The design and use of multiple cusum scheme

*Chuen-Sheng Cheng, *Sheng-Su Cheng
* Department of Industrial Engineering and Management
Yuan-Ze University
135 Yuan-Tung Road Taoyuan, Taiwan, R.O.C.

Key Words : Shewhart charts; CUSUM; Average run length (ARL).

1. Introduction

Control charts are widely used for both manufacturing and service industries.

Traditional Shewhart control chart is a very simple procedure, it works well for detecting large process shifts. A criticism of Shewhart charts is that they only use the information about the process contained in the last observation. One method of increasing the sensitivity of Shewhart charts to smaller process shifts is by introducing what are known as supplementary rules [1~2,3].

However, the use of these rules reduces the simplicity and ease of use of the Shewhart charts. An effective alternative to Shewhart control charts is the CUSUM chart proposed by Page[4]. Algorithmic versions of CUSUM schemes are based on the following recursive statistics,

$$S_H(i) = \max\{0.0, z_i - k + S_H(i-1)\}$$
$$S_L(i) = \max\{0.0, z_i - k + S_L(i-1)\}.$$

where z_i is the standardized statistic of a single observation or the average of several observations taken at sampling time i. The parameter k is usually called the reference values and h the decision interval. An out-of-control signal is triggered either when S_H or S_L exceeds h. The upper one-sided CUSUM chart is intended for detecting an upward shift in the process mean, while a lower one-sided CUSUM chart is intended for detecting a downward shift in the process mean. The staring values of CUSUM statistics are usually set to $S_H(0) = S_L(0) = 0$. However, Lucas and Crosier[5] recommended the use of CUSUM charts with $S_H(0) \neq 0$ and $S_L(0) \neq 0$ when a process starts out with a mean shift.

In general, CUSUM control schemes require the setting of reference value k and decision interval h before implementation. The proper selection of k and h is quite important, as it has substantial impact on the performance of the CUSUM. Page[4] provided some guidance on the design of CUSUM charts for monitoring of the process mean. The parameter k is usually set equal to half the process shift which is important to detect quickly while the parameter h is chosen to get an adequate in-control ARL. In theory, CUSUM control schemes can be used to detect any magnitude of process changes.

CUSUM charts with large values of k are more efficient for signaling larger shifts. On the other hand, smaller values of k are usually useful for detecting smaller shifts. Note that the traditional 3-sigma Shewhart control chart is a CUSUM scheme with $k = 3$ and $h = 0$.

Implementing CUSUM chart requires the knowledge of the future shifts in order to get the quickest detection. If the process shifts are equal to $\Delta\sigma$, with Δ known in advance, then we can optimize the performance of the CUSUM. In reality we never know the magnitude of future shifts. As such, the optimal design of the CUSUM chart for a future shifts is difficult. In addition, often practitioners wish to signal a range of expected shifts efficiently. Therefore, a control procedure is needed that can perform well on average over a range of expected process shifts.

In this paper, a multiple CUSUM scheme is proposed that are robust at signaling a range of process shifts. The multiple CUSUM-scheme is modeled by applying three standard CUSUM charts simultaneously.

The parameters of the multiple CUSUM-scheme are pre-determined for a wide range of process shifts.

2. The Design of a Multiple CUSUM Scheme

The implementation of a CUSUM chart involves the choice of k and h. Although there is a clear formula to calculate the reference value k, however, the CUSUM chart can be optimized when we have accurate information on the magnitude of shifts. If the magnitude of process shifts is unknown, the user of CUSUM charts might select the parameters based on their experience. Note that the value of k will determine the range of process shifts that the CUSUM can produce an efficient monitoring procedure. Therefore, it is intuitive to use several CUSUM charts with different k values to extend the range of shifts that can be detected quickly. We will denote this approach as the multiple CUSUM. It should be noted that the combined Shewhart-CUSUM proposed by Lucas[6] could be viewed as a multiple CUSUM with two CUSUMS applied simultaneously. The multiple CUSUM is designed to be efficient at signaling a wide range of future expected but unknown shifts. In designing a multiple CUSUM, the following issues must be addressed:

1. The number of simultaneous CUSUM.

2. The setting of parameters.

The number of CUSUM charts is determined by the size of range of shifts that need to be detected efficiently. Increasing the number of CUSUM will increase the size of range of shifts that can be detected quickly.

But with this comes greater complexity in its design and introduces much more computational effort. The multiple CUSUM proposed in this paper is the simultaneous application of three CUSUM charts.

In this paper, we formulate the multiple CUSUM by choosing k_1 = 1.0, k_2 = 0.5, and k_3 = 0.25. This produces a monitoring procedure that can perform well for a broad range of shifts. Next we consider the choice of h. In CUSUM schemes the value of h will determine the in-control ARL. We can find the h_j, j = 1, 2, 3 values such that the CUSUM charts will have the same in-control ARL when the CUSUMS are applied individually. However, the overall in-control ARL will be reduced when CUSUMS are applied jointly. It seems reasonable to use higher h values for compensation. The parameters h_j's are adjusted by a factor ρ ($0 < \rho < 1$). The new parameters will be $h_j^* = h_j / \rho$. The following steps are recommended for the design of multiple CUSUM.

Step 1. Specify the in-control ARL for the multiple CUSUM.

Step 2. Determine the decision interval values, h_1, h_2, h_3, so that the standard CUSUM will have the same in-control ARL when the CUSUMS are applied individually. The values of h_j can be obtained by using computer program described in[7,8].

Step 3. Set the factor ρ = 0.5.

Step 4. Let $h_j^* = h_j / \rho$.

Step 5. Compute the in-control ARL of the multiple CUSUM. If the multiple CUSUM produces the in-control ARL specified in Step 1, then stop. Otherwise, go to step 6.

Step 6. Adjust the ρ value by bisection methods and go to step 4.

A simulation program used in step 5 is available from the authors. During implementation, h_j^* will be used in the multiple CUSUM.

3. Supplementary Rules

The multiple CUSUM-scheme is modeled by applying several standard CUSUM charts simultaneously. The advantage of this is the efficient signaling of a range of process shifts. However, this approach also increases the efforts of control charting. It will be better to use a single CUSUM chart when the magnitude of process shifts becomes obvious.

Let $r_j(i) = \max\,[S_{Hj}(i)/h_j^*, S_{Lj}(i)/h_j^*]$ and $r\,max = \max[r_1(i), r_2(i), r_3(i)]$. The following rules are proposed to provide more information about the magnitude of shifts.

Rule 1: if there are 4 successive samples such that $rmax$ comes from the CUSUM with reference value k_1.

Rule 2: if there are 5 successive samples such that $rmax$ comes from the CUSUM with reference value k_2.

If any one rule is met, only single CUSUM scheme with a specific reference value will be active. The monitoring procedure will issue an out-of-control signal whenever $r_j >$ 1.0. The above rules will reduce the complexly of the multiple CUSUM, in addition, they also improve the sensitivity of the multiple CUSUM. Note that there are many ways to form a set of supplementary rules. The rules mentioned above are designed to improve the responsiveness to medium and large process shifts. We will examine their contributions in a later section.

4. Discussion of Results

Control charts are often evaluated by calculating their average run length (ARL).

The ARL of a control chart is defined as the average number of samples taken before an out-of-control signal is obtained. The ARL should be large when the process is in-control and short when the process shifts to undesirable level.

In designing the multiple CUSUM, we assume that all process shifts are equally important. Table 1 compares the ARLs performance of three standard CUSUM charts with k_1 = 1.0, h_1 = 2.63, k_2 = 0.5, h_2 = 5.0, and k_3 = 0.25. h_3 = 8.45, the multiple CUSUM and the combined procedure of multiple CUSUM and supplementary rules. Note that the ρ value of the combined procedure is adjusted in order to maintain the same in-control ARL. The ARLs for different Δ's ranging from 0 to 4 are presented. The ARL's of the multiple CUSUM were estimated using simulation.

The estimated standard error for the ARL is approximately 1% of the ARL value. Figure 1 depicts the ARLs of various CUSUM schemes.

The individual CUSUM is optimal in signaling a specific magnitude of process shifts, so it is worth comparing the control schemes from the global point of view (i.e., over a range of expected shifts). Let ARL* be the average run length obtained from the optimal CUSUM (i.e., setting $k = \Delta / 2$) for a given Δ. For m equally spaced values of different Δ's, an index η is defined with smaller values the better. The index η is defined as follows:

$$\eta = \frac{1}{m} \sum_{i=1}^{m} \frac{ARL_i - ARL_i^*}{ARL_i^*}$$

The η values for different procedures are recorded in the last row of Table 1. Within the range of shifts in Table 1, η of the multiple CUSUM is smaller than that of the CUSUM charts with k = 0.25, 0.5 and 1.0.

The comparison indicates that the multiple CUSUM signals a broader range of shifts more efficiently than conventional CUSUM charts. As can be seen in Table 1, the supplementary rules can improve the efficiency of signaling process shifts. This is evident from the lowest η in Table 1. The results also show that a CUSUM with k =

Table 1 The Comparisons of Global Performance and ARLs for Single and Multiple CUSUM

Δ	ARL *	Standard CUSUM			Multiple CUSUM	Multiple CUSUM combined with
		k = 0.25	k = 0.5	k = 1.0	ρ = 0.875	the supplementary rules ρ = 0.96
0.000	465.0	467.0	465.4	465.8	462.2	465.0
0.250	81.36	92.21	139.5	235.9	103.9	109.5
0.500	30.55	30.55	38.0	78.34	32.60	31.54
0.750	16.39	17.27	17.05	30.10	17.31	16.12
1.000	10.38	12.0	10.38	14.40	11.06	10.16
1.250	7.245	9.197	7.393	8.402	7.769	7.175
1.500	5.381	7.473	5.747	5.677	5.821	5.397
1.750	4.192	6.307	4.714	4.237	4.566	4.224
2.000	3.378	5.469	4.009	3.378	3.712	3.470
2.250	2.795	4.839	3.499	2.818	3.112	2.910
2.500	2.363	4.350	3.114	2.428	2.682	2.516
2.750	2.034	3.959	2.813	2.142	2.365	2.223
3.000	1.776	3.640	2.573	1.923	2.124	1.986
3.250	1.572	3.375	2.381	1.746	1.933	1.806
3.500	1.413	3.155	2.227	1.599	1.776	1.656
3.750	1.289	2.967	2.107	1.473	1.640	1.524
4.000	1.196	2.800	2.013	1.364	1.519	1.409
η		66.97%	32.55%	35.06%	15.27%	8.840%

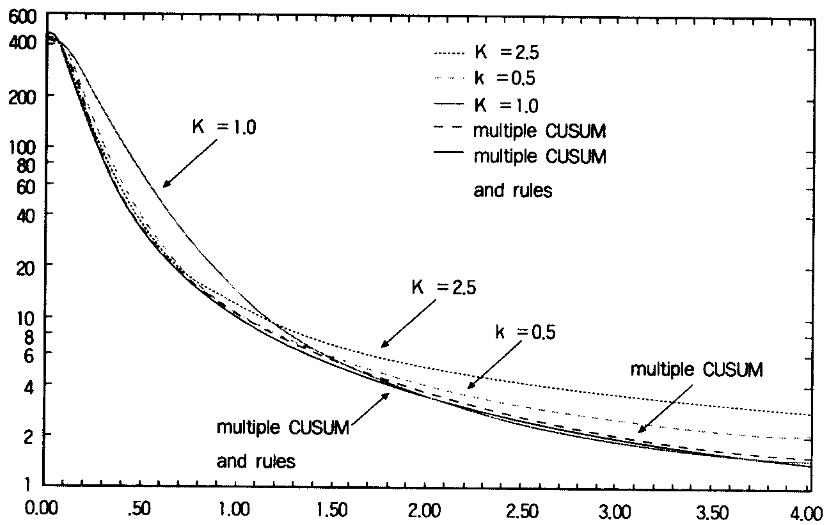

Fig. 1 Average run lengths

0.5 is more robust than the CUSUM with k = 0.25 or 1.0. When a single CUSUM is applied and the magnitude of future shifts is unknown, it seems that k = 0.5 will be the appropriate choice.

In summary, the multiple CUSUM will improve the conventional CUSUM while requiring little additional computational effort. The multiple CUSUM combined with supplementary rules is a more robust procedure than the conventional CUSUM charts at signaling the range of shifts considered.

5. Example

In this section, a selected example from the industry is used to illustrate the advantages of the multiple CUSUM. The data is taken from the process producing bearings. The measurements are made on the inside diameter of the bearings. There are 45 samples available, each containing three observations on the quality characteristic. Figure 2 provides a graphical display of the sample averages. The process is in control for the first 30 samples with μ = 54 and $\sigma_{\bar{x}}$ = 0.35 .

The shift in the mean (about 1.5 $\sigma_{\bar{x}}$) occurs from the sample 31 onwards. Table 2 shows the operation of the multiple CUSUM using standardized values z_i. In Table 2, one can see that individual CUSUM charts with (k = 0.25, h = 8.45), (k =0.5, h =5.0) and (k =1.0, h =2.63) will signal the shifts at observations 41, 39 and 38, respectively. It seems that a CUSUM chart with k = 1.0 will provide a better protection if the shift can be efficiently estimated. In practice, it is very difficult to choose an optimal k value when future shifts are unknown. In this example a multiple CUSUM is applied to monitoring the process. The multiple CUSUM was designed with ρ = 0.875, k_1 = 1.0, h_1^* =2.63/0.875=3.006, k_2 =0.5, h_2^*=5.0/0.875=5.714, k_3 =0.25, h_3^* =8.45/0.875=9.657. This multiple CUSUM provides an in-control ARL equal to 465. The multiple CUSUM procedure will issue an out-of-control signal at the 39th observation ($S_{H2}(39)$ = 5.722> h_2^*=5.714). The result is the same as operating a single CUSUM chart with k =0.5 and h =5.0. Next we examine the contribution of the supplementary rules. It is seen from Table 2 that there are 4 successive samples with r max belonging to S_{H1} from observations 31 to 34. Hence, only the CUSUM chart with k =1.0 will be active starting from the 35th

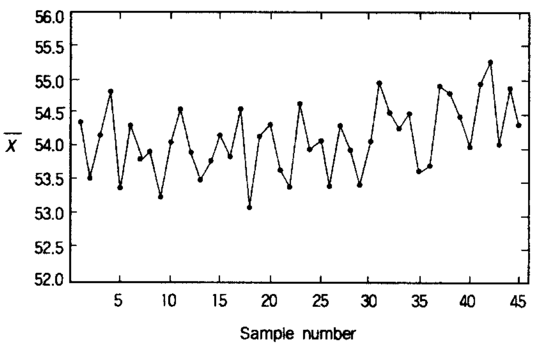

Fig. 2 Plot of sample average

Table 2 Operation of Multiple CUSUM Schemes

No.	z_i	Standard CUSUM						Combined multiple CUSUM and rules		
		$S_{H1}(i)$	$S_{L1}(i)$	$S_{H2}(i)$	$S_{L2}(i)$	$S_{H3}(i)$	$S_{L3}(i)$	r_1	r_2	r_3
1	1.052	0.052	0.000	0.552	0.000	0.802	0.000	0.019	0.106	0.091
2	-1.362	0.000	0.362	0.000	0.862	0.000	1.112	0.132	0.166	0.126
3	0.157	0.000	0.000	0.000	0.205	0.000	0.705	0.000	0.039	0.080
4	2.324	1.324	0.000	1.824	0.000	2.074	0.000	0.483	0.350	0.236
5	-1.854	0.000	0.854	0.000	1.354	0.000	1.604	0.312	0.260	0.182
6	0.966	0.000	0.000	0.466	0.000	0.716	0.388	0.000	0.089	0.081
7	-0.562	0.000	0.000	0.000	0.062	0.000	0.700	0.000	0.012	0.080
8	-0.238	0.000	0.000	0.000	0.000	0.000	0.688	0.000	0.000	0.078
9	-2.184	0.000	1.184	0.000	1.684	0.000	2.622	0.432	0.323	0.298
10	0.147	0.000	0.037	0.000	1.037	0.000	2.225	0.014	0.199	0.253
11	1.559	0.559	0.000	1.059	0.000	1.309	0.416	0.204	0.203	0.149
12	-0.254	0.000	0.000	0.304	0.000	0.804	0.420	0.000	0.059	0.091
13	-1.470	0.000	0.470	0.000	0.970	0.000	1.640	0.172	0.186	0.186
14	-0.658	0.000	0.128	0.000	1.128	0.000	2.048	0.047	0.217	0.233
15	0.461	0.000	0.000	0.000	0.167	0.211	1.337	0.000	0.032	0.152
16	-0.506	0.000	0.000	0.000	0.173	0.000	1.593	0.000	0.033	0.181
17	1.578	0.578	0.000	1.078	0.000	1.328	0.000	0.211	0.207	0.151
18	-2.656	0.000	1.656	0.000	2.156	0.000	2.406	0.604	0.414	0.273
19	0.423	0.000	0.233	0.000	1.233	0.173	1.733	0.085	0.237	0.197
20	0.902	0.000	0.000	0.402	0.000	0.825	0.581	0.000	0.077	0.094
21	-1.040	0.000	0.040	0.000	0.540	0.000	1.371	0.015	0.104	0.156
22	-1.764	0.000	0.804	0.000	1.804	0.000	2.885	0.293	0.346	0.328
23	1.809	0.809	0.000	1.309	0.000	1.559	0.826	0.295	0.251	0.177
24	-0.133	0.000	0.000	0.676	0.000	1.176	0.709	0.000	0.130	0.134
25	0.210	0.000	0.000	0.386	0.000	1.136	0.249	0.000	0.074	0.129
26	-1.758	0.000	0.758	0.000	1.258	0.000	1.757	0.277	0.242	0.200
27	0.836	0.000	0.000	0.336	0.000	0.586	0.671	0.000	0.065	0.076
28	-0.191	0.000	0.000	0.000	0.000	0.144	0.612	0.000	0.000	0.070
29	-1.685	0.000	0.685	0.000	1.185	0.000	2.047	0.250	0.228	0.233
30	0.193	0.000	0.000	0.000	0.492	0.000	1.604	0.000	0.094	0.182
31	2.662	1.662	0.000	2.162	0.000	2.412	0.000	***0.606***	0.415	0.274
32	1.424	2.086	0.000	3.086	0.000	3.586	0.000	***0.761***	0.593	0.407
33	0.723	1.809	0.000	3.309	0.000	4.059	0.000	***0.660***	0.635	0.461
34	1.419	2.229	0.000	4.229	0.000	5.229	0.000	***0.813***	0.812	0.594
35	-1.114	0.114	0.114	2.614	0.614	3.864	0.864	0.042	—	—
36	-0.886	0.000	0.000	1.229	1.000	2.729	1.500	0.000	—	—
37	2.516	1.516	0.000	3.245	0.000	4.995	0.000	0.553	—	—
38	2.229	2.745	0.000	4.974	0.000	6.974	0.000	1.002	—	—
39	1.248	2.993	0.000	5.722	0.000	7.972	0.000	1.093	—	—
40	-0.067	1.926	0.000	5.155	0.000	7.655	0.000	0.703	—	—
41	2.659	3.585	0.000	7.314	0.000	10.064	0.000	1.309	—	—
42	3.619	6.204	0.000	10.433	0.000	13.433	0.000	2.265	—	—
43	0.069	5.273	0.000	10.002	0.000	13.252	0.000	1.925	—	—
44	2.460	6.733	0.000	11.962	0.000	15.462	0.000	2.458	—	—
45	0.948	6.681	0.000	12.410	0.000	16.160	0.000	2.439	—	—

observation. The procedure of combinedmultiple CUSUM (ρ = 0.96) and supplementaryrules signals an out-of-control situation at observation 38 ($r_1(38)$ =1.002>1.0). The result is the same as a correct choice of the reference value (i.e., k =1.0 in this case) when operating a CUSUM control chart. This example clearly demonstrates that the multiple CUSUM can be further improved by using supplementary rules.

6. Conclusions

The conventional CUSUM is optimal with $k = \Delta/2$, if it is known that the process shift is $\Delta\sigma$. In practice, the value of Δ is seldom known in applications. In this paper, we propose a multiple CUSUM that can efficiently signal over a wide range of process shifts. The multiple CUSUM is thought of as the simultaneous use of several standard CUSUM schemes. A design procedure is proposed to aid in the design of multiple CUSUM. A set of rules is also proposed to simplify the computational effort of the multiple CUSUM. The results indicate that the multiple CUSUM performs better than a single CUSUM for most process shifts. If the magnitude of the process shifts is unknown, then the proposed approach is expected to perform well.

Acknowledgement

This research was supported in part by the National Science Council of the Republic of China under the Grant No. NSC-89-2213-E-155-004.

References

1. Nelson, L.S., Interpreting Shewhart Control Charts, *Journal of Quality Technology*, 17, 114-116 (1985).
2. Nelson, L.S., The Shewhart Control Chart-Tests for Special Cause, *Journal of Quality Technology*, 16, 237-239 (1984).
3. Western Electric Company, *Statistical Quality Control Handbook*, Indianapolis, Western Electric Co. Inc., Indiana (1958).
4. Page, E.S., Continuous Inspection Schemes, *Biometrika*, 41, 100-115 (1954).
5. Lucas, J.M., and Crosier, R. B., Fast Initial Response for CUSUM Quality Control Schemes: Give Your CUSUM a Head Start, *Technometrics*, 24,199-205 (1982).
6. Lucas, J.M., Combined Shewhart-CUSUM Quality Control Schemes, *Journal of Quality Technology*, 14, 51-59 (1982).
7. Charles, W.C., and Rigdon, S. E., A Comparison of the Markov Chain and the Integral Equation Approaches for Evaluating the Run Length Distribution of Quality Control Charts, *Communication in Statistics-Simulation and Computation*, 20, 191-204 (1991).
8. Gan, F.F., The Run Length Distribution of a Cumulative Sum Control Chart, *Journal of Quality Technology*, 25, 205-215 (1993).