Chiang Mai J. Sci. 2016; 43(4) : 914-929
http://epg.science.cmu.ac.th/ejournal/
Contributed Paper

Three Non-parametric Control Charts Based on Ranked Set Sampling

Waliporn Tapang, Adisak Pongpullponsak* and Sukuman Sarikavanij
Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand.
*Author for correspondence; e-mail: adisak.pon@kmutt.ac.th

Received: 24 May 2015
Accepted: 19 October 2015

ABSTRACT

In this study, the ranked set sampling (RSS) is applied to the three nonparametric control charts: the Mann-Whitney statistics (U-chart), the Wilcoxon rank sum test (W-chart) and the Hodges-Lehmann estimator (HL-chart) in order to reduce bias in sampling procedure. The data are generated from normal distribution, uniform distribution and 11 shapes of 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 the in control average run length (ARL) of the three control charts with RSS are higher than the Shewhart X bar chart at +/-3 standard deviation control limits. When the shift occurs, the out of control ARL of HL-chart is the lowest in every mean shift.

Keywords: ranked set sampling, nonparametric control charts, wilcoxon rank sum test, mann-whitney statistics, hodges-lehmann estimator

1. INTRODUCTION

Quality control charts are extensively used in observing and testing a production process. The control via the charts involves monitoring process changes and identifying abnormalities in the process. This makes possible the diagnosis for some problems in the production, losses reduction and, substantial improvements in product quality. However, quality control charts literally are based on simple random sampling (SRS) in quality control which may not be satisfied since it gives samples with high variation resulting in wide control limits, thereby low quality control efficiency.

In addition to SRS, quality control charts has also been developed by using ranked set sampling. Ranked set sampling (RSS) was proposed by McIntyre in 1952[1]. The development of RSS is believed to be able to solve the problem of high cost in sampling or taking long time in measuring the true. In 1966, Hall and Dell [2] observed that using RSS the samples can be ranking more efficient than using SRS when they are measured at the actual sample size in the same situation. Besides, the quality control charts with RSS yields higher efficiency compared with the actual measure under the same sample size because the control limits are narrower that are leading to better quality control. Muttlak and Al-Sabah [3] developed quality control charts for finding the mean of population by comparing the mean numbers where the values fall within the control limits before falling out of average run length (ARL) between perfect RSS, imperfect RSS and SRS.

Using RSS, Pongpullponsak and Sontisamran[4] created a quality control charts model for high variance statistical process, in order to filter and reduce the variation of product's quality. In that report, the variety of product quality was ranked by using multiple characteristics.

Most of the methods that have been reported in many literatures are often assumed to have behaviors following to normal distribution. This brought into some limitations in these methodologies. The most serious limitation is that the methods have been constructed based on asymptotic properties where it is not true when item sizes are large. A more practical approach was introduced by Chen and Bai [5] in parametric settings who considered that unbalanced RSS would fix the problem. In this kind of situation, development and application of control charts are not depended on normality and that is why nonparametric control charts can serve this purpose.

The main advantage of nonparametric control charts is flexibility. They have been derived from the idea that parametric probability distribution for the underlying process does not need to be considered; at least as far as establishing and implementing the charts are concerned. Obviously, this is very beneficial in the field of process control, particularly start-up situations where not much data is available for a parametric (such as normal theory) procedure. Also, the nonparametric charts are likely to share the robustness properties of nonparametric tests at confidence intervals and are, therefore, far more likely to be less impacted by outliers.

There is increasing knowledge about nonparametric control charts in statistical process control and monitoring. In 1979, Bakir and Reynolds [6] proposed the cumulative sum chart (CUSUM), which was established on the basis of the Wilcoxon signed-rank statistic. McDonald [7] considered a CUSUM procedure for individual observations, which was further adapted from the statistic called sequential ranks. Another nonparametric control chart type is exponentially-weighted moving average (EWMA) chart for individual observations. Hackel and Ledolter[8] 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, the authors recommended to employ the ranking on previously collected reference data. However, the in-control average run length or IC ARL could be substantially large if the unknown in-control parameters must be estimated. Later, Woodall and Montgomery [9] foresaw an increasing role for nonparametric methods in control charts application. Chakraborti et al. [10] gave an overview and discussed the advantages of several nonparametric control charts over their normal theory counterparts.In the same year, Bakir [11] complied and classified several nonparametric control charts according to the driving nonparametric idea behind each one of them. In 2004, Chakraborti and Van de Wiel [12] developed new nonparametric control charts based on the well-known Mann–Whitney (U-Test) statistic. It was proposed for process location that maintains its in-control properties and thus could be justifiably used for any continuous process distribution. Also, control limits of the proposed charts are provided for practical implementation. Comparisons of some performance criteria related to the run length distribution show that while the proposed chart has clearly superior and stable in-control run length properties, the chart is nearly as effective in detecting shifts as the Shewhart $\overline{X}$ chart when the process is normal, but is more effective than the Shewhart $\overline{X}$ chart for a heavy-tailed distribution such as the Laplace and for a skewed distribution such as the Gamma (2,2). The 5th percentile of the (conditional) ARL0(X) distribution can be a useful chart design criterion, in addition to the traditional (unconditional) ARL0.

Many statisticians considered the U-test, established by Gibbons and Chakraborti [13], as the best nonparametric test for location of process. The U-test is a direct competitor to the normal theory based on two-independent-sample t-tests. Remarkably, even when the underlying distributions are normal, the U-test is about 96% as efficient as t-test for moderately large sample sizes and yet, unlike the t-test, it does not require normality to be valid. Moreover, for some heavy-tailed distributions such as the Laplace (double exponential), the logistic distribution, or skewed distributions such as the exponential, the U-test is known to be more efficient than the t-test. In short, the U-test is the practitioners' choice when not much is known about the shape of the underlying distributions. In 2014, Jayathavaj and Pongpullponsak[14] compared performances of the three dual-scheme variable parameters (VP) nonparametric (NP) control charts: the Sign Test (ST), the Mann-Whitney Test (MW) and the Hodges-Lehmann estimator (HL) using the Markov chain approach. The results show that among the three VP (n_1 = 10, n_2 = 10, 11, …, 20), ST has more variation in the average run lengths (ARL) than MW and HL. In each NP control chart, the average number of samples to signal lies between the ARL of their individual schemes, ST has highest variation while MW and HL have varied in a smaller band and HL has better performances than MW in almost every sample size.

For this study, we establish three non-parametric control charts based on RSS viz. the Wilcoxon rank sum charts (W-Chart), the Mann–Whitney chart (U-Chart) and the Hodges–Lehmann estimator chart (HL-Chart). The data are generated from normal distribution and uniform distribution and 11 shapes of Weibull distributions with the process shift in δ times of standard deviation (δ=0.50, 1.0, 1.5, 2.0, 2.5 and 3.0) are selected for using in the study.

The paper is organized as follows; Section 2 industrial application, Section 3 three non- parametric control charts based on RSS, Section 4 performance comparisons from simulation results and Section 5 conclusions.

2. Industrial Application of Control Charts Based on Ranked Set Sampling and Non-Parametric Control Charts

Industrial processes are generally monitored by using Statistical Process Control (SPC). Generally, SPC monitors a process to determine whether it is operated under statistical control or not. The good example is evaluation of average dried weight of leaves by Ridout[15]. Since it is very difficult to measure the precise dry weight of leaves, Ridout suggested that the spray deposits of water on both sides and the upper and lower surfaces of leaves, can be used in estimation. Also, Ridout [15] employed characteristics of interest in RSS from the products containing multiple characteristics in order to reduce error from ranking and increase efficiency of sampling. In his report, it was found that RSS is not appropriate for controlling production of goods with multiple characteristics. However, in 2013 we (Pongpullponsak and Sontisamran [4]) constructed statistical quality control based on ranked set sampling for multiple characteristics (RSSMC). Importantly, when compared with RSS, SRS, MRSS and ERSS approaches, our control chart using RSSMC demonstrated more efficient and satisfy than those of control charts. Alloway and Raghavachari[16] successfully applied their technique to primer thickness data that was taken from Ford Motor Company. Albers et al. [17] showed an application of nonparametric control chart for thickness of electric shaver razor head produced by an electrochemical process. Hackl and Ledolter[8] applied their procedure on weights of tomato cans. Pyzdek[18] revealed that the thickness of alloy layers which was produced by a hot deep galvanizing process do not follow a normal distribution. Therefore, application of SPC technique assumed to follow normal distribution can lead to a misleading conclusion. For the sake of industrial situation, it is essential to check the normality of the data whether it is following the typical bell-shaped normal distribution or not, if not, one should not go for conventional Shewhart-type control chart. The beauty of a non-parametric control chart is that it does not require any distributional assumption. In fact, through simulation study we have revealed that non-parametric control charts are also useful for normal distribution.

3. Nonparametric Control Charts Based on Ranked Set Sampling

3.1 Ranked Set Sampling Method

Ranked set sampling (RSS) has been proposed by McIntyre [1]. The samples obtained by this method will be ranked using other variables that relate to the variable of interest or the variable to be actual measurement. The steps in random ranked set sampling are described as below;

Step 1: Randomly select n^2 sample units from the population

Step 2: Allocate the n^2 selected units as randomly as possible into m sets, each of size n

Step 3: Without yet knowing any values for the variable of interest, rank the units within each set based on a perception of relative values for this variable, which may be based on personal judgment or done with measurements of a covariate that is correlated with the variable of interest

Step 4: Choose a sample for actual analysis by including the smallest ranked unit in the first set, then the second smallest ranked unit in the second set, continuing in this fashion until the largest ranked unit is selected in the last set

Step 5: Repeat steps 1 through 4 for r cycles until the desired sample size, N = nr, is obtained for analysis

To explain more for this method, assuming that 3 sample sets are random sampling to collect 3 samples/set and the sampling are repeated 4 cycles $r = 4$. This can be concluded as shown in Figures 1-2;

Cycle	Rank 1	Rank 2	Rank 3
1	⊙	·	·
	·	⊙	·
	·	·	⊙
2	⊙	·	·
	·	⊙	·
	·	·	⊙
3	⊙	·	·
	·	⊙	·
	·	·	⊙
4	⊙	·	·
	·	⊙	·
	·	·	⊙

Figure 1. Sample units for RSS.

Cycle 1

$$X_{(1:3)1} \leq X_{(2:3)1} \leq X_{(3:3)1} \rightarrow X_{(1:3)1}$$
$$X_{(1:3)1} \leq X_{(2:3)1} \leq X_{(3:3)1} \rightarrow X_{(2:3)1}$$
$$X_{(1:3)1} \leq X_{(2:3)1} \leq X_{(3:3)1} \rightarrow X_{(3:3)1}$$

Cycle 2

$$X_{(1:3)2} \leq X_{(2:3)2} \leq X_{(3:3)2} \rightarrow X_{(1:3)2}$$
$$X_{(1:3)2} \leq X_{(2:3)1} \leq X_{(3:3)2} \rightarrow X_{(2:3)2}$$
$$X_{(1:3)2} \leq X_{(2:3)1} \leq X_{(3:3)2} \rightarrow X_{(3:3)2}$$

… … … … … … …

Cycle r

$$X_{(1:3)r} \leq X_{(2:3)r} \leq X_{(3:3)r} \rightarrow X_{(1:3)r}$$
$$X_{(1:3)r} \leq X_{(2:3)r} \leq X_{(3:3)r} \rightarrow X_{(2:3)r}$$
$$X_{(1:3)r} \leq X_{(2:3)r} \leq X_{(3:3)r} \rightarrow X_{(3:3)r}$$

Figure 2. Ranked sample units for RSS.

From Figure 1, a ranked set sample design with set size n = 3 and number of sampling cycles r = 4 is demonstrated. Although 36 sample units have been selected from the population, only 12 circled units are actually included in the final sample for quantitative analysis. For n = 3, the cycle then repeats r times, the sampling procedure is illustrated in Figure 2.

Let $X_{(i:n)j}$ denote the ith order statistic from the ith sample of size n in the jth cycle, then the unbiased estimator for the population mean, see Takahasi and Wakimoto[19], is defined as

$$\overline{X}_{rss,j} = \frac{1}{nr}\sum_{j=1}^{r}\sum_{i=1}^{n} X_{(i:n)j}$$

The variance of $\overline{X}_{rss,j}$ is given

by $Var\left(\overline{X}_{rss,j}\right) = \frac{1}{nr}\sum_{i=1}^{n}\sigma^2_{(i:n)}$, where

$\sigma^2_{(i:n)} = E\left[X_{(i:n)} - E\left(X_{(i:n)}\right)\right]^2$ is the population variance of the ith order statistic.

3.2 The Mann-Whitney Statistic (U-test)

The Mann-Whitney test is the most popular nonparametric in the statistical practice. This test is defined as the number of times that a $Y = (Y_1, Y_2, \ldots, Y_n)$ and $X = (X_1, X_2, \ldots, X_m)$ are preceded in the combined ordered arrangement of two independent random samples and the same distribution. We denoted by $Y = (Y_1, Y_2, \ldots, Y_n)$ is available from an in-control process sample of size n and $X = (X_1, X_2, \ldots, X_m)$ denotes an independent random tested sample of size m. Into a single sequence of $m + n = N$, so the possibility of $X_i = Y_j$ for some (i, j) does not need to be considered. If the mn indicator random variables are defined as

$$D_{ij} = \begin{cases} 1 & if\ X_i < Y_j\ \ for\ all\ j = 1,2,\ldots,n \\ 0 & if\ X_i > Y_j\ \ for\ all\ i = 1,2,\ldots,m \end{cases}$$

then, a symbolic representation of the Mann-Whitney U statistic is presented by

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

Since U is defined as a linear combination of these mn random variables, the mean and variance of U can be written as following

$$E(U) = \sum_{i=1}^{m}\sum_{j=1}^{n} E\left(D_{ij}\right) = \frac{mn}{2}$$

$$\mathrm{var}(U) = \frac{mn(N+1)}{12}$$

3.3 The Wilcoxon Rank Sum Test (W-test)

The Wilcoxon rank sum test is a nonparametric alternative to the two sample t-test which is based solely on the order in which the observations from the two samples fall. It does not need to concern whether populations have normal distribution or not. Suppose, that we have samples of observations from each of two populations X and Y containing n_X and n_Y observations, respectively. And both populations contain two conditions which have equal variances and are independence of random samples. The two population distributions are assumed to be identical under the null hypothesis, in which independent random samples from the two populations should be similar if the null hypothesis is true.

Because we are now allowing the population distributions to non-normal, the rank sum procedure must deal with the possibility of extreme observations in data. We combined samples from both populations, one way to samples containing extreme values is to replace each data value with its rank from lowest to highest. The smallest value in the combined sample is assigned the rank of 1 and the largest value is assigned the rank of $N = n_X + n_Y$.

The calculation of the rank sum statistic consists of the following step.

Step 1: List the data values for both samples from smallest to largest

Step 2: Assign the numbers 1 to N to the data values with 1 to the smallest value and N to the largest values, which are the ranks of the observations

Step 3: If there are ties, duplicated values, in the combined data set, the ranks for the observations in a tie will be taken to be the average of the ranks for those observations.

Step 4: Let W denote the sum of the ranks for the observations form population X.

Under the null hypothesis, the sampling distribution of W has mean and variance can be given by

$$\mu_w = \frac{n_X\left(n_X + n_Y + 1\right)}{2} \text{ and}$$

$$\sigma_W = \sqrt{\frac{n_X n_Y\left(n_X + n_Y + 1\right)}{12}}$$

The hypothesis of wilcoxon rank sum test is

H_0: The two populations are identical.

H_1: 1. Population X is shifted to the right of population Y.

2. Population X is shifted to the left of population Y.

3. Population X and Y are shifted from each other.

Case 1: $n_X \leq 10$, $n_Y \leq 10$ and W is the sum of ranks in sample X. For $\alpha = 0.05$, the table of wilcoxon will be used to find critical values for W_u and W_l;

1. Reject H_0 if $W > W_u$
2. Reject H_0 if $W < W_l$
3. Reject H_0 if $W > W_u$ or $W < W_l$

Case 2: $n_X > 10$, $n_Y > 10$ W is Normal approximation for larger samples, we can treat the distribution of W_X as if it is/are normal $\left(\mu_X, \sigma_X\right)$, where

$$z = \frac{w_X - \mu_X}{\sigma_X} \text{ and } z = \text{Normal } (0, 1).$$

1. Reject H_0 if $z \geq z_\alpha$
2. Reject H_0 if $z \leq -z_\alpha$
3. Reject H_0 if $\geq z_{\alpha/2}$

3.4 The Hodges-Lehmann Estimator

In 1963, Hodges and Lehmann [20] proposed Hodges-Lehmann Estimator (HL.E.) as an estimator for the point of symmetry θ of a continuous and symmetric distribution. Initially, the HL.E. is a nonparametric estimator based on the Wilcoxon signed-rank statistic. However, Lehmann [21] showed later that this estimator belongs to the class of robust R-estimators.

The computations of the HL.E. follow four steps.

Step 1: Let $Y_1, Y_2, \ldots, Y_n$ be a random sample obtained from some distributions, which is continuous and symmetrical about θ. Then, compute

$$M = \frac{n(n+1)}{2}$$

Step 2: Compute the Walsh averages,

$$W_r = \frac{Y_i + Y_j}{2}$$

where $r = 1, 2, \ldots, M$ and $i \leq j = 1, 2, \ldots, n$

Step 3: Reorder the Walsh averages in ascending order, that is $W_{(1)} \leq W_{(2)} \leq \ldots \leq W_{(M)}$

Step 4: The HL.E. for the point of symmetry θ of a continuous and symmetric distribution is defined as:

$HL.E. = \text{median } \{ W_{(1)}, W_{(2)}, \ldots, W_{(M)} \}$

or

$$HL.E. = \begin{cases} W_{(k+1)} & \text{if } M \text{ is odd} \\ \dfrac{W_k + W_{k+1}}{2} & \text{if } M \text{ is even} \end{cases}$$

where

$$k = \begin{cases} (M-1)/2 & \text{if } M \text{ is odd} \\ M/2 & \text{if } M \text{ is even} \end{cases}$$

In addition to the above computations of the HL.E., the main properties of this estimator are given below:

(i) The asymptotic relative efficiency of the HL.E. relative to the sample mean is 0.955, if the underlying distribution is Normal (Gaussion). However, the asymptotic relative efficiency of the HL.E. is often greater than unity, if the underlying distribution is non-normal. Alloway and Raghavachari [16] stated that the asymptotic properties of the HL.E. are impressive,

(ii) The asymptotic relative efficiency for the HL.E. is the same as the Wilcoxon is signed 1 rank test and it is asymptotically normally distributed. Besides, it is robust against gross errors.

(iii) The HL.E. is unbiased and translation invariant. Also, it gives reasonable results for distributions in the neighborhood of the Normal (Gaussion) distribution. Additionally, Alloway and Raghavachari [16] mentioned that the performances of robust estimators are often better than traditional measures for heavy tailed distributions and the HL.E. properties are reasonable and easy to explain to users.

3.5 Control Charts for Mean Using Mann Whitney Statistic (U-test) Based on RSS

Tapang W. and Pongpullponsak A. [22] proposed a Mann-Whitney control chart base on RSS.

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

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

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

Assume that the test samples are themselves independent and are independent of the reference sample.

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

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

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

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 R1,…, Rn are the rank of the n observations $\overline{Y}_{rss,1},\overline{Y}_{rss,2},\ldots,\overline{Y}_{rss,n}$ in the complete sample of m + n observations.

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

When the process is in control, 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 indicate 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.

3.6 Construction of the Wilcoxon Rank Sum Test (W-test) Control Chart Based on RSS

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

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

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

Assume that the test samples are themselves independent and are independent of the reference sample.

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 $\overline{Y}_{rss,1},\overline{Y}_{rss,2},\ldots,\overline{Y}_{rss,n}$ in the complete sample of m + n observations.

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

When the process is in control, 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 indicate 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.

3.7 Control Chart for Mean Using Hodges-Lehmann Estimators (HL) Based on RSS

Step 1: Randomly reference sample of size m by RSS, denoted by $Y=\left(\overline{X}_{rss,1},\overline{X}_{rss,2},\ldots,\overline{X}_{rss,m}\right)$, in which a minimum size of 10 is required to achieve a significance level equal to 3 where sigma limits are traditionally used

Step 2: Use the table of positions of the two ordered Walsh averages to determine the subgroup control limit values, which will subsequently yield the upper and lower control values of Walsh averages

Step 3: Compute the Hodges–Lehmann estimator, when

- The center line for the control chart, the center line for the control chart will be the average of the Hodges–Lehmann estimators from each subgroup from step (3)

- The upper and lower control limits, the upper and lower control limits will be the median of the upper and lower values from step (2) above for all subgroups, where these are the upper control limit and the lower control limit

4. Performance of W-Chart, U-Chart and HL- Chart by Simulation

In evaluating the ARL performance of the three nonparametric 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 20 and 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.

4.1 The Mann-Whitney Control Charts (U-Charts)

From the Mann-Whitney control charts with sample sizes from control group in which *m* = 10, 15, 20 and the treatment group at

n = 10, 15, 20 determined control limits with α is closed to 0.0027 (closed to ± 3σ control limits in Shewhart $\bar{X}$ chart), denoted by –3 and 3 for action limits and ARL of selected schemes for standard normal data, uniform data and 11 shapes of data are shown in Table 1.

The performance of Mann-Whitney control chart (U-Chart) based on RSS for normal, uniform and Weibull distributions with the sample size = 10, 15 and 20 are compared as show in Table 1. The highest ARLs of normal and uniform distributions are at 789.47 and 697.67. The highest ARL for Weibull distribution is 769.23 at beta 3.2219 and n = 10. This show that at n = 10, the ARLs are around 700 which is about double the ARL of chart at +/- 3 standard deviation control limits (a = 0.0027, ARL = 370). We show graph of the control charts at sample size of 10 in Figure 3 (a)-(c).

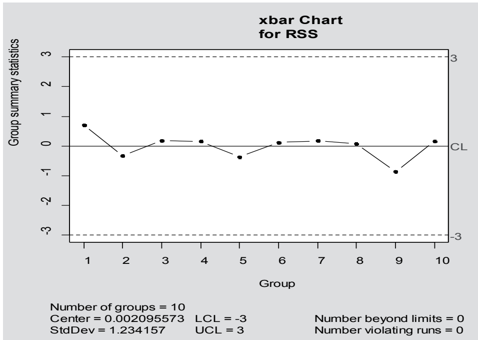
4.2 The Wilcoxon Rank Sum Test Control Chart W-Chart)

From the Wilcoxon rank sum test control charts with sample sizes from 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 (closed to ±3σ control limits in Shewhart $\bar{X}$ chart), denoted by –3 and 3 for action limits and ARL of selected schemes for standard normal data, uniform data and 11 shapes of Weibull data are shown in Table 2.

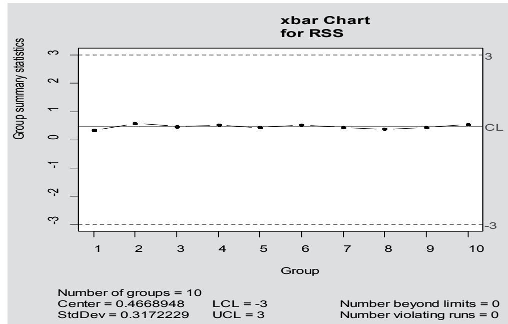
The performance of Wilcoxon rank sum test control chart (W-Chart) based on RSS for normal, uniform and Weibull distributions with the sample size = 10, 15 and 20 are compared as show in Table 2. The highest ARLs of normal and uniform distributions are at 625.18 and 731.71. The highest ARL for Weibull distribution is 769.23 at beta 0.6478 and n =

Table 1. Average run length (ARL) of the Mann-Whitney control chart (U-Chart) with Normal distribution, Uniform distribution and Weibull distribution.

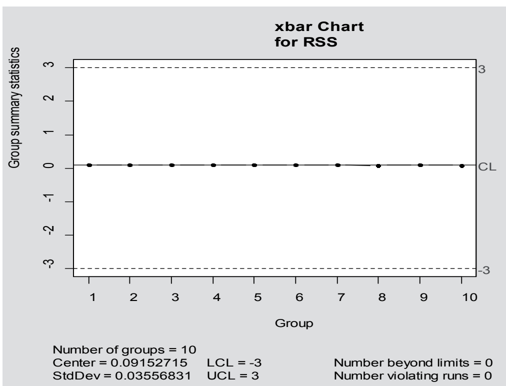
Distribution	Beta	10			15			20		
		ARL	SD	Median	ARL	SD	Median	ARL	SD	Median
Normal		789.47	0.42	0.14	422.54	0.39	0.05	491.8	0.34	0.03
Uniform		697.67	0.26	0.50	491.8	0.26	0.51	535.71	0.29	0.48
Weibull	3.2219	769.23	0.01	0.09	416.67	0.01	0.09	476.19	0.01	0.09
	2.211	555.56	0.06	0.43	400.00	0.09	0.39	384.62	0.09	0.43
	1.563	500.00	0.24	0.89	384.62	0.28	0.88	416.67	0.31	0.91
	1.00	633.33	0.67	1.98	666.67	1.02	2.31	454.55	0.90	1.43
	0.7686	543.07	0.49	0.73	588.24	0.56	1.07	500.00	0.45	0.87
	0.6478	709.09	0.78	1.25	555.56	0.78	1.25	500.00	0.67	0.99
	0.5737	588.24	1.12	1.11	555.56	1.12	1.56	526.32	1.30	0.93
	0.5237	633.33	2.23	1.72	454.55	3.16	0.88	384.62	3.16	0.88
	0.4873	588.24	1.31	1.91	343.78	1.46	1.39	400.00	2.15	1.59
	0.4596	479.26	0.90	1.66	476.19	1.73	1.32	526.32	1.18	1.31
	0.4376	555.56	3.48	1.74	656.67	2.58	1.67	588.24	2.08	1.67



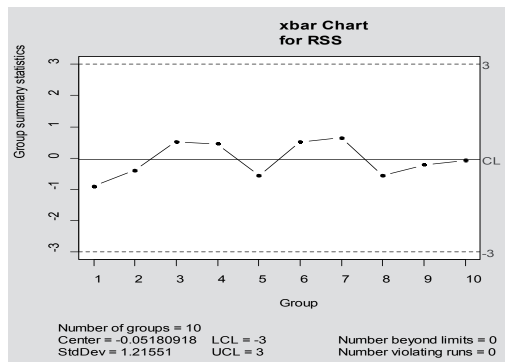
(a) U-Chart from Normal data.



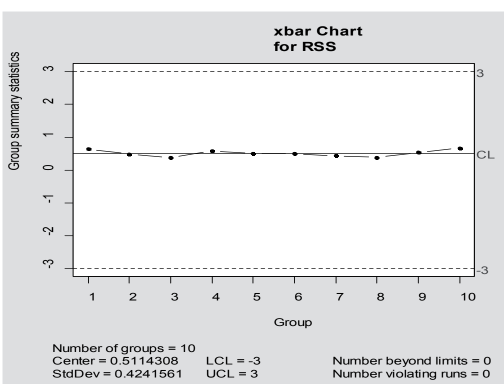
(b) U-Chart from Uniform data.



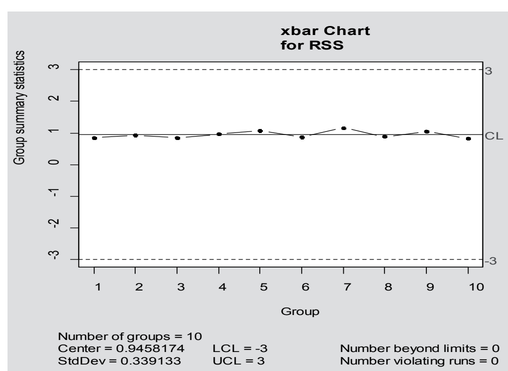
(c) U-Chart from Weibull data at beta 3.2219.



(d) W-Chart from Normal data.

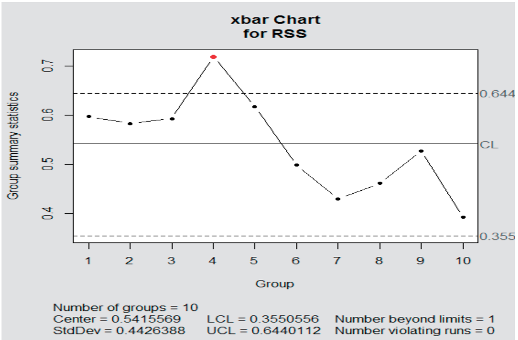


(e) W-Chart from Uniform data.

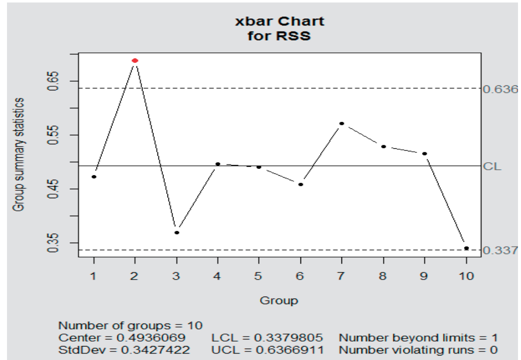


(f) W-Chart from Weibull data at beta 3.2219.

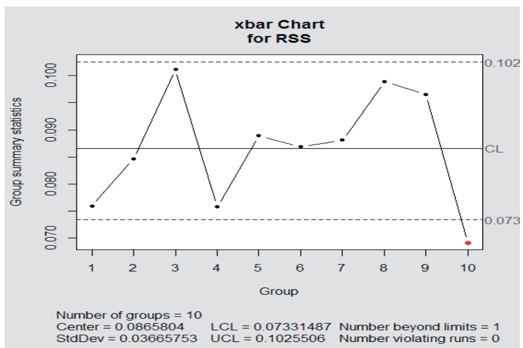
Figure 3. Graph of three control charts established from data with Normal distribution, Uniform distribution and Weibull distribution at the sample size of 10.



(g) HL-Chart from Normal data.



(h) HL-Chart from Uniform data.



(i) HL-Chart from Weibull data at beta 3.2219.

Figure 3. (Continued).

Table 2. Average run length (ARL) of the Wilcoxon rank sum test control chart (W-Chart) with Normal distribution, Uniform distribution and Weibull distribution.

Distribution	Beta	10			15			20		
		ARL	SD	Median	ARL	SD	Median	ARL	SD	Median
Normal		652.18	0.55	-0.14	428.57	0.37	0.13	447.76	0.34	0.02
Uniform		731.71	0.10	0.50	517.24	0.08	0.52	545.45	0.09	0.54
Weibull	3.2219	588.24	0.11	0.91	416.67	0.11	0.87	454.55	0.13	0.84
	2.211	625.10	0.18	0.77	476.19	0.18	0.89	526.32	0.16	0.89
	1.563	714.28	0.28	0.81	454.55	0.22	0.85	588.24	0.24	0.83
	1.00	526.32	0.37	0.93	476.19	0.72	0.93	370.37	0.37	0.99
	0.7686	666.67	0.91	0.84	476.19	0.56	0.84	400.00	0.73	0.87
	0.6478	769.23	0.62	1.52	416.17	0.64	0.91	526.32	1.29	0.89
	0.5737	555.56	0.97	1.36	526.32	0.80	1.09	400.00	0.93	0.92
	0.5237	588.24	1.01	1.44	500.00	0.54	1.11	476.19	1.97	1.35
	0.4873	666.67	2.04	0.56	666.67	1.21	1.80	625.00	1.79	1.25
	0.4596	414.29	3.76	1.54	454.55	2.69	1.14	370.37	4.15	1.42
	0.4376	476.19	1.04	1.40	526.32	2.18	1.26	526.32	3.20	1.16

10. This show that at n = 10, the ARLs are around 700 which is about double the ARL of chart at +/- 3 standard deviation control limits (a = 0.0027, ARL = 370).

We show graph of the control charts at sample size of 10 in Figure 3 (d)-(f).

4.3 The Hodges-Lehmann Estimator Control Chart (HL-Chart)

The ARL performance by 30,000 run lengths simulation of the Hodges-Lehmann estimator control chart for Normal data, Uniform data and Weibull data. The control limits are computed by the original method proposed by Alloway et al. [16] as shown in Table 3.

As can be seen in Table 3, at the sample size n=10, the Hodges-Lehmann control limit=(2,53) and probability $\alpha = 0.00390$. At the sample size of n=15, the control limit = (12,108) has probability $\alpha = 0.00336$ and when the sample size n=20, the control limit = (29,181), then probability $\alpha = 0.00272$.

The performance of Hodges-Lehmann estimator control charts (HL-Chart) based on RSS for normal, uniform and Weibull distributions with the sample size = 10, 15 and 20 are compared as show in Table 4. The highest ARL of normal distributions is 714.29 at n = 15. The highest ARL of uniform distributions is 769.23 at n = 20. The highest ARL for Weibull distribution is 666.67 at beta 3.2219 and 2.211 at n = 10. This show that the ARLs are around 700.

We show graph of the control charts at sample size of 10 in Figure 3 (g)-(i).

4.4 Comparisons of Efficiency of the U-Chart, W-Chart and HL- Chart Based on RSS when Sample Size is 10, 15, 20 and has Normal Distribution.

Figure 4 (a)-(c) shows ARL curve of three different methods when the distribution is normal. When comparing the shift mean graph at the sample size of 10, it is seen that when the shift is 0.0 and 1.0, the U-chart is most performance whereas the W-chart is most performance at the shift of 0.5, 1.5, 2.0, 2.5 and 3.0 (Figure 4 (a)). The Figure 4 (b) shows the performance

Table 3. The control limit of the HL-Chart denoted by lc and uc.

Distribution	Beta	10			15			20		
		ARL	uc	lc	ARL	uc	lc	ARL	uc	lc
Normal		588.24	0.64	0.35	714.29	0.34	-0.37	625.00	0.28	-0.36
Uniform		588.24	0.63	0.33	384.62	0.60	0.40	769.23	0.58	0.40
Weibull	3.2219	666.67	0.10	0.07	357.14	1.02	0.77	476.19	0.99	0.80
	2.211	666.67	1.10	0.68	476.19	1.04	0.71	400.00	1.04	0.75
	1.563	454.55	1.18	0.60	434.78	1.11	0.66	500.00	1.10	0.71
	1.00	384.62	1.48	0.52	357.14	1.35	0.65	357.14	1.29	0.66
	0.7686	370.37	2.12	0.38	357.14	1.75	0.64	384.62	1.78	0.66
	0.6478	322.58	2.36	0.43	322.58	2.24	0.59	357.14	2.06	0.63
	0.5737	384.62	3.02	0.46	357.14	2.91	0.55	434.78	2.66	0.66
	0.5237	357.14	5.68	0.44	526.32	3.33	0.55	357.14	2.95	0.73
	0.4873	344.83	4.37	0.41	303.03	3.77	0.69	357.14	3.55	0.69
	0.4596	333.33	1.70	0.39	454.55	4.67	0.60	434.78	4.25	0.74
	0.4376	366.92	6.06	0.41	434.78	5.18	0.58	344.83	4.11	0.77

Table 4. Average run length of the Hodges-Lehmann estimator control charts with normal distribution, uniform distribution and Weibull distribution.

Distribution	Beta	10			15			20		
		ARL	SD	Median	ARL	SD	Median	ARL	SD	Median
Normal		588.24	0.14	0.00	714.29	0.11	0.00	625.00	0.10	0.00
Uniform		588.24	0.04	0.50	384.62	0.03	0.50	769.23	0.03	0.50
Weibull	3.2219	666.67	0.04	0.89	357.14	0.03	0.90	476.19	0.03	0.90
	2.211	666.67	0.06	0.88	476.19	0.05	0.88	400.00	0.04	0.88
	1.563	454.55	0.08	0.89	434.78	0.07	0.89	500.00	0.06	0.89
	1.00	384.62	0.14	0.96	357.14	0.12	0.96	357.14	0.10	0.96
	0.7686	370.37	0.22	1.07	357.14	0.18	1.07	384.62	0.15	1.07
	0.6478	322.58	0.31	1.19	322.58	0.24	1.20	357.14	0.21	1.20
	0.5737	384.62	0.40	1.32	357.14	0.32	1.33	434.78	0.28	1.33
	0.5237	357.14	0.82	1.81	526.32	0.40	1.46	357.14	0.34	1.46
	0.4873	344.83	0.62	1.57	303.03	0.49	1.58	357.14	0.41	1.59
	0.4596	333.33	0.72	4.17	454.55	0.57	1.71	434.78	0.49	1.70
	0.4376	366.92	0.85	1.80	434.78	0.66	1.83	344.83	0.56	1.84

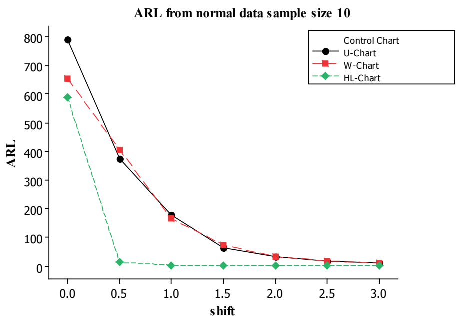
of the control charts at the sample size of 15. The graph demonstrates that at the shift of 0.0, the HL-chart is most performance whilst the W-chart is most performance when the shift is between 0.5 and 3.0. Showing in Figure 3 (c) is the performance at the sample size of 20. While the HL-chart is most performance at the shift of 0.0, the W-chart is most performance at the shift of 0.5 and the U-chart is most performance at the shift from 1.0 to 3.0.

4.5 Comparisons of Efficiency of the W-Chart, U-Chart and HL- Chart Based on RSS when Sample Size is 10, 15, 20 and has Uniform Distribution

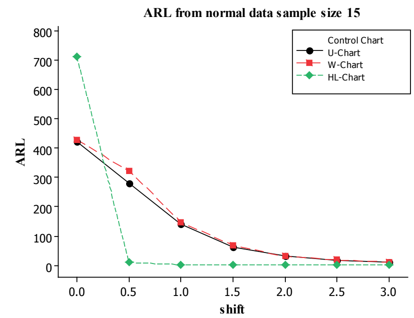
Figure 4 (d)-(f) shows ARL curve of three different methods when the distribution is uniform. Comparisons of the shift mean graph of ARL of the U-chart, W-chart and HL-chart (Figure 4 (d)) reveals that when the sample size is 10, the W-chart is most performance at all shift (0.0 to 3.0). Likewise, in Figure 4 (e), when the sample size is 15, the W-chart is most performance at all shift (from 0.0 to 3.0). Figure 4 (f), demonstrates the performance of the control charts at the sample size of 20. The graph shows that at the shift of 0.0, the HL-chart is most performance. And the U-chart is most performance when the shift is at 0.5, 1.0 and 1.5, whereas the W-chart is most performance when the shift is between 2.0 and 3.0.

4.6 Comparisons of Efficiency of the W-Chart, U-Chart and HL- Chart Based on RSS when Sample Size is 10, 15, 20 has Weibull Distribution

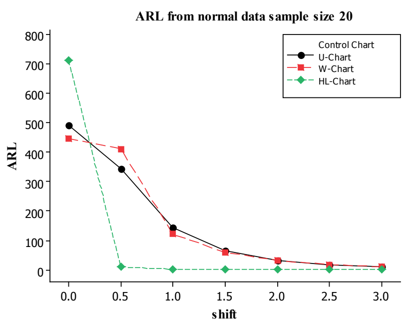
Figure 4 (g)-(i) shows ARL curve of three different methods when the distribution is Weibull and the beta is 3.2219. We compare the three control chart from Weibull data at this beta because it's close to Normal distribution. Comparisons of the shift mean graph of ARL of the U-chart, W-chart and HL-chart (Figure 4 (g) reveals that when the sample size is 10, the U-chart is most performance at the shift



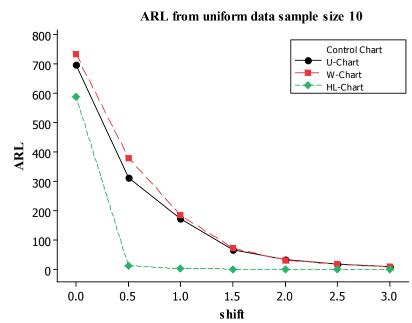
(a)



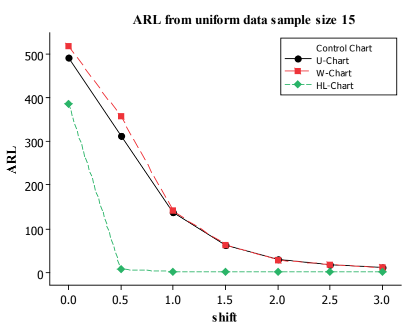
(b)



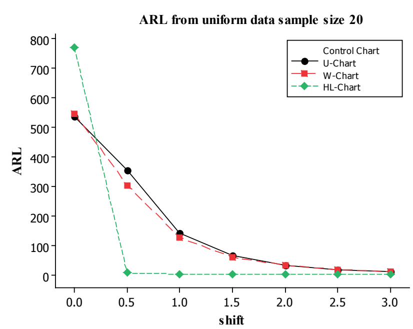
(c)



(d)

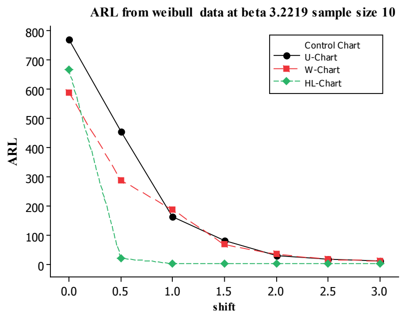


(e)

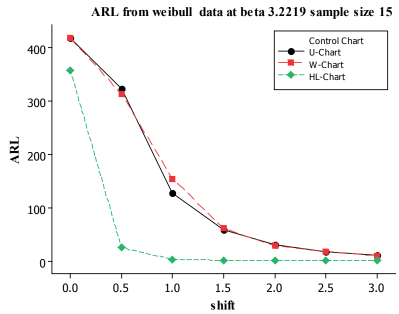


(f)

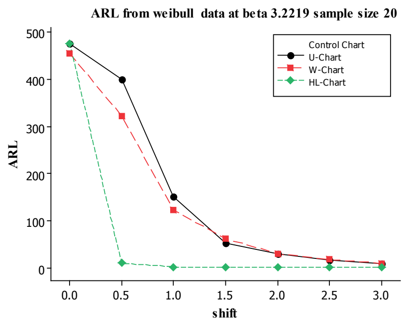
Figure 4. ARLs of three control charts established from data with normal distribution, uniform distribution and Weibull distribution at the beta of 3.2219 and the sample size of 10, 15 and 20.



(g)



(h)



(i)

Figure 4. (continued).

of 0.0, 0.5 and 1.5 whilst the W-chart is most performance when the shift is at 1.0, 2.0, 2.5 and 3.0. Figure 4 (h) shows the performance of the control charts at the sample size of 15. The W-chart is most performance when the shift is at 0.0, 1.0, 1.5, 2.0, 2.5 and 3.0. The U-chart is most performance at the shift of 0.5. Finally, in Figure 4 (i), the performance of the control charts at the sample size of 20 is presented. The graph shows that at the shift of 0.0, the HL-chart is most performance. And while the U-chart is most performance at the shift of 0.5 and 1.0, the W-chart is most performance at the shift ranging from 1.5 to 3.0.

5. Conclusions

In this study, we use RSS in developing non-parametric control charts for sample mean, the Mann-Whitney statistics (U-chart), the Wilcoxon rank sum test (W-chart) and the Hodges-Lehmann estimator (HL-chart). The data used in estimation are generated from normal distribution, uniform distribution and 11 shapes of Weibull distributions with the process shift in δ times of standard deviation (δ=0.50, 1, 1.5, 2, 2.5 and 3) are selected for the study. A minimum size of 10 is required to achieve a significant level equal to $\pm 3\sigma$ control limits which are traditionally used.

We finds that the ARL performance of the non-normal process distributions is higher than normal because the Rank Set Sampling (RSS) forms the new sample by choosing observations from the samples that arranged 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.

When we compared these control charts with the control charts mentioned above, we observe that the W-chart is best performance at the sample

size of 10 and 15 whereas the U-chart is best performance at the sample size of 20.

ACKNOWLEDGMENTS

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] McIntyre G.A., *Aust. J. Agr. Res.*, 1952; **3**: 385-390. DOI 10.1071/AR9520385.

[2] Halls L.K. and Dell T.R., *Forest Sci.*, 1966; **12**: 22-26.

[3] Muttlak H.A. and Al-Sabah W.S., *J. Appl. Stat.*, 2003; **30**: 1055-1078. DOI 10.1080/0266476032000076173.

[4] Pongpullponsak A. and Sontisamran P., *Chiang Mai J. Sci.*, 2013; **40**(3): 485-498.

[5] Chen Z. and Bai Z. H., *Sankhya A*, 2000; **62**: 178-192.

[6] Bakir S.T. and Reynolds M.R., *Technometrics*, 1979; **21**: 175-183. DOI 10.1080/00401706.1979.10489747.

[7] McDonald D., *Nav. Res. Logist.*, 1990; **37**: 627-646.

[8] Hackl P. and Ledolter J., *J. Qual. Technol.*, 1991; **23**(2): 117-126.

[9] Woodall W.H. and Montgomery D.C., *J. Qual. Technol.*, 1999; **31**: 376-386.

[10] Chakraborti S. et al., *J. Qual. Technol.*, 2001; **33**: 304-315.

[11] Bakir S.T., Classification of distribution free control charts, *Proceedings of Annual Meeting of American Statistical Association (Section Quality and Productivity)*, August 2001; 5-9.

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

[13] Gibbons J.D. and Chakraborti S., *Nonparametric Statistical Inference*, 4th Edn., Marcel Dekker, New York, 2003.

[14] Jayathavaj V. and Pongpullponsak A., *Chiang Mai J. Sci.*, 2014; **41**(5.2): 1457-1472.

[15] Ridout M.S. and Cobby J.M., *Appl. Stat.*, 1987; **36**: 145-152.

[16] Alloway J.A. and Raghavachari M., *J. Qual . Tech.*, 1991; **23**: 336-347.

[17] Albers W. et al., *Normal, Parametric and Nonparametric Control Charts, A Kata Driven Choice*, Memorandum No 1674, University of Twente, Netherlands, 2003.

[18] Pyzdek T., *Qual. Eng.*, 1995. DOI 10.1080/08982119508918823.

[19] Takahasi K. and Wakimoto K., *Ann. Inst. Stat. Math.*, 1968; **20**: 1-31. DOI 10.1007/BF02911622.

[20] Hodges J.L. and Lehmann E.L., *Ann. Math. Stat.*, 1963; **34**: 598-611. DOI 10.1214/aoms/1177704172.

[21] Lehmann E.L., *Theory of Point Estimation*, John Wiley and Sons, New York, 1983.

[22] Tapang W. and Pongpullponsak A., *East-West J. Math.*, 2012; a special volume: 1-10.

[23] MINITAB Thailand., *MINITAB 16 Order Number* 100004968850, Single License, February 02, 2010.