A Study on the Run Length Distribution of Synthetic $\bar{X}$ Chart

Z. L. Chong, Michael B. C. Khoo, and H. W. You

Abstract—**Quality is a vital element in every field. In order to compete among the competitors, the manufacturer or service provider should produce or provide quality product or service. However, before a product or service can be considered as having an acceptable quality level, all the processes producing the product (or providing the service) should be in-control. The synthetic $\bar{X}$ chart consists of two sub-charts, i.e. the $\bar{X}$ /S sub-chart and CRL/S sub-chart. The synthetic $\bar{X}$ chart is used to detect shifts in the process mean. In this paper, we study the percentage points (percentiles) of the run length distribution of the synthetic $\bar{X}$ chart so as to provide practitioners with a more complete understanding of the behaviour of the chart, instead of the sole dependence on the average run length (ARL).**

Index Terms—**Average run length (ARL), conforming run length (CRL), run length distribution, synthetic $\bar{X}$ chart.**

I. Introduction

Quality is an important element for a consumer when choosing among competing products or services. A quality product or service reduces problems that it may pose to a consumer. The concept of quality is widely implemented in modern industries and services. In general, a consumer will consider a product or service to be of good quality when the characteristics of the product or service can satisfy his/her needs or requirements. Apart from this, a quality product or service must perform its intended function properly.

In reality, no process can be in a stable condition forever. Hence, statistical process control (SPC) needs to be implemented. SPC is a collection of problem-solving tools that are used in achieving stability and improving capability through the reduction of variability in a process. Among the SPC tools, a control chart is the most powerful tool to control the quality of a process and improve the quality of products or services. Control charts enable us to investigate whether a process is stable or not. Once a process is stable, the performance of the process will be evaluated and eventually process improvement can be carried out.

The Shewhart $\bar{X}$ chart or simply the $\bar{X}$ chart was proposed by Walter A. Shewhart. It is used to monitor the stability and performance of the process mean. It is known that the $\bar{X}$ chart can quickly detect large shifts in the process mean. However, it is insensitive to small and moderate mean shifts. This is because the $\bar{X}$ chart uses only the most recent sample to decide whether a process is in-control or out-of-control.

The conforming run length (CRL) chart was first introduced by [1]. The CRL chart is an attributes chart to detect shifts in the fraction non-conforming, *p*. The CRL chart determines whether a process is in-control or not, based on the number of conforming units between two non-conforming units. Therefore, an attributes chart is simpler for process monitoring than a variables chart. Due to this advantage, the CRL chart is studied by many researchers, for example [2], [3].

A CRL value is the number of conforming units between two non-conforming units, including the ending non-conforming unit. This chart works based on the idea that the distribution of the CRL changes with *p*, where the CRL decreases as *p* increases and the CRL increases as *p* decreases.

The idea to use the percentage points of the run length distribution to assess a chart's performance was introduced by [4], [5], even though it is easier to use the average run length (ARL) to study a chart's performance compared to a study of the entire run length distribution.

According to [6], the run length distribution is skewed to the right and it follows a geometric distribution. Thus, the entire run length distribution should be studied as it provides useful information regarding the performance of a control chart. A study of the run length distribution of a chart provides a complete understanding on the behavior of a process. In conjunction with this, [7] published tables of the percentage points of the run length distribution for the Shewhart chart with supplementary rules.

The run length distribution is characterized by the percentage points (percentiles) of its distribution, such as the median. The percentage points are more informative than the ARL, which are explained as follows. First, the lower percentage points of the run length distribution when the process is in-control give information about the early false alarm rates. Second, a detailed study of the percentage points provides vital information about the spread of the run length distribution.

The organization of this paper hereafter is as follows: Section II reviews the synthetic $\bar{X}$ chart charting procedure. Optimal design of the synthetic $\bar{X}$ chart and its performance are discussed in Section III. Section IV discusses the results and findings of this research. Finally, conclusions are drawn in Section V.

Manuscript received February 22, 2015; revised May 16, 2015. This research is supported by the Universiti Sains Malaysia (USM), Research University Grant (RUI), no. 1001/PMATHS/811263.

The authors are with the School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia (e-mail: chong.zhilin@gmail.com, mkbc@usm.my, you.woon@yahoo.com).

DOI: 10.7763/IJET.2016.V8.915

II. The Synthetic $\bar{X}$ Chart

The synthetic $\bar{X}$ chart is a combination of the $\bar{X}$ /S sub-chart and the conforming run length (CRL) sub-chart. The synthetic $\bar{X}$ chart was suggested by [8]. Costa and Rahim [9] stated that an increased interest in researching on the synthetic $\bar{X}$ chart, compared to other types of charts, may be due to the fact that an analyst prefers waiting until the presence of a second point falling outside the control limits before taking action to investigate the assignable causes.

The $\bar{X}$ /S sub-chart contains two control limits, i.e. the lower control limit $\left(\mathrm{LCL}_{\bar{X}/S}\right)$ and the upper control limit $\left(\mathrm{UCL}_{\bar{X}/S}\right)$. [10] Castagliola and Khoo [10] noted that CRL is the number of inspected sample points between two consecutive non-conforming ones, including the non-conforming point at the end. The CRL/S sub-chart only has one control limit, i.e. the lower limit, denoted as *L*, where *L* is a positive integer. According to [11] when *L* increases, the synthetic $\bar{X}$ chart behaves approximately like the usual Shewhart $\bar{X}$ chart.

Wu and Spedding [8] showed that the synthetic $\bar{X}$ chart is quicker in detecting small mean shifts compared to the Shewhart $\bar{X}$ chart. In addition, [8] also pointed out that the synthetic $\bar{X}$ chart is less effective than the Exponentially Weighted Moving Average (EWMA) chart in detecting small shifts. However, it is usually not advisable to detect very small shifts in a process to avoid too frequent process interruption. According to [8], the control limits for the $\bar{X}$ /S sub-chart are as follows:

$$\mathrm{LCL}_{\bar{X}/S} = \mu_0 - k\frac{\sigma_0}{\sqrt{n}} \tag{1}$$

$$\mathrm{CL}_{\bar{X}/S} = \mu_0 \tag{2}$$

$$\mathrm{UCL}_{\bar{X}/S} = \mu_0 + k\frac{\sigma_0}{\sqrt{n}} \tag{3}$$

where *k* controls the width of the limits. To construct the synthetic $\bar{X}$ chart, first a sample of size *n* is taken and the sample mean is computed and denoted as $\bar{X}$. If $\mathrm{LCL}_{\bar{X}/S} \le \bar{X} \le \mathrm{UCL}_{\bar{X}/S}$, the sample is considered as conforming, otherwise the sample is non-conforming. The lower control limit of the CRL/S sub-chart is set as *L*, where *L* is an integer. The number of samples between the present and last non-conforming sample is counted and is taken as the CRL value. Then if $\mathrm{CRL} > L$, the process is classified as in-control, otherwise the process is out-of-control.

It is worth noting that if a sample point $\bar{X}$ plots beyond the limits of the $\bar{X}$ /S sub-chart, we cannot immediately conclude that the process is out-of-control. Instead, it just indicates that a non-conforming sample exists.

III. Optimal Design of the Synthetic $\bar{X}$ Chart

An optimal design of the synthetic $\bar{X}$ chart aims at finding a set of optimal parameters that will minimize ARL, when ARL_0 is fixed at a desired value. For the synthetic $\bar{X}$ chart, there are two optimal parameters, i.e. *L* (lower control limit of the CRL/S sub-chart) and *k* (control limits constant of the $\bar{X}$ /S sub-chart) that must be determined using the Mathematica program.

The user needs to specify the sample size, *n*, size of optimal shift of interest, δ_d and desired in-control ARL, ARL_0 before the computation of optimal *L* and *k* values can be made. However, the target values of μ and σ need not be specified since they are not required in the computation of the optimal *L* and *k* values.

In the simulation study using the Mathematica program simulated by computer, $n = 3$, 5, 7 and 10 are chosen. δ_d is set as 1, while ARL_0 is fixed as 370. The optimization procedure works by first initializing *L* as 1. Therefore, *L* is the only independent design variable. Then the value of *k* is computed using (4)

$$\frac{1}{P0exact1} \times \frac{1}{1-(1-P0exact1)^L} == ARL0, \{P0exact1, 1\times 10^{-8}\} \tag{4}$$

where $P0exact1 = 2\Phi(-k)$. In this syntax, *L* is the independent design variable, which means that we can fix the value of *L*. In this program, *L* is initialized as 1. Then *L is* increased incrementally by one, starting from one until fifty. From the fifty pairs of (*L*, *k*) and their corresponding ARL_1 values, the pair that produces the smallest ARL_1 is taken as the optimal pair. It is very likely that the smallest ARL_1 is obtained by the time $L = 50$.

The ARL_1 value can be computed from any (*L*, *k*) pair using the following syntax:

$$ARLdelta = \frac{1}{Pdelta} \times \frac{1}{1-(1-Pdelta)^L}\ , \tag{5}$$

where $Pdelta$ is the power of the $\bar{X}$ */S* sub-chart and $Pdelta$ can be obtained by solving the following equation in the Mathematica program:

$$\begin{aligned} Pdelta &= 1 - CDF\left[Normal[0,1], k-\delta\sqrt{n}\right] \\ &\quad + CDF\left[Normal[0,1], -k-\delta\sqrt{n}\right], \end{aligned} \tag{6}$$

After finding the optimal parameters, *L* and *k*, we can proceed to find the percentage points (percentiles) of the run length distribution using the Statistical Analysis System (SAS) program. Since the run length distribution is highly skewed, examining the entire run length distribution provides a better understanding about a chart's performance.

The simulation study considers both increasing and decreasing shifts in the mean. An increasing shift occurs when the mean of the underlying process increases from μ_0 to $\mu_0 + \delta\sigma$, while a decreasing shift occurs when the mean decreases from μ_0 to $\mu_0 - \delta\sigma$.

In this study, $n \in \{3, 5, 7, 10\}$ and $\delta \in \{0, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3\}$ are considered. The percentage points of the run length distribution for the synthetic $\bar{X}$ chart are computed based on 50000 simulation trials. The percentage points of the run length distribution for the in-control and out-of control processes are computed for the 0.1th, 1st, 5th, 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th and 90th percentiles. The underlying process is assumed to follow a normal distribution.

A SAS program is written to compute the percentage points of the run length distribution for the synthetic $\bar{X}$ chart. Here, mu0 = 0 and sigma = 1 are used in the program, where mu0 and sigma represent μ_0 and σ, respectively. Four parameters are specified in this SAS program. They are δ, *n*, *L* and *k*. The *L* and *k* values are computed by the Mathematica program mentioned above. In SAS, the following equations give the lower and upper control limits of the $\bar{X}$/S sub-chart:

$$\text{LCL} = \text{mu0} - k * \text{sigma} / \text{sqrt}(n), \tag{7}$$

$$\text{LCL} = \text{mu0} + k * \text{sigma} / \text{sqrt}(n). \tag{8}$$

Note that the values of the random variable *X*, which follows a normal distribution are generated as follows:

$$X = \text{delta} + \text{sigma} * \text{rannor}(33333), \tag{9}$$

where 33333 is a random seed used in the program. A random seed value is used to generate a series of random numbers. The in-control observations are generated by setting delta = 0, while the out-of-control observations are simulated by letting delta = δ, where $\delta > 0$.

IV. Results and Findings

Table I shows the optimal parameters computed using the Mathematica program, for sample sizes *n* = 3, 5, 7 and 10. From Table I, it is observed that as the sample size *n* increases, the corresponding values of the optimal parameters, *L* and *k* decrease.

TABLE I: Optimal Design Parameters

n	*L*	*k*
3	6	2.29367
5	4	2.21855
7	3	2.16382
10	2	2.08459

After obtaining the optimal parameters, the SAS program is used to compute the percentage points of the run length distribution of the synthetic $\bar{X}$ chart. Tables II, III, IV and V give the percentage points for $n \in \{3, 5, 7, 10\}$.

From Table II to Table V, it is observed that when the process is in-control, i.e. $\delta = 0$, the value of $\text{ARL}_0 = 370$ falls between the 60th and 70th percentiles of the run length distribution. This shows that the in-control run length distribution of the synthetic $\bar{X}$ chart is skewed to the right. These tables also show that the skewness of the run length distribution decreases with an increase in the magnitude of the shift δ. This indicates that the skewness of the run length distribution changes with the magnitude of the shift, hence interpretation of a chart's performance based solely on ARL is insufficient. Instead, studying a chart's performance based on the percentiles of the run length distribution give more meaningful interpretation and is easier to understand.

TABLE II: Simulation Results for the Synthetic $\bar{X}$ Chart When $n = 3$, $L = 6$, $k = 2.29367$

δ	Percentage points of the run length distribution (%)											
	0.1	1	5	10	20	30	40	50	60	70	80	90
0.00	1	1	3	5	47	102	166	240	333	451	617	902
0.25	1	1	2	3	12	35	62	95	134	185	257	380
0.50	1	1	1	2	3	5	10	18	28	41	58	88
0.75	1	1	1	1	2	3	3	4	6	11	16	25
1.00	1	1	1	1	1	2	2	3	3	4	5	10
1.50	1	1	1	1	1	1	1	2	2	3	3	4
2.00	1	1	1	1	1	1	1	1	1	2	2	3
2.50	1	1	1	1	1	1	1	1	1	1	1	2
3.00	1	1	1	1	1	1	1	1	1	1	1	1

TABLE III: Simulation Results for the Synthetic $\bar{X}$ Chart When $n = 5$, $L = 4$, $k = 2.21855$

δ	Percentage points of the run length distribution (%)											
	0.1	1	5	10	20	30	40	50	60	70	80	90
0.00	1	1	2	4	54	108	170	244	334	451	615	895
0.25	1	1	1	3	8	24	43	65	91	126	174	257
0.50	1	1	1	1	2	3	4	9	13	20	29	44
0.75	1	1	1	1	1	2	2	2	3	4	7	11
1.00	1	1	1	1	1	1	1	1	2	2	3	4
1.50	1	1	1	1	1	1	1	1	1	1	1	2
2.00	1	1	1	1	1	1	1	1	1	1	1	1
2.50	1	1	1	1	1	1	1	1	1	1	1	1
3.00	1	1	1	1	1	1	1	1	1	1	1	1

TABLE IV: Simulation Results for the Synthetic $\bar{X}$ Chart When $n = 7$, $L = 3$, $k = 2.16382$

δ	Percentage points of the run length distribution (%)											
	0.1	1	5	10	20	30	40	50	60	70	80	90
0.00	1	1	2	10	57	110	171	244	333	450	613	890
0.25	1	1	1	2	6	18	31	47	67	92	128	189
0.50	1	1	1	1	2	2	3	5	8	12	17	27
0.75	1	1	1	1	1	1	1	2	2	3	3	7
1.00	1	1	1	1	1	1	1	1	1	2	2	2
1.50	1	1	1	1	1	1	1	1	1	1	1	1
2.00	1	1	1	1	1	1	1	1	1	1	1	1
2.50	1	1	1	1	1	1	1	1	1	1	1	1
3.00	1	1	1	1	1	1	1	1	1	1	1	1

TABLE V: Simulation Results for the Synthetic $\bar{X}$ Chart When $n = 10$, $L = 2$, $k = 2.08459$

δ	Percentage points of the run length distribution (%)											
	0.1	1	5	10	20	30	40	50	60	70	80	90
0.00	1	1	2	15	62	115	176	248	337	451	610	884
0.25	1	1	1	1	5	13	22	33	47	64	89	131
0.50	1	1	1	1	1	1	2	2	5	7	10	16
0.75	1	1	1	1	1	1	1	1	1	2	2	4
1.00	1	1	1	1	1	1	1	1	1	1	1	2
1.50	1	1	1	1	1	1	1	1	1	1	1	1
2.00	1	1	1	1	1	1	1	1	1	1	1	1
2.50	1	1	1	1	1	1	1	1	1	1	1	1
3.00	1	1	1	1	1	1	1	1	1	1	1	1

In addition, the lower percentage points of the run length distribution, for the in-control process, such as the 5th, 10th, and 20th percentiles provide information about the early false alarm rates. For example, based on Table II, we can conclude that there is a 10% chance that a false out-of-control signal will occur by the 5th sample point, given that the process is in-control. The higher percentage points of the run length distribution enable practitioners to state with a high probability, that an out-of-control will be signaled by a certain sample point, when the process shifts by a certain magnitude. From Table II, when a process shifts by a magnitude, say $\delta = 0.5$, one can state with a probability of 0.9 that this shift will be detected by the 88th sample point.

V. Conclusions

The results from the SAS program show that when the process is in-control, the in-control ARL falls between the 60th and 70th percentiles of the run length distribution, i.e. the difference between the in-control ARL and in-control median run length (MRL) (50th percentiles of the run length distribution for the in-control case) is somewhat large. However, as the magnitude of the shift increases, the out-of-control ARL becomes closer to the out-of-control MRL. This shows that the in-control run length distribution is more rightly skewed compared to the out-of-control run length distribution. Furthermore, the skewness of the run length distribution decreases with the magnitude of the shift. Thus, a more meaningful interpretation and analysis of the synthetic chart's performance can be made via a study of the percentage points of its run length distribution and not merely based on the ARL.

References

[1] P. D. Bourke, "Detecting a shift in fraction nonconforming using run-length control charts with 100% inspection," *Journal of Quality Technology*, vol. 23, no. 3, pp. 225-238, 1991.

[2] M. B. C. Khoo, H. C. Lee, Z. Wu, C. H. Chen, and P. Castagliola, "A synthetic double sampling control chart for the process mean," *IIE Transactions*, vol. 43, no. 1, pp. 23-38, 2011.

[3] F. Aparisi and M. A. De Luna, "Synthetic- $\bar{X}$ control charts optimized for in-control and out-of-control regions," *Computers & Operations Research*, vol. 36, no. 12, pp. 3204-3214, 2009.

[4] G. A. Barnard, "Control charts and stochastic processes," *Journal of the Royal Statistical Society, Series B (Methodological)*, vol. 21, no. 2, pp. 239-271, 1959.

[5] A. F. Bissell, "CUSUM techniques for quality control," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, vol. 18, no. 1, pp. 1-30, 1969.

[6] S. Chakraborti, "Run length, average run length and false alarm rate of Shewhart X-bar chart: exact derivations by conditioning," *Communications in Statistics–Simulation and Computation*, vol. 29, no. 1, pp. 61-81, 2000.

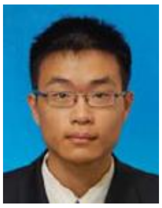
[7] A. C. Palm, "Tables of run length percentiles for determining the sensitivity of Shewhart control charts for averages with supplementary runs rules," *Journal of Quality Technology*, vol. 22, no. 4, pp. 289-298, 1990.

[8] Z. Wu and T. A. Spedding, "A synthetic control chart for detecting small shifts in the process mean," *Journal of Quality Technology*, vol. 32, no. 1, pp. 32-38, 2000.

[9] A. F. B. Costa and M. A. Rahim, "A synthetic control chart for monitoring the process mean and variance," *Journal of Quality in Maintenance Engineering*, vol. 12, no. 1, pp. 81-88, 2006.

[10] P. Castagliola and M. B. C. Khoo, "A synthetic scaled weighted variance control chart for monitoring the process mean of skewed populations," *Communications in Statistics–Simulation and Computation*, vol. 38, no. 8, pp. 1659-1674, 2009.

[11] R. B. Davis and W. H. Woodall, "Evaluating and improving the synthetic control chart," *Journal of Quality Technology*, vol. 34, no. 2, pp. 200-208, 2002.



Z. L. Chong is a full time Ph.D. student in the School of Mathematical Sciences, Universiti Sains Malaysia (USM). He obtained his bachelor of science with education (Honours) in mathematics [distinction] from USM in 2012. His research interest is statistical process control, focusing on quality control charts.

Michael B. C. Khoo is a professor in the School of Mathematical Sciences, Universiti Sains Malaysia (USM). He specializes in statistical process control. He publishes extensively in international journals. He also serves as a member of the editorial boards of several International journals. His research interest is statistical quality control, focusing on quality control charts. He is a member of the American Society for Quality (ASQ) and a life member of the Malaysian Mathematical Sciences Society.

H. W. You is a Ph.D. student in the School of Mathematical Sciences, Universiti Sains Malaysia (USM). She holds a bachelor of applied science degree in applied statistics from USM. Her research interest is statistical process control.