Journal of Academia and Industrial Research (JAIR)
Volume 3, Issue 1 June 2014

ISSN: 2278-5213

RESEARCH ARTICLE

Nonparametric CUSUM Charts for Process Variability

D.M. Zombade[1] and V.B. Ghute[2*]
[1]Dept. of Statistics, Walchand College of Arts and Science, Solapur, (MS), India
[2]Dept. of Statistics, Solapur University, Solapur-413255, (MS), India
vbghute_stats@rediffmail.com*, digambarzombade@gmail.com; +91 9422650105*, 9881372729

Abstract

In this study, two nonparametric CUSUM type control charts are developed for monitoring the process variability. The proposed charts are based on nonparametric two sample tests for testing equality of variance. The average run length (ARL) performance of the proposed CUSUM-type charts is evaluated through a simulation study and compared with the corresponding Shewhart-type nonparametric charts. The study indicates that the proposed CUSUM-type nonparametric control charts are better in detecting small shifts in process variability, while for larger shifts Shewhart-type control charts has better performance.

Keywords: Control chart, average run length, process variability, CUSUM-type charts, nonparametric tests.

Introduction

Control charts are statistical process control tools that are widely used for monitoring mean and variability of a process. Most of the control charts that have been developed in literature are designed and evaluated under the assumption that the underlying distribution of the quality characteristic is normal. In real applications, there are many situations in which the process data come from a non-normal distribution which need to be monitored by appropriate control charts. To monitor such type of data, development of control charts that do not depend on a particular distributional assumption is desirable. Nonparametric control charts can serve this purpose. The main advantage of a nonparametric control chart is that it does not assume any probability distribution for the characteristic of interest. A formal definition of nonparametric or distribution-free control chart is given in terms of its in-control run length distribution. If the in-control run length distribution is same for every continuous distribution then the chart is called distribution-free.

In literature, several nonparametric control charts are proposed for monitoring location of a univariate process. Some of these are based on signs and/or rank statistics by assuming a known in-control target value for process location. Amin *et al.* (1995) developed Shewhart and cumulative sum (CUSUM) control charts based on sign test statistic. Bakir and Reynolds (1979) developed a nonparametric CUSUM to monitor a process center based on with-in group signed-ranks. Amin and Searcy (1991) used with-in group signed ranks to develop exponentially weighted moving average (EWMA) control chart. Bakir (2004) developed a distribution-free Shewhart control chart for monitoring process center based on the signed-ranks of grouped observations. Bakir (2006) proposed Shewhart, CUSUM and EWMA control charts based on signed-rank-like statistics of grouped data for monitoring a process center when in-control target center was not specified and studied the robustness of the charts against outliers. Chakraborti *et al*. (2001) presented an extensive review of the literature on univariate nonparametric control charts. Ghute and Shirke (2012) developed nonparametric control chart based on bivariate signed-rank test to monitor the changes in the location of a bivariate process. Ghute (2013) developed distribution-free control chart based on Hodges sign test to monitor the location of a bivariate process. There exist only few articles on nonparametric charts for monitoring process variability. Lehmann (1975) suggested using non-parametric tests for the equality of two variances for use as control statistics in nonparametric control charts for variability. Control charts using tests statistics for comparing two variances would require obtaining an initial sample (of size m) when the process is considered to be in-control. Then at each sample time i, a sample of size n is obtained from the process and the pooled sample of size ($m + n$) is obtained. The observations in the pooled sample are then ranked from smallest to largest and some statistic based on the ranks of the observations is calculated. Das and Bhattacharya (2008) proposed a nonparametric control chart for monitoring process variability based on Conver's squared rank test for variance. Das (2008) developed two nonparametric control charts for monitoring process variability based on two nonparametric tests. Murakani and Matsuki (2010) proposed a nonparametric control chart for dispersion based on the rank sum statistic. Since few works are reported in the literature on nonparametric control charts for monitoring process variability, the purpose of this study is to develop CUSUM-type nonparametric control charts for monitoring process variability for the case that the location parameter is under control.

The proposed nonparametric control charts are based on two sample nonparametric tests proposed by Mood (1954) and Sukhatme (1956). These are most powerful test statistics for detecting scale shifts. The performance of the proposed charts is assessed for both the in-control state and out-of-control state under different underlying distributions.

Materials and methods

Shewhart-type nonparametric control chart based on Sukhatme test: Suppose we want to compare two independent random samples $X=(X_1,X_2,...,X_m)$ and $Y=(Y_1,Y_2,...,Y_n)$ which are drawn from absolute continuous distributions and differ only in the scale parameters. Let σ_X and σ_Y be the arbitrary measures of dispersion of X and Y respectively then problem of testing of hypothesis is $H_0:\sigma_X=\sigma_Y$ against $H_1:\sigma_X\neq\sigma_Y$.

The Sukhamte test statistic for testing null hypothesis is defined as:

$$T=\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}D(X_i,Y_j)\,, \tag{1}$$

$$\text{where}\quad D(X,Y)=1 \quad \text{if either } 0<X<Y \text{ or } Y<X<0$$
$$=0 \quad \text{otherwise}$$

We reject hypothesis if T is too large or too small. The mean and variance of the statistic T is given by:

$$E(T)=\frac{1}{4}\ \text{and}\ Var\,(T)=\frac{(m+n+7)}{48mn}$$

For a large sample,

$$Z=\frac{T-E(T)}{\sqrt{Vat\,(T)}} \tag{2}$$

has a standard normal distribution and the test is performed on the basis of tabulated values of the standard normal distribution.

We consider Z as the control chart statistic for the Shewhart type nonparametric control chart for monitoring process variability and the chart is referred as NP-S chart. We consider $X=(X_1,X_2,...,X_m)$ as reference sample of size *m* from an in-control process and that $Y=(Y_1,Y_2,...,Y_n)$ be an arbitrary test sample of size *n*. The sample statistics Z computed from independent observations from the process are plotted against an upper control limit UCL = 3 and LCL = -3. The process is considered out-of-control when a plotted point lies above UCL or below LCL.

Shewhart-type nonparametric control chart based on Mood test: Suppose $X=(X_1,X_2,...,X_m)$ and $Y=(Y_1,Y_2,...,Y_n)$ are two independent random samples drawn from absolute continuous distributions and differ only in the scale parameters. We wish to test $H_0:\sigma_X=\sigma_Y$ against $H_1:\sigma_X\neq\sigma_Y$. Let $R_1<R_2<....<R_m$ be the combined samples ranks of the X-values in increasing order of magnitude. The Mood test statistic for testing null hypothesis is defined as:

$$M=\sum_{i=1}^{m}\left(R_i-\frac{N+1}{2}\right)^2,\quad \text{where}\quad N=m+n \tag{3}$$

The mean and variance of the statistic M is given as:

$$E(M)=\frac{m(N^2-1)}{12}\quad\text{and}\quad Var(M)=\frac{mn(N+1)(N^2-4)}{180}$$

For N greater than or equal to 30, we may consider the normalized random variable W

$$W=\frac{M-E(M)}{\sqrt{Var(M)}} \tag{4}$$

and perform the test on the basis of tabulated values of the standard normal distribution.

We consider W as the control chart statistic for the Shewhart-type nonparametric control chart for monitoring process variability and the chart is referred as NP-M chart. We consider $X=(X_1,X_2,...,X_m)$ as reference sample of size *m* from an in-control process and that $Y=(Y_1,Y_2,...,Y_n)$ be an arbitrary test sample of size *n*. The sample statistics W computed from independent observations from the process are plotted against an upper control limit UCL = 3 and LCL = -3. The process is considered out-of-control when a plotted point lies above UCL or below LCL.

CUSUM-type nonparametric control charts: It is well known that the Shewhart-type control charts are relatively inefficient in detecting small shifts of the process parameters. Memory based control charts such as cumulative sum (CUSUM), exponentially weighted moving average (EWMA) are developed as alternative to the Shewhart charts for the detection of small process shifts in the process parameters. They use the additional information from recent history of process hence are more effective than a Shewhart control chart in detecting small process shifts. Therefore, in this study, two nonparametric CUSUM procedures are developed to monitor process variability. The proposed CUSUM procedures are based on two sample nonparametric tests proposed by Mood (1954) and Sukhatme (1956).

To develop nonparametric CUSUM chart to monitor a process variability, nonparametric test statistic T is obtained from two independent random samples $X=(X_1, X_2, ..., X_m)$ and $Y=(Y_1, Y_2, ..., Y_n)$. Then one sided nonparametric CUSUM chart statistic based on nonparametric statistic T is given by:

$$S_t = Max(0, S_{t-1} + T_t - k), \qquad (5)$$

Where $S_0 = 0$ and k > 0. The CUSUM scheme signals at first t for which $S_t > h$, where h > 0 and k > 0 are parameters of the procedure. The parameter k is the reference value and h is the decision interval for the CUSUM.

The one sided nonparametric CUSUM chart statistic based on Sukhatme statistic Z is given by:

$$S_t = Max(0, S_{t-1} + Z_t - k), \qquad (6)$$

and chart based on this statistic is referred as NPCSM-S chart.

The one sided nonparametric CUSUM chart statistic based on Mood statistic W is given by:

$$S_t = Max(0, S_{t-1} + W_t - k), \qquad (7)$$

and chart based on this statistic is referred as NPCSM-M chart.

Results and discussion

Performance of the proposed control charts: To examine the ability of proposed NPCSM-S and NPCSM-M charts to detect variability shift in a process, we consider underlying process distributions as normal, double exponential and uniform with mean zero and variance one. The uniform distribution is considered as process distribution to see the effect of a light tailed distribution and double exponential distribution is considered to see the effect of heavy tailed distribution on the performance of proposed nonparametric control charts.

Consider a process where quality characteristic of interest X is distributed with mean μ and standard deviation σ. Let μ_0 and σ_0 be the in-control values of μ and σ respectively. When a shift in process standard deviation occurs, we have change from the in-control value σ_0 to the out-of-control value $\sigma_1 = \delta\sigma_0$ $(0 < \delta \neq 1)$. Therefore, when control chart for variability is employed, the process shifts are measured through $\delta = \frac{\sigma_1}{\sigma_0}$.

When $\delta = 1$, the process is considered to be in-control.

For $\delta > 1$ an increase in σ occurs and for $\delta < 1$, decrease in σ occurs. Computer programs written in C language are used to study the performance of the proposed control charts. In the (upper) one-sided, CUSUM procedure the reference value k is taken as $\frac{1}{2}\sigma_0$. Using this value of k, the value of h should then be chosen to achieve the desired in-control ARL. The in-control and out-of-control ARL values of the proposed control charts are computed using 10000 simulations for sample size of *n* = 15 and 20.

Table 1 and 2 provide the ARL values of the Shewhart-type and CUSUM-type nonparametric control charts based on Sukhatme test statistic when the underlying process data actually follows normal, double exponential and uniform distributions with sample sizes *n* =15 and 20 respectively.

Examinations of Table 1 and 2 lead to the following findings:

- In-control ARL values of the proposed NP-S and NPCSM-S control charts for different process distributions are approximately same.
- For small shifts, out-of-control ARL values of NPCSM-S chart are smaller than that of the NP-S chart. Therefore, NPCSM-S chart is more efficient than NP-S chart for detecting small shifts in process when underlying process distribution is normal, light tailed uniform and heavy tailed double exponential.
- For uniformly distributed data, both NPCSM-S and NP-S charts perform better than normally and doubly exponential data.

Table 1. ARL values of NP-S and NPCSM-S charts when *n* = 15.

Shift δ	Normal		Double exponential		Uniform	
	NP-S	NPCSM-S	NP-S	NPCSM-S	NP-S	NPCSM-S
1.0	285	285	285	285	285	285
1.2	113.27	26.62	142.82	40.52	71.31	14.39
1.4	47.15	9.76	69.43	14.38	24.64	6.16
1.6	23.78	6.11	39.53	8.62	11.98	4.33
1.8	14.15	4.62	24.77	6.27	7.20	3.50
2.0	9.27	3.85	16.93	5.12	4.90	3.03
3.0	2.96	2.53	5.12	3.09	2.03	2.25
4.0	1.84	2.18	2.86	2.52	1.49	2.05
5.0	1.49	2.05	2.03	2.28	1.30	1.97

Table 2. ARL values of NP-S and NPCSM-S charts when *n* = 20.

Shift δ	Normal		Double exponential		Uniform	
	NP-S	NPCSM-S	NP-S	NPCSM-S	NP-S	NPCSM-S
1.0	285	285	285	285	285	285
1.2	96.72	19.84	125.26	31.08	54.71	10.81
1.4	33.36	7.54	54.50	10.94	15.36	4.89
1.6	14.99	4.85	26.83	6.71	4.12	3.50
1.8	8.45	3.77	15.62	4.94	4.12	2.89
2.0	5.43	3.19	9.99	4.10	2.86	2.53
3.0	1.84	2.18	2.97	2.56	1.39	1.99
4.0	1.28	1.95	1.75	2.17	1.15	1.83
5.0	1.14	1.84	1.36	2.00	1.07	1.72

Table 3. ARL values of NP-M and NPCSM-M charts when *n* = 15.

Shift δ	Normal		Double exponential		Uniform	
	NP-M	NPCSM-M	NP-M	NPCSM-M	NP-M	NPCSM-M
1.0	415	412	415	410	412	414
1.2	137.21	49.33	184.70	62.19	66.75	31.74
1.4	46.02	28.65	78.83	36.21	17.62	19.76
1.6	20.50	21.38	39.69	26.81	7.49	15.60
1.8	11.25	17.73	23.02	21.95	4.53	13.52
2.0	7.10	15.52	14.85	19.07	3.19	12.28
3.0	2.23	11.23	4.05	13.13	1.53	9.90
4.0	1.49	9.89	2.26	11.14	1.23	8.99
5.0	1.28	9.05	1.71	10.18	1.16	8.58

Table 4. ARL values of NP-M and NPCSM-M charts when *n* = 20.

Shift δ	Normal		Double exponential		Uniform	
	NP-M	NPCSM-M	NP-M	NPCSM-M	NP-M	NPCSM-M
1.0	418	418	414	414	416	415
1.2	109.29	38.84	166.97	49.28	45.11	24.65
1.4	31.19	22.25	58.02	28.18	10.23	15.18
1.6	12.73	16.53	26.67	20.82	4.42	11.99
1.8	6.61	13.70	14.51	17.00	2.65	10.42
2.0	4.09	12.02	8.92	14.72	1.95	9.46
3.0	1.47	8.70	2.44	10.13	1.16	7.55
4.0	1.15	7.67	1.51	8.61	1.05	6.96
5.0	1.07	7.17	1.24	7.86	1.03	6.63

Table 3 and 4 provide the ARL values of the Shewhart-type and CUSUM-type nonparametric control charts based on Mood test statistic when the underlying process data actually follows normal, double exponential and uniform distributions with sample sizes *n* =15 and 20 respectively. Examinations of Table 3 and 4 lead to the following findings:

- In-control ARL values of the proposed NP-M and NPCSM-M control charts for different process distributions are approximately same.
- For small shifts, out-of-control ARL values of NPCSM-M chart are smaller than that of the NP-M chart. Therefore, NPCSM-M chart is more efficient than NP-M chart for detecting small shifts in process when underlying process distribution is normal, light tailed uniform and heavy tailed double exponential.
- For uniformly distributed data, both NPCSM-M and NP-M charts perform better than normally and doubly exponential data.

Conclusion

In this study, two nonparametric CUSUM-type control charts are developed for monitoring process variability. The performance of the proposed control charts is studied by simulation and compared with corresponding Shewhart-type control charts under normal, light tailed and heavy tailed distributions. Our simulation study indicates that the NPCSM-M and NPCSM-M control charts are more efficient than NP-S and NP-M control charts for detecting small shifts in process variability for different process distributions. Both NP-M and NP-S control charts perform better when underlying process distribution is light tailed.

Acknowledgements

Authors would like to thank the Editor for comments and suggestions. The first author would like to acknowledge the support by University Grants Commission, New Delhi for award of Teacher Fellowship under Faculty Improvement Program during the XIIth plan.

References

1. Amin, R. and Searcy, A. 1991. A nonparametric exponentially weighted moving average control scheme. *Commun. Stat. Simulation Comput.* 20: 1049-1072.
2. Amin, R., Reynolds, M. and Bakir, S. 1995. Nonparametric quality control charts based on the sign statistic. *Commun. Stat. Theory Meth.* 24: 1597-1623.
3. Bakir, S. 2004. A distribution-free Shewhart quality control chart based on signed-ranks. *Quality Engg.* 16: 613-623.
4. Bakir, S. 2006. Distribution-free quality control charts based on signed-rank-like statistics. *Commun. Stat. Theory Meth.* 35: 743-757.
5. Bakir, S. and Reynolds, M. 1979. A non-parametric procedure for process control based on within-group ranking. *Technometrics*. 21: 175-183.
6. Chakraborti, S., Vanderlaan, P. and Bakir, S. 2001. Nonparametric control charts: An overview and some results. *J. Qual. Technol*. 3: 304-315.
7. Chakraborti, S., Vanderlaan, P. and Vandewiel, M. 2004. A class of distribution-free control charts. *J. Royal Stat. Soc. Series C: Appl. Stat.* 53(3): 443-462.
8. Das, N. 2008. Non-parametric control chart for controlling variability based on rank test. *Econ. Qual. Control*. 2: 227-242.
9. Das, N. and Bhattacharya, A. 2008. A new nonparametric control chart for controlling variability. *Qual. Technol. Quant. Management.* 5(4): 351-361.
10. Ghute, V. 2013. Distribution-free control chart for bivariate process. *J. Acad. Indus. Res.* 1(11): 703-705.
11. Ghute, V. and Shirke, D. 2012. A nonparametric signed-rank control chart for bivariate process location. *Qual. Technol. Quant. Management.* 9(4): 317-328.
12. Lehmann, E. 1975. Nonparametric Statistical Methods Based on Ranks. Holden-Day, San Fransisco, California.
13. Mood, A. 1954. On the asymptotic efficiency of certain nonparametric two-sample tests. *Ann. Math. Stat.* 25: 514-522.
14. Murakami, H. and Matsuki, T. 2010. A nonparametric control chart based on the Mood statistic for dispersion. *Int. J. Advan. Manuf. Techn.* 49: 757-763.
15. Sukhamte, B. 1956. On certain two-sample nonparametric tests for variances. *Ann. Math. Stat.* 28(1): 188-194.