NON – PARAMETRIC CONTROL CHARTS BASED ON MEDIAN RANKED SET SAMPLING

[1]WALIPORN TAPANG, [2]ADISAK PONGPULLPONSAK

[1,2]Faculty of Science, King Mongkut's University of Technology Thonburi, Thailand
E-mail: [1]walitapang@hotmail.com, [2]adisak.pon@kmutt.ac.th

Abstract— In this paper, we study the average run length (*ARL*) performance of the two non – parametric control charts (Mann-Whitney control chart (U – Chart) and Wilcoxon rank sum test control chart (W – Chart)).The data are generated from normal distribution, uniform distribution, and Weibull distributions with the process mean shift in δ times of standard deviation (δ = 0.50, 1, 1.5, 2, 2.5, and 3). The results show that *ARL* of non – parametric control chart base on median ranked set sampling (MRSS) has superior performance than the simple random sampling.

Index Terms— Average Run Length, Median Ranked Set Sampling, Non – Parametric Control Chart

I. INTRODUCTION

The statistical quality control (SQC) methods used to help improving the quality of products by identifying and possibly reducing any assignable sources of variation that might be present.For this methods, the distribution of chance causes is often assumed to follow some parametric distribution, we cancall this parametric SQC. This method includes the cases ranging from where the chance distribution is completely that is unknown to where it may be partly known but it is far from being normal. Moreover, there is a need for some flexible and robust control charts that do not require normality or any other specific distributional assumption. Non-parametric control charts can be served this broader purpose and can thus offering the useful alternatives to apply the parametric control charts.

There is an increasing knowledge about non-parametric control charts in statistical process control and monitoring. In 1979, Bakir and Reynolds [1] proposed the cumulative sum chart (CUSUM).McDonald [2] also considered a CUSUM procedure for individual observations.Hackel and Ledolter[3] constructed the standardized ranks of observations and defined by the in-control distributions. In utilization of the EWMA chart, if there is no data available. Woodall and Montgomery [4] foresaw an increasing role for non-parametric methods in the control charts application. Chakraborti et al. [5] gave an overview and discussed the advantages of several non-parametric control charts over their normal theory counterparts.Bakir [6] complied and classified several non-parametric control charts according to the driving non-parametric idea behind each one of them. In 2004,Chakraborti and Van de Wiel [7] developed new non-parametric control charts based on the Mann–Whitney(U – Test) statistic. In 2014,Jayathavaj and Pongpullponsak [8]compared the performances of the three dual – scheme variable parameters (VP) non-parametric (NP) control charts.

There are many literatures that is establishedthese methods are based on simple random sampling (SRS).In 1952,the ranked set sampling (RSS) was offered by McIntyre [9]. Hall and Dell [10] observed that using RSS the samples can be ranked more efficient than using SRS.Muttlak and Al – Sabah [11] developed the quality control charts for finding the mean of population.Pongpullponsak and Sontisamran[12] created the quality control charts model for high variance statistical process, in order to filter and reduce the variation of product's quality.

For this study, we establish two non-parametric control charts based on median ranked set sampling MRSS viz. the Wilcoxon rank sum charts (W – Chart) and the Mann–Whitney chart (U – Chart). The data is generated from normal distribution, uniform distribution, and Weibull distributions using the process shift in δ times of standard deviation (δ=0.50, 1.0, 1.5, 2.0, 2.5, and 3.0) that are selected for using in the study. The performance of the established two non-parametric control charts are the average run length.

The paper is organized as follows; Section 2 median ranked set sampling, Section 3 two non- parametric control charts based on MRSS, Section 4 performance comparisons from simulation results and Section 5 conclusions.

II. MEDIAN RANKED SET SAMPLING

Median ranked set sampling (MRSS) was proposedby Muttlak[13]. Using this method, the sample at the median of the sets is selected, if the set size n is odd. In case of the set size with even number, sample selection is from the $(n/2)^{th}$order in the first half and the $((n+2)/2)^{th}$ order in the second half of the set. The method of MRSS can be concluded as follows.

Step 1 select *n* sample units per set from the total *n* sample sets

Step 2 allocate sample units into the set by using a variable related to a variable of interest in ranking

Step 3 choose the sample units for actual measurement by selecting the smallest rank in the $((n+1)/2)^{th}$ orderfrom the sample sizes with an odd number. For the sample sets with an even number, the smallest rank in the $(n/2)^{th}$ order of the first half and the smallest rank in the $((n+2)/2)^{th}$ order in the second half are chosen.

Step 4 repeat steps 1 through 3 for *r* cycles until the desired sample size is obtained (s = nr)

Seen in Figure I, samples are randomly selected for 3 sets, where each set are contained 3 sample units, then repeated this procedure for 4 times.

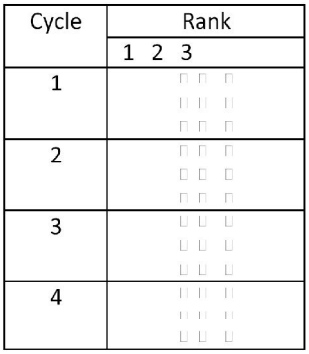

Fig.I The MRSS sample units for case I

In Figure I, it contains ranked sample units for each row, and only the sample units marked as that is chosen for actual measurement. Therefore, from the total 36 sample units which are randomly selected for 4 cycles, only 12 units will be used for measuring a variable of interest.

As seen in Figure II, the samples are randomly selected for 4 sets, where each set are contained 4 sample units, then repeated this procedure for 4 times.

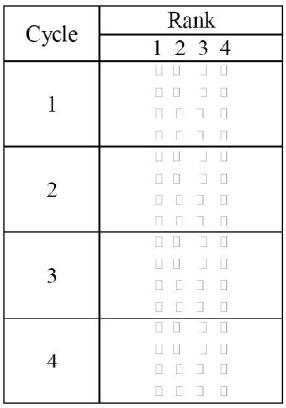

Fig. IIThe MRSS sample units for case II

In the illustration above, it containsranked sample units for each row, and only the sample units is marked as that is chosen for actual measurement. Therefore, from the total 36 sample units which are randomly selected for 4 cycles, only 16 units will be used for measuring a variable of interest.

For the sample sizes with an odd number, given $X_{(i:m)j}$ is the (n/2) order statistic of the ith order from sample size n in the jth cycle.In case of an even number, $X_{(i:m)j}$ is the $(n/2)^{th}$ order statistic of the ith order from sample size n $(i=1,2,...,L)$ where $L=n/2$,andthe $((n+2)/2)^{th}$ order statistic of the ith order from sample size n $(i=L+1,L+2,...,n)$,where Muttlak [13] proposed an estimator of population mean for MRSS as below;

$$\overline{X}_{mrss,j}=\frac{1}{n}\sum_{i=1}^{n}X_{(i:m)j} \quad ;j=1,2,...,r \quad \cdots(1)$$

The variance of $\overline{X}_{mrss,j}$ is defined by

$$\text{var}\left(\overline{X}_{mrss,j}\right)=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{(i:m)} \quad \cdots \quad (2)$$

III. NON – PARAMETRIC CONTROL CHART

A. Man Whitney control chart(U – chart) based on MRSS

Step 1: Random reference sample of size m by MRSS, denoted by $X=\left(\overline{X}_{mrss,1},\overline{X}_{mrss,2},...,\overline{X}_{mrss,m}\right)$, is available in an in – control process.

Step 2: Random test sample of size n by MRSS, denoted by $Y=\left(\overline{Y}_{mrss,1},\overline{Y}_{mrss,2},...,\overline{Y}_{mrss,n}\right)$ is available in an in – control process.

Step 3: The superscript t is used to denote the tth test sample, $Y^t=\left(\overline{Y}_{mrss,1},\overline{Y}_{mrss,2,...},\overline{Y}_{mrss,n}\right)$, t = 1,2,...

Step 4: The U – test is constructed from using the total number of X–Y pairs where the Y observation is larger than that of the X. This can be written as

$$U=\sum_{i=1}^{m}\sum_{j=1}^{n}D_{ij} \text{ where}$$

$$D_{ij}=\begin{cases}1 & \text{if } \overline{X}_{mrss,i}<\overline{Y}_{mrss,j} \text{ for all } j=1,2,...,n\\ 0 & \text{if } \overline{X}_{mrss,i}>\overline{Y}_{mrss,j} \text{ for all } i=1,2,...,m\end{cases}$$

At this stage, it can be assumed that the Wilcoxon rank –

sum test is equivalent to the U - test by the relationship below;

$$U = W_n - \frac{n(n+1)}{2}$$

where $W_n = \sum_{j=1}^{n} R_j$

and $R_1,\ldots, R_n$ are the rank of the n observations $\overline{Y}_{mrss,1}, \overline{Y}_{mrss,2}, \ldots, \overline{Y}_{mrss,n}$ in the complete sample of m + n observations.

Step 5: Calculation of the $E(U)$ *and* $Var(U)$ is carried out as following.

The expectation and variance of U are given by

$$E(U) = \frac{mn}{2} \quad and \quad Var(U) = \frac{mn(m+n+1)}{12}$$

But in this study, the sample numbers are large so the standardized U- Test is defined by

$$Z = \frac{U - E(U)}{\sqrt{Var(U)}}$$

Step 6: The control chart based on Z following the usual Shewhart scheme is designed. The control limits are given by

$$UCL = 3,\ CL = 0,\ LCL = -3$$

Step 7: The Z values is plotted in the control chart. If any point goes beyond the limit it will be indicated that the process is out of control with respect to variability. For each pair of consecutive samples of size m + n, the realization z of the test statistic Z is calculated. If z < -3 or z> 3, an alarm will be triggered and a search for an assignable cause will be undertaken.

B. Wilcoxon rank sum test control chart (W – chart) based on MRSS

Step 1: Random reference sample of size m by MRSS, denoted by $B = \left(\overline{X}_{mrss,1}, \overline{X}_{mrss,2}, \ldots, \overline{X}_{mrss,m}\right)$, is available in an in – control process.

Step 2: Random test sample of size n by MRSS, denoted by $A = \left(\overline{Y}_{mrss,1}, \overline{Y}_{mrss,2}, \ldots, \overline{Y}_{mrss,n}\right)$ is available in an in – control process.

Step 3: The superscript t is used to denote the tth test sample, $A^t = \left(\overline{Y}_{mrss,1}, \overline{Y}_{mrss,2}, \ldots, \overline{Y}_{mrss,n}\right)$, t = 1,2,...

Step 4: The W – test is constructed from using the total number of X +Y where the Y observation is larger than that of the X. This can be written as

$$W = W_n - \frac{n(n+1)}{2}$$

where $W_n = \sum_{i=1}^{n} R_i$ and $R_1,\ldots,R_n$ are the rank of the n observations $\overline{Y}_{mrss,1}, \overline{Y}_{mrss,2}, \ldots, \overline{Y}_{mrss,n}$ in the complete sample of m + n observations.

Step 5: Calculation of the $E(W)$ *and* $Var(W)$ is carried out as below.

The expectation and variance of W are given by

$$E(W) = \frac{n(n+m+1)}{2} \quad and \quad Var(W) = \frac{mn(m+n+1)}{12}$$

But in this study, the sample numbers are large so the standardized W – test is defined by

$$Z = \frac{W - E(W)}{\sqrt{Var(W)}}$$

Step 6: The control chart based on Z following the usual Shewhart scheme is designed. The control limits are given by

$$UCL = 3,\ CL = 0,\ LCL = -3$$

Step 7: The Z values is plotted in the control chart. If any point goes beyond the limit, it will be indicated that the process is out of control with respect to variability.

For each pair of consecutive samples of size m + n, the realization z of the test statistic Z is calculated. If z< -3 or z > 3, an alarm will be triggered and a search for an assignable cause will be undertaken.

IV. PERFORMANCE OF U-CHART AND W - CHART

In evaluating the ARL performance of the two non-parametric control charts, the 30,000 runs ARL simulation from the Uniform, the standard Normal, and the Weibull distributions with sample sizes n =10, 15, and 20and the mean shift in δ time of standard deviation from 0.0 to 3.0 steps by 0.5 are used. All the programs are written in R.

A. Man Whitney control chart (U – chart) based on MRSS

From the Mann-Whitney control charts with sample sizes from the control group in which m = 10, 15, 20 and the treatment group at n = 10, 15, 20 determined the control limits with α is closed to 0.0027, denoted by –3 and 3 for action limits shown in TableI.

Table I Average run length (ARL) of the Mann-Whitney control chart (U – Chart) with Normal distribution, Uniform distribution, and Weibull distribution

Distribution	Normal			Uniform			Weibull (beta = 3.2219)		
Sample size	10	15	20	10	15	20	10	15	20
0.0	566.67	454.55	526.32	614.03	492.61	431.03	769.23	416.67	476.19
0.5	384.62	312.50	277.78	398.41	290.70	304.88	454.55	322.58	400.00
1.0	156.25	131.58	151.52	169.78	143.06	143.06	161.29	126.58	151.52
1.5	74.07	60.24	58.82	70.37	58.24	60.83	78.13	58.82	52.08
2.0	30.21	30.67	32.26	30.64	28.89	31.45	28.57	30.4	29.94
2.5	15.82	17.21	16.86	16.38	17.04	16.80	15.8	17.36	16.92
3.0	9.13	9.85	9.63	9.13	9.77	9.47	9.32	9.85	9.48

The performance of Mann – Whitney control chart (U – Chart) based on MRSS for normal, uniform, and Weibull distributions with the sample size = 10, 15 and 20 are compared as show in Table I. The highest ARLs of the sample size 10 is 769.23 from Weibull distribution, the sample size 15 is 492 from uniform distribution and the sample size 20 is 526.32 form normal distribution.

C. Wilcoxon rank sum test control chart (W – chart) based on MRSS

From the Wilcoxon ranked sum test control charts with sample sizes from the control group at m = 10, 15, 20 and the treatment group at n = 10, 15, 20 determined control limits with α is closed to 0.0027, denotedby –3 and 3 for action limits, and ARL of selected schemes for standard normal data,uniform data, and Weibull data are shown in Table II.

Table II Average run length (ARL) of the Wilcoxon rank sum test control chart (W – Chart) with Normal distribution, Uniform distribution, and Weibull distribution

Distribution	Normal			Uniform			Weibull (beta = 3.2219)		
Sample size	10	15	20	10	15	20	10	15	20
0.0	526.32	588.24	454.55	588.24	633.33	526.32	555.56	400.00	384.62
0.5	303.03	333.33	270.27	434.78	434.78	400.00	333.33	370.37	357.14
1.0	172.41	123.46	131.58	169.49	126.58	128.21	172.42	125.6	125.00
1.5	62.11	64.52	70.92	80.00	52.08	62.50	68.96	63.29	58.14
2.0	30.21	30.67	31.35	37.17	28.57	33.22	32.68	28.33	33.67
2.5	17.67	17.36	16.29	16.10	17.57	16.29	15.06	16.58	16.34
3.0	9.02	10.03	9.52	9.28	9.72	9.13	9.26	9.99	8.56

The performance of W – Chartbased on RSS for normal, uniform, and Weibull distributions with the sample size 10, 15 and 20 are compared as show in Table II. The highest ARLs of all sample size from uniform distribution, the sample size 10 is 588.24, the sample size 15 is 633.33 and the sample size 20 is526.32.

CONCLUSION

The average run lengths of the two non – parametric control chart are high performance because the median ranked sampling (MRSS) forms the new sample by choosing the observations from the samples that are arranged the observations in ordered, this method bring out the real process characteristics better than simple random sampling (SRS). The performance from simulation study with n=10, 15 and 20 also confirmed that the ARLs of non-normal distributions are higher than normal in every sample size.

ACKNOWLEDGEMENT

This work was partially supported by the Higher Education Research Promotion and National

Research University Project of Thailand, Office of the Higher Education Commission. Also, the first author would like to thank the Office of the Higher Education Commission, Thailand, for the financial support of the Ph.D. Program at KMUTT.

REFERENCES

[1] S.T. Bakir and M.R. Reynolds, "A non Parametric Procedure for Process Control Based onwithin Group Ranking",Technometrics, Vol.21, pp. 175 – 183, 1979

[2] D. McDonald, "A CUSUM procedure based on sequential ranks", NavalResearch Logistics, Vol. 37, pp. 627–646, 1990.

[3] P. Hackl and J. Ledolter, "A control chart based on ranks", Journal ofQuality Technology, Vol. 23, pp. 117–126, 1991.

[4] W.H. Woodall and D.C. Montgomery,"Research issues and ideas in statistical process control",Journal ofQuality Technology, Vol. 31, pp. 376–386, 1999.

[5] S. Chakraborti et al., "Nonparametric Control Charts: An Overview and Some Results", Journal ofQuality Technology, Vol. 33, pp. 304-315, 2001.

[6] S.T.Bakir, "Classification of distribution free control charts", Proceedings of annual meeting of American Statistical Association (Section Quality and Productivity), August 2001; 5-9.

[7] S. Chakarborti and M.A.Van de Wiel, "A nonparametric control chart based on the Mann-Whitney stistic",Available at: http://www.win.tue.nl/bs/spor/2003-24.pdf

[8] V. Jayathavaj and A. Pongpullponsak, "Comparative Performances of the Variable Parameters Nonparametric Control Charts using the Markov Chain Approach", Chiang Mai Journal ofScience, Vol. 41, No. 5.2, pp. 1457 – 1472, 2014.

[9] G. A. McIntyre, "A method for unbiased selective sampling using ranked sets", Australian Journal of Agricultural Research, Vol. 3, pp. 385-390, 1952.

[10] L.K. Halls and T.R. Dell, "Trial of ranked set sampling for forage yields", Forest Science, Vol. 12, pp. 22 – 26, 1966.

[11] H.A. Muttlak and W.S. Al-Sabah, "Statistical quality control based on ranked set sampling", Journal ofApplied Statistics, Vol. 30, pp. 1055 – 1078,2003.

[12] A. Pongpullponsak and P. Sontisamran, "Statistical quality control based on Ranked Set Sampling for Multiple Characteristics", Chiang Mai Journal ofScience, Vol. 40, No. 3, pp. 485 – 498, 2013.

[13] H.A. Muttlak, "Median ranked set sampling", Journal ofApplied Statistics, Vol. 6, pp. 245-255,1997.

★★★