 2, \cdots \quad (11)$$

where $Z_0 = \mu_0$, and the smooth parameter, β ($0 < \beta \le 1$), is determined by the practitioner. The time-varying control limits of the EWMA- $\tilde{X}$ control chart are

$$CL = \mu_0 \quad (12)$$

$$UCL / LCL = \mu_0 \pm L\sqrt{\frac{\beta(1-(1-\beta)^{2i})}{2-\beta}} \times \sigma_0\tilde{\sigma}_{0,1} \quad (13)$$

where L determines the width of control limits. When $\beta = 1$, the plotted statistic in (11) will be $Z_i = \tilde{X}_i$, and the control limits are

$$CL = \mu_0 \quad (14)$$

$$UCL / LCL = \mu_0 \pm L\,\sigma_0\tilde{\sigma}_{0,1}. \quad (15)$$

The EWMA- $\tilde{X}$ control chart reduces to the Shewhart- $\tilde{X}$ control chart.

III. Description of the Tabular CUSUM- $\tilde{X}$ Control Chart

From (1) and (2), the CUSUM- $\tilde{X}$ control statistics, η_i^+ and η_i^-, can be represented as

$$\eta_i^+ = \max[0, \tilde{X}_i - (\mu_0 + K) + \eta_{i-1}^+] \quad (16)$$

$$\eta_i^- = \max[0, (\mu_0 - K) - \tilde{X}_i + \eta_{i-1}^-] \quad (17)$$

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The starting values are $\eta_0^+ = \eta_0^- = 0$. From (3), the reference value, *K*, can be presented as

$$K = \frac{|\mu_1 - \mu_0|}{2} = \frac{\delta_0 \sigma}{2} = (\frac{\delta_0}{2\tilde{\sigma}_{0,1}})(\sigma\tilde{\sigma}_{0,1}) = d\sigma_{\tilde{X}} \tag{18}$$

where

$$d = \delta_0 / (2\tilde{\sigma}_{0,1}) \tag{19}$$

Let *M* be the decision interval. From (4), we can define

$$M = m\sigma_{\tilde{X}} = m\sigma\tilde{\sigma}_{0,1} \tag{20}$$

Under the selected *d*, we should choose *m* to achieve the desired in-control ARL_0. If either η_i^+ or η_i^- exceeds *M*, the process is considered to be out of control.

Without loss of generality, throughout this paper, we assume that in the absence of a special cause of variation, $X_{i,j}$, $i = 1, 2, 3, \cdots$, $j = 1, 2, \cdots, n$, are independent and have a common normal distribution with mean μ_i (the in-control $\mu_i = \mu_0 = 0$) and variance $\sigma^2 = 1$. For simplicity, we assume *n* = 5 for the performance comparison. Then, we can get $\tilde{\sigma}_{0,1} = 0.536$ [1]. The simulation is used to estimate the values of *m*, *h* and ARL_1. Each simulation runs 100,000 iterations. Table 1 shows the simulation results of *h* and *m* of CUSUM-, $\bar{X}$ and $\tilde{X}$ charts, respectively, under various interested mean shifts $\delta_0 \in$ {0.1, 0.3, 0.5, 0.7, 1.0, 1.5}.

Table 1. Values of *h* and *m* (*n* = 5, $ARL_0 \cong 500$)

CUSUM	δ_0 = 0.1	0.3	0.5	0.7	1.0	1.5
$\bar{X}$ chart	*h*=14.271	7.158	4.749	3.495	2.439	1.522
$\tilde{X}$ chart	*m*=15.457	8.106	5.508	4.129	2.935	1.927

IV. THE ARL PERFORMANCE

In order to evaluate the *ARL* of different charts in the presence of outliers, the contaminated normal distribution used in [6] is adopted. A contaminated normal distribution is that the observations $(100 - \theta)$ % come from (*N*(0, 1)) normal distribution and θ% come from (*N*(0, σ_o^2)) normal distribution, where θ denotes the level of contamination and σ_o^2 denotes the variance of an outlier. We assumed that the outliers occur due to the common causes of variation and lead to a temporary shift. We are interested in detecting a permanent shift. Two kind of contaminated normal distributions, $(\sigma_o, \theta) \in$ {(2.5, 6), (2.0, 10)}, are used herein to evaluate the *ARL* and *AQC* of various control charts.

The $\tilde{X}$ charts used for comparison are the Shewhart-$\tilde{X}$ chart (with *L* = 3.128), the CUSUM-$\tilde{X}$ chart (with δ_0 = 0.3, *m* = 8.106), and the EWMA-$\tilde{X}$ chart (with β = 0.1, *L* = 2.827). The $\bar{X}$ charts used for comparison are the Shewhart-$\bar{X}$ chart (with *L* = 3.090), the CUSUM-$\bar{X}$ chart (with δ_0 = 0.3, *h* =7.158), and the EWMA-$\bar{X}$ chart (with α = 0.1, *L* = 2.835). Vis simulation, the values of *L*, *m* and *h* for those control charts are based on the data which 100% come from the (*N*(0, 1)) normal distribution (i.e., $(\sigma_o, \theta) = (0, 0)$), with a desired $ARL_0 \cong 500$.

In Tables 2 to 4, various combinations of (σ_o, θ) denote various contaminated normal distributions. When $(\sigma_o, \theta) = (0, 0)$, the data 100% come from the normal distribution *N*(0, 1). When $(\sigma_o, \theta) \in \{(2.5, 6), (2, 10)\}$, the ARL_0s of $\tilde{X}$ charts are greater than that of the $\bar{X}$ charts. For instance, when $(\sigma_o, \theta) = (2.5, 6)$, the ARL_0 of the CUSUM-$\tilde{X}$ chart is 430.0 is greater than those of three $\bar{X}$ charts (87.1, 265.4, 186.1, respectively). It means that the $\bar{X}$ charts are sensitive to outliers and the CUSUM-$\tilde{X}$ chart is outliers-resistant. This robustness is the strong point of the CUSUM-$\tilde{X}$ chart. However, under various process mean shifts, all of the ARL_1s of $\tilde{X}$ charts are greater than that of the corresponding $\bar{X}$ charts. For instance, when (σ_o, θ) = (2.5, 6) and δ = 0.1, the ARL_1 of the CUSUM-$\tilde{X}$ chart (= 139.4) is greater than that of the CUSUM-$\bar{X}$ charts (= 97.4). The shift-detecting ability of $\tilde{X}$ charts is worse than that of $\bar{X}$ charts without respect to outliers.

Table 2. *ARL*s of different charts with contaminated data ((σ_o, θ) = (0, 0), desired $ARL_0 \cong 500$)

control chart	shift (δ) =	0.0	0.1	0.3	0.5	0.7	1.0	1.5
$\tilde{X}$	Shewhart	500.1	439.2	175.3	69.6	28.3	9.7	2.7
	CUSUM	502.0	151.4	26.2	13.1	8.7	5.7	3.8
	EWMA	499.2	165.4	24.4	10.3	6.4	4.0	2.7
$\bar{X}$	Shewhart	500.2	405.3	128.1	41.5	16.3	5.0	**1.7**
	CUSUM	501.2	**130.0**	20.2	9.9	6.5	4.5	2.9
	EWMA	501.0	136.3	**19.2**	**8.5**	**5.2**	**3.4**	2.4

Table 3. *ARL*s of different charts with contaminated data ((σ_o, θ) = (2.5, 6), desired $ARL_0 \cong 500$)

control chart	shift (δ) =	0.0	0.1	0.3	0.5	0.7	1.0	1.5
$\tilde{X}$	Shewhart	264.5	236.8	126.6	48.6	24.1	9.8	4.0
	CUSUM	**430.0**	139.4	26.8	12.8	8.6	5.8	3.8
	EWMA	343.9	127.4	23.3	10.0	6.2	4.0	2.7
$\bar{X}$	Shewhart	87.1	**78.0**	48.2	24.5	13.8	5.7	2.7
	CUSUM	265.4	97.4	19.2	9.9	6.6	4.4	2.9
	EWMA	186.1	85.2	**17.1**	**8.0**	**5.0**	**3.4**	**2.4**

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table 4. *ARL*s of different charts with contaminated data $((\sigma_o, \theta)$ = (2.0, 10), desired $ARL_0 \cong 500)$

control chart	shift (δ) =	0.0	0.1	0.3	0.5	0.7	1.0	1.5
$\tilde{X}$	Shewhart	218.8	204.5	119.8	56.3	23.8	10.0	3.9
	CUSUM	**395.4**	136.2	25.7	13.0	8.6	5.8	3.8
	EWMA	314.3	137.2	22.4	10.1	6.4	4.1	2.6
$\overline{X}$	Shewhart	106.9	106.5	61.8	27.4	12.4	6.0	2.7
	CUSUM	269.8	**98.4**	19.7	9.8	6.5	4.4	2.9
	EWMA	198.3	104.0	**16.6**	**8.0**	**5.1**	**3.3**	**2.4**

V. Conclusion

A novel CUSUM- $\tilde{X}$ control chart is employed to monitor the process mean. Under some contaminated normal distributions, the EWMA- $\overline{X}$ chart can detect the process shifts faster. The shift-detecting ability of the proposed CUSUM- $\tilde{X}$ chart is similar to that of the EWMA- $\tilde{X}$ chart. In outliers-resistance, the CUSUM- $\tilde{X}$ chart performs best. This conclusion is valuable for the practitioner when the process presents outliers.

References

[1] P. Castagliola, "Approximation of the normal sample median distribution using symmetrical Johnson S_U distributions: Application to quality control," *Communications in statistics: Simulation and Computation*, 1998, 27, 289–301.
[2] P. Castagliola, "An $(\tilde{X}/R)$ -EWMA control chart for monitoring the process sample median," *International Journal of Reliability, Quality and Safety Engineering*, 2001, 8(2), 123–135.
[3] S.V. Crowder, "Design of exponentially weighted moving average schemes," *Journal of Quality Technology*, 1989, 21, 155–162.
[4] F.F. Gan, "An optimal design of cusum quality control charts," *Journal of Quality Technology*, 1991, 23(4), 279–286.
[5] D.M. Hawkins, "Cumulative sum control charting: an underutilized SPC too," *Quality Engineering*, 1993, 5(3), 463–477.
[6] M.B.C. Khoo, "A Control chart based on sample median for the detection of a permanent shift in the process mean," *Quality Engineering*, 2005, 17, 243–257.
[7] J.M. Lucas, "The design and use of V-mask control schemes," *Journal of Quality Technology*, 1976, 8(1), 1–12.
[8] D.C. Montgomery. *Introduction to Statistical Quality Control*. New York: Wiley, USA, 2005, ch. 8.
[9] D.C. Montgomery, J.C.-C. Torng, J.K. Cochran, and F.P. Lawrence, "Statistically constrained economic design of the EWMA control chart," *Journal of Quality Technology*, 1995, 27, 250–256.
[10] E.S. Page, "Continuous inspection schemes," *Biometrics*, 1954, 41(1), 100–115.
[11] E.S. Page, "Cumulative sum control charts," *Technometrics*, 1961, 3(1), 1–9.
[12] S.M. Ross, *A Course in Simulations*. Macmilan, New York, 1990.
[13] S.H. Sheu, and T.C. Lin, "The generally weighted moving average control chart for detecting small shifts in the process mean," *Quality Engineering*, 2003. 16(2), 209–231.
[14] S.H. Sheu, and L. Yang, "The generally weighted moving average median control chart," *Quality Technology and Quantitative Management*, 2006, 3, 455–471.
[15] W.H. Woodall, and B.M. Adams, "The statistical design of CUSUM charts," *Quality Engineering*, 1993, 5(4), 559–570.

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)