A nonparametric control chart based on the Mann-Whitney statistic

Citation for published version (APA):

Document status and date:
Published: 01/01/2003

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 01. Apr. 2019
A Nonparametric Control Chart based on the Mann-Whitney Statistic

S. Chakraborti                      M.A. van de Wiel
Department of Information Systems, Statistics                Department of Mathematics
and Management Science          and Computer Science
University of Alabama           Eindhoven University of Technology
Tuscaloosa, Alabama                 Eindhoven
U.S.A.                                              The Netherlands

Keywords: Statistical process control; Distribution-free; Shewhart $\bar{X}$ chart; Parameter estimation; Reference sample; Average run length and run length percentiles; Heavy-tailed and skewed distributions; Saddlepoint and Edgeworth approximations; Monte Carlo estimation.

Abstract

Nonparametric or distribution-free charts can be useful in statistical process control when there is limited knowledge about the underlying process. In this paper a Shewhart-type chart is considered for the location, based on the Mann-Whitney statistic. The control limit calculations use Lugannani-Rice saddlepoint, Edgeworth and other approximation methods along with Monte Carlo estimation and take account of the dependence resulting from the use of a reference sample from the in-control process. Interactive software, made available on a website, enables a complete implementation of the procedure by remote users. An illustrative numerical example is presented. Performance comparisons, in terms of the in-control unconditional ARL and the 5th percentile of the conditional run length distribution show that the nonparametric chart has superior in-control properties with regard to the classical Shewhart $\bar{X}$ chart. Further comparisons of the out-of-control unconditional ARL and the 95th percentiles of the conditional run length distributions show that the proposed chart is almost as good as the $\bar{X}$ chart for the normal distribution, but is far more powerful for a heavy-tailed distribution such as the Laplace, or for a skewed distribution such as the Gamma.

Introduction

In statistical process control (SPC), the pattern of chance causes is often assumed to follow the normal distribution. It is well recognized however, that in many applications the underlying process distribution is not known sufficiently to assume normality (or any other parametric distribution), so that statistical properties of commonly used charts, designed to perform best under the assumed distribution (such as normality), could be potentially (highly)

---

1 Corresponding author: e-mail: m.a.v.d.wiel@TUE.nl
affected. In situations like this, development and application of control charts that do not depend on normality, or more generally, on any specific parametric distributional assumptions seem highly desirable. Distribution-free or nonparametric control charts can serve this purpose.

There is increasing awareness about the value of nonparametric methods in statistical process control and monitoring. Woodall and Montgomery (1999) state that “There would be appear to be an increasing role for non-parametric methods, particularly as data availability increases.” Chakraborti, Van der Laan and Bakir (hereafter CVB) (2001) presented an overview of the literature on univariate nonparametric control charts for variables. In addition to providing extensive references, they listed six primary advantages of nonparametric control charts: (i) simplicity, (ii) no need to assume a particular distribution for the underlying process, (iii) the in-control run length distribution is the same for all continuous distributions (the same is true for the false alarm rate; and thus different nonparametric charts can be compared more easily), (iv) more robust and outlier resistant, (v) more efficiency in detecting changes when the true distribution is markedly non-normal, particularly with heavier tails, and (vi) no need to estimate the variance to set up charts for the location parameter.

To summarize, a key advantage of the nonparametric charts is the flexibility that no particular distribution (such as normal) for the underlying process needs to be assumed, in advance, in order to implement the charts. Also, generally speaking, the nonparametric charts are likely to inherit the robustness properties of nonparametric tests and are, therefore, more likely to be less impacted by outliers. So for unknown yet expected to be heavy-tailed or skewed distributions, these procedures are especially preferable. On the downside, one might suspect that all of this “good” might come at a “loss of power.” On this point CVB state “It should be noted that nonparametric methods can be somewhat less efficient than their parametric counterparts, provided of course that one has a complete knowledge of the underlying stochastic process for which the particular parametric method was specifically designed; however, the reality is that such information is seldom, if ever, available to the quality practitioner.” All in all, nonparametric methods provide a set of tools flexible enough to be useful in many practical situations and we strongly advocate that when available, the practitioners consider using these methods in process control and monitoring problems.
Terminology and Problem

An important problem in statistical process control is the problem of monitoring the center or the location of a process. The location parameter could be the mean or the median or some percentile of the distribution; the latter two are often more attractive when the underlying distribution is known to be skewed. Among the available control charts for the mean of a process, the classical Shewhart $\bar{X}$ chart is the most popular because of its inherent simplicity and practical appeal. Although this chart is quite versatile, as noted in CVB, there are situations in practice where a nonparametric chart is expected to be a better alternative. In many applications of a control chart, the process parameters are often assumed specified or known. If this is not the case and parameters are estimated, there is growing evidence in the recent literature that most charts including the Shewhart $\bar{X}$ chart, behave quite differently with respect to their false alarm rates and run length. We do not assume that the process parameters are specified or that the process distribution is known, instead, we assume that a reference sample is available from the in-control process.

There are some nonparametric charts available in the literature; the reader is referred to CVB for references and detailed accounts. Among the Shewhart type charts, for example, Amin, Reynolds and Bakir (1995) considered charts based on the sign test when the process median is specified. This is not the problem considered here. Willemein and Runger (1996) considered the same setting as ours and proposed using two order statistics from the reference sample to define the control limits. Chakraborti, Van der Laan and Van de Wiel (2003) (hereafter CVV) considered the so-called precedence charts. These charts are similar to the Willemein-Runger charts. It has been shown that the precedence charts are good alternatives to the $\bar{X}$ chart in some situations.

While the precedence charts are a step in the right direction, it is known that the nonparametric test underlying this chart, the so-called precedence test, is neither the most powerful test (for location) nor the most commonly used nonparametric test in practice. In this paper we consider a chart based on the popular and more powerful Mann-Whitney (1947) (hereafter MW) test. Some readers might be more familiar with the Wilcoxon (1945) rank-sum test, which is equivalent to the MW test. For a discussion of these tests and their properties, the reader is referred to any nonparametrics book, e.g., by Gibbons and Chakraborti (2003) or any
standard statistical methods book. Gibbons and Chakraborti (2003; page 279) state: “Many statisticians consider the Mann-Whitney test the best nonparametric test for location.” The MW test a direct competitor to the normal theory based, two-independent-sample t-test. Remarkably, even when the underlying distributions are normal, the MW test is about 96% as efficient (again see e.g., Gibbons and Chakraborti, 2003; pp. 278-279) as (the most efficient) 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 like the Laplace (double exponential) or the logistic distribution, or skewed distributions like the exponential, the MW test is known to be more efficient than the t-test. In short, MW test is the practitioners’ choice when not much is known about the shape of the underlying distributions.

The MW Control Chart

Suppose that a reference sample of size $m$, denoted by $X=(X_1,...,X_m)$, is available from an in-control process and that $Y=(Y_1,...,Y_n)$ denotes an arbitrary, test sample of size $n$. The superscript $h$ is used to denote the $h^{th}$ test sample, $Y^h=(Y_1,...,Y_n), \ h=1,2,...$. Assume that the test samples are themselves independent and are independent of the reference sample. The MW test is based on the total number of $X$-$Y$ pairs where the $Y$ observation is larger than the $X$. This is the statistic

$$M_{XY} = \sum_{i=1}^{m} \sum_{j=1}^{n} I(X_i < Y_j)$$

$$= \sum_{j=1}^{n} \{I(Y_j > X_1) + ... + I(Y_j > X_m)\}, \ \ (1)$$

where $I(X_i < Y_j)$ is the indicator function for the event $\{X_i < Y_j\}$. Note that $M_{XY}$ lies (attains values) between 0 and $mn$ and large values of $M_{XY}$ indicate a positive shift, whereas small values indicate a negative shift.

The proposed MW chart uses $M_{XY}^h$ as the charting statistic and signals on the $h^{th}$ test sample if

$$M_{XY}^h < L_{mn} \ \ \text{or} \ \ M_{XY}^h > U_{mn},$$
where \( L_{mn} \) and \( U_{mn} \) are the lower control limit (LCL) and the upper control limit (UCL), respectively. The distribution of \( M_{XY} \) is known to be symmetric about \( mn/2 \) when the process is in-control, so that it is reasonable to take \( L_{mn} = mn - U_{mn} \). Thus the proposed charting procedure is quite simple: at every new test sample, calculate the charting statistic \( M_{XY} \) between the test sample and the reference sample and compare it to the control limits \( U_{mn} \) and \( mn - U_{mn} \). The process is declared to be out-of-control when \( M_{XY} \) is either above \( U_{mn} \) or below \( L_{mn} \). In this paper we focus on two-sided charts. Naturally, one-sided charts can be developed analogously.

Altukife (2003) considered a control chart based on the “sum of ranks.” However, his chart is not distribution-free since his chart constants are derived for each of four parametric distributions. Also, the problem considered in this paper appears to be different from ours in that a reference sample is not used for prospective process monitoring.

**Design and Implementation**

Practical implementation of the chart requires the upper control limit \( U_{mn} \) \( (L_{mn} = mn - U_{mn}) \). Typically, the control limits (or charting constants) are found for some specified in-control average run length (ARL) value, say 370 or 500. To this end note that if the successive charting statistics, \( M_{XY}^1, M_{XY}^2, \ldots \) corresponding to test sample 1, 2, \ldots, would be statistically independent, finding control limits would be a relatively easy task, because then, as in the standards known case (see e.g., Montgomery, 2001, p. 228), the \( ARL_0 \) would be equal to the reciprocal of the false alarm rate, given by \( p_0 = 2P(M_{XY} \geq U_{mn}) \). In that case the upper control limit \( U_{mn} \) would be equal to the two-sided critical value of a MW test with size equal to \( 1/ARL_0 \) and one can use tables or software to find \( U_{mn} \). Note that even in this simple case, a practical problem is that in a typical control charting applications \( ARL_0 = 370 \), which means that the UCL is the upper critical value for a MW test with size \( \frac{1}{2}(1/370) = 0.00135 \). Such critical values would require calculation of extreme right-tail probabilities and will most likely not be found in standard tables for the MW test (usually constructed for size 0.05 or 0.01).

Even if such critical values can be found, the main problem with this approach is that in our case the successive statistics \( M_{XY}^1, M_{XY}^2 \ldots \) are dependent, and in general, this dependence
can not be ignored (see for example, Quesenberry (1993) and Chakraborti (2000) for the Shewhart chart). The dependence arises due to the fact that all test samples are compared with the same reference sample $X$ and it affects all operational and performance characteristics of the chart, such as false alarm rate, $ARL$, etc.

In order to calculate the control limits for a specified $ARL_0$, first we develop an efficient method to calculate the $ARL$ (and hence the $ARL_0$). The same method is used to compute certain run length related performance characteristics of the chart such as some percentiles and out-of-control $ARL$'s.

**Calculation of $ARL$**

Let $F$ and $G$ represent the cdf of $X$ and $Y$, respectively. Suppose that $F$ and $G$ are continuous, so that “ties” between the $X$’s and the $Y$’s, as well as within the $X$’s and the $Y$’s themselves, do not occur theoretically. We will say more about the ties later. In this section we focus on the computation of the average run length for both the in and out-of-control cases. In order to take proper account of the dependence mentioned earlier, the $ARL$ is derived by conditioning on the reference sample, an idea first explicitly introduced in Chakraborti (2000).

To this end, observe that the probability of signal with any test sample, given the reference sample $X = \bar{x}$, is

$$p_G(\bar{x}) = P_G(M_{\bar{x}Y} < mn - U_{mn}) + P_G(M_{\bar{x}Y} > U_{mn}).$$  \hspace{1cm} (2)

Let $N$ denote the run length random variable for the chart. Given the reference sample $X = \bar{x}$, say, the independence of two arbitrary test samples $Y^h$ and $Y^l$ implies the independence of $M_{\bar{x}Y}^h$ and $M_{\bar{x}Y}^l$. Hence,

$$ARL = E(N)$$

$$= E_f[E_G(N \mid X = \bar{x})]$$

$$= E_f\left(\frac{1}{p_G(\bar{x})}\right)$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{p_G(\bar{x})} dF(x_1) \cdots dF(x_m)$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} v(G(x_1), \ldots, G(x_m)) dF(x_1) \cdots dF(x_m)$$
where $1/p_G(\bar{x})$ is written as $\nu(P(Y_j < x_1),...,P(Y_j < x_m)) = \nu(G(x_1),...,G(x_m))$, say, for some (implicitly defined) deterministic function $\nu$. The third equality in (3) follows since given $X = \bar{x}$, $N$ is geometrically distributed with probability $p_G(\bar{x})$.

The random variable $E_G(N \mid X)$ is of interest in itself. It is the conditional average run length given the random reference sample. From (3) we observe that the unconditional ARL is the mean (expectation) of $E_G(N \mid X)$. Later, we show that other characteristics of the distribution of $E_G(N \mid X)$, in particular percentiles, are very useful to study control chart performance, and we develop procedures to compute these as well. However, for the time being we concentrate on the unconditional ARL to explain the proposed chart.

When the process is in-control, that is when $X$’s and $Y$’s come from the same distribution $F = G$, we have:

$$ARL_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{p_F(\bar{x})} dF(x_1)...dF(x_m)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu(F(x_1),...,F(x_m)) \ dF(x_1)...dF(x_m)$$

$$= \int_{0}^{1} \int_{0}^{1} \nu(x_1,...,x_m) \ dx_1...dx_m$$

$$= \int_{0}^{1} \int_{0}^{1} \frac{1}{p_U(\bar{x})} \ dx_1...dx_m,$$

(4a)

Equation (4a) shows that the in-control ARL does not depend on $F$, in fact the same argument can be used to show that the in-control run-length distribution does not depend on $F$ and hence the proposed chart is distribution-free. Equation (4b) shows that in the in-control situation, we may assume $F$ to be the cdf of the uniform$(0,1)$ distribution, which simplifies calculations considerably. Note that $p_U(\bar{x}) = P_U(M_{\bar{x}Y} < mn - U_{mn}) + P_U(M_{\bar{x}Y} > U_{mn})$ is the conditional probability of a signal at any test sample given the reference sample, when the process is in-control, and the subscript $U$ denotes that both the reference sample and the test sample are obtained from the uniform$(0,1)$ distribution.

Clearly (4) is a special case of (3). We need to calculate (4) to implement the chart and (3) to evaluate and compare chart performance. For both of these objectives there are two
problems. First, an explicit formula for $p_G(\bar{x})$ (or $p_U(\bar{x})$), in general, is not known, which prevents a direct computation. Second, we have an $m$-dimensional integration in both (3) and (4). Our approach is basically to calculate $p_G(\bar{x})$ (and $p_U(\bar{x})$) (exactly or approximately), using a fast algorithm, and then use that to approximate the integral in (3 and 4) with a Monte Carlo estimate to get estimates

$$ARL \approx \frac{1}{K} \sum_{i=1}^{K} \frac{1}{p_G(\bar{x}_i)},$$

and

$$ARL_0 \approx \frac{1}{K} \sum_{i=1}^{K} \frac{1}{p_U(\bar{x}_i)},$$

where $\bar{x}_i = (x_{i1},...,x_{im})$ is the $i^{th}$ Monte Carlo sample, $i = 1,...,K$, of which each component is drawn from some specified $F$ for the $ARL$ and from the uniform(0,1) for the $ARL_0$ and $K$ denotes the number of Monte Carlo samples. For an accurate approximation of the $ARL$ or the $ARL_0$, $K$ needs to be sufficiently large and therefore fast computation of $p_G(\bar{x})$ (and $p_U(\bar{x})$) for an arbitrary reference sample $\bar{x}$ is essential for the practical use of this approximation.

The $ARL$ calculations proceed in two steps. The first step is a fast and efficient calculation of the signal probability $p_G(\bar{x})$.

**Fast Computation of Signal Probability**

Given the application environment and the potentially large number of reference sample observations $m$, the need for speed and efficiency is self-evident. We detail the procedure for fast computation of $p_G(\bar{x})$; calculation of $p_U(\bar{x})$ follows along similar lines. From (1) it is seen that the MW statistic is simply the sum $\sum_{j=1}^{m} C_j$, where $C_j$ represents the total number of times (an arbitrary) $Y_j$ exceeds $X_i$, $i = 1,...,m$. (or equivalently, $C_j$ represents the number of $X$’s that precede $Y_j$, $j = 1,...,n$). From (2) it follows that the calculation of $p_G(\bar{x})$ essentially requires calculation of the upper-tailed probability $P_G(M_{\bar{x}Y} > U_{mn})$ say, and this requires an efficient enumeration of all $n$-tuples $\{C_1,...,C_n\}$ for which the sum is greater than $U_{mn}$ and the summation of the probabilities for such tuples.
Note that $P(C_j = l)$ is equal to $P(X_{(i)} < Y_j < X_{(i+1)})$, where $X_{(i)}$ denotes the $i^{\text{th}}$ ordered observation in the reference sample for $i = 1, ..., m$ with $X_{(0)} = -\infty$ and $X_{(m+1)} = \infty$. Given the reference sample, the last probability is simply $P(x_{(l)} < Y_j < x_{(l+1)})$, which is denoted by $a_l$, $l = 0, ..., m$. The conditional probability generating function (pgf) of $C_j$ is

$$H_1(z) = \sum_{l=0}^{m} P(C_j = l)z^l = \sum_{l=0}^{m} a_l z^l. \quad (6)$$

Again, given (conditioned on) the reference sample $X = \bar{x}$, the random variables $C_j$ are independent, since $Y_j$ are independent. Therefore, the conditional pgf of $M_{xy}$ (the sum of the $C_j$s), is simply the product of the pgfs in (6)

$$H_2(z) = \sum_{j=0}^{mn} P(M_{xy} = j)z^j = (\sum_{j=0}^{m} a_j z^j)^n. \quad (7)$$

In principle, the probability $P_G(M_{xy} > U_{mn})$ can be calculated by expanding the power in (7) and collecting the coefficients of all terms with degree greater than $U_{mn}$. However, for $m$ moderate to large (say $m \geq 100$) and the test sample size $n$ not very small (say, $n \geq 5$) this takes a considerable amount of time, especially since the procedure has to be repeated $K$ (a large number) times once for each Monte Carlo sample. Fortunately, alternative, faster, methods can be used for this purpose. One such method is a “branch-and-bound” algorithm, based on Mehta et al. (1988), which is useful to shorten the intermediate expressions that result from expanding (6), saves considerable time. For example, suppose we have expanded $H_{2,i}(z) = (\sum_{l=0}^{m} a_l z^l)^i$ for $i < n$. To obtain $H_{2}(z)$ we will need to multiply $H_{2,i}(z)$ by $H_{2,n-i}(z)$. Since the degree of $H_{2,n-i}(z)$ equals $m^*(n-i)$, we know that all terms in $H_{2,i}(z)$ for which the exponent is smaller than $U_{mn} - m^*(n-i)$ will not contribute to $p_G(\bar{x})$ and therefore such terms can be deleted from $H_{2,i}(z)$. This method can be applied iteratively for $i = 1, ..., n$. As noted earlier, in a typical control charting application the necessary upper control limit $U_{mn}$ is expected to be quite large (found in the upper tail of the in-control distribution of the charting statistic $M_{xy}$) so as to guarantee a sufficiently large $ARL_0$. Thus the branch-and-bound procedure can be quite efficient here since a large number of intermediate terms can deleted from evaluation.
Even with the branch-and-bound technique, \( m \) and/or \( n \) might be just too large to allow for exact computations within a practically reasonable amount of time and hence a good approximation to the control limits may be necessary. From the central limit theorem for a sum of independent random variables we observe that \( M_{\overline{x}_Y} \), which equals \( \sum_{j=1}^{n} C_j \), converges to a normal random variable when \( n \to \infty \). This can in principle be used to find a normal approximation to \( P_G(M_{\overline{x}_Y} > U_{mn}) \). However, it is well known that the normal approximation is not very accurate in the tails for a finite \( n \). In our context, \( U_{mn} \) is typically far in the upper tail of the distribution of \( M_{\overline{x}_Y} \) and \( n \) is usually not very large, so the normal approximation is not very useful. Instead, we find that the Lugannani-Rice formula (hereafter LR-formula; see Jensen 1995, page 74) for the upper-tail probability for the mean of i.i.d. discrete random variables (which is a "saddlepoint" approximation formula) is particularly suitable for our purpose. This formula is known to be more accurate than the normal approximation in the tails of a distribution and is based on knowledge of the cumulant generating function of \( C_j \), which is readily found from the pgf in (5):

\[
k(t) = \log[G(z)_{z=e^t}] = \log[\sum_{i=0}^{m} a_i e^{it}].
\]

Let \( m(t) \) and \( \sigma^2(t) \) denote the first and the second derivative of \( k(t) \), respectively. Further let \( u = (U_{mn} + 1) / n \) and \( \overline{M}_{\overline{x}_Y} = M_{\overline{x}_Y} / n \). The saddlepoint \( \gamma \) is the solution to the equation \( m(t) = u \); the solution is determined numerically, using the Newton-Raphson method.

Using (3.3.17) in (Jensen, 1995, page 79) we obtain an approximation:

\[
P_G(M_{\overline{x}_Y} > U_{mn}) = P_G(\overline{M}_{\overline{x}_Y} > U_{mn} / n) = P_G(\overline{M}_{\overline{x}_Y} \geq u) = 1 - \Phi(u) + \phi(u)(\frac{1}{\lambda} - \frac{1}{r}),
\]

where

\[
\lambda = n^{1/2}(1 - e^r)\sigma(\gamma), \quad r = \text{sgn}(\gamma)[2n(\mu - k(\gamma))]^{1/2}
\]

and \( \Phi(.) \) and \( \phi(.) \) are, respectively, the cdf and the pdf of the standard normal distribution.

**Monte Carlo estimation of ARL**

Having computed (or approximated) \( p_G(\overline{x}) \) efficiently in step one, we apply Monte Carlo simulation to approximate the \( ARL \) in step two. The main question here is regarding the
size of $K$, so that the Monte Carlo error is acceptably small. Note that since the $ARL$ is the average of the conditional $ARL_G(X)$ over all possible $X$’s, and the $K$ simulated reference samples are independent, the Monte Carlo standard error of the estimate $\hat{ARL}$ equals 

$$\sigma_{mc} = \sigma(ARL_G(X))/\sqrt{K},$$

where $\sigma(ARL_G(X))$ is the standard deviation of $ARL_G(X)$. We do not know this standard deviation, but we may estimate it by the sample standard deviation $s(ARL_G(X))$ and obtain an estimate of $\sigma_{mc}$, say $s_{mc}$. Then we can try to guarantee that

$$s_{mc} = s(ARL_G(X))/\sqrt{K} \leq D,$$  \hspace{1cm} (9)

where $D$ is either a specified number or a percentage of the current estimate $\hat{ARL}$. We can start with say $K = 100$, increase $K$, compute $s_{mc}$, and repeat the process until the specification $s_{mc} \leq D$ is met.

The same procedure works for the calculation of $\hat{ARL}_0$. Suppose we would like the estimated the in-control $ARL$, $\hat{ARL}_0$, to be approximately 500. Then we may set $D$ to be 2% of 500, that is, equal to 10. With this 2% “error” specification we observed that $K$ was usually under 1000. We calculate the Monte Carlo estimate $\hat{ARL}_0$ of (unconditional) $ARL_0$ using (5b) with $K = 1000$ and using five different methods to compute or approximate $p_U(\bar{x})$. We kept $K$ constant for these computations to obtain a fair comparison between the computing times. The five methods are:

1. Exact: Monte Carlo simulation and exact computation of $p_U(\bar{x})$. Monte Carlo error decreases with order $\sqrt{K}$.
2. LR-formula: Monte Carlo simulation and formula (8) to approximate $p_U(\bar{x})$.
3. Normal: Monte Carlo simulation and normal approximation for $p_U(\bar{x})$.
4. Fixed reference sample: Fix reference sample to $\bar{q} = (1/(m+1),...,m/(m+1))$ and approximate $ARL_0$ by $1/p_U(\bar{q})$.
5. 1/(false alarm rate): approximate $ARL_0$ by $1/(2P_0(M_{xy} > U_{mn}))$, where the subscript 0 denotes the in-control case $F = G$. 

Approximations 4 and 5 will be explained later in this section. Table 2 below displays the estimated $ARL_0$ values and the upper control limits for fifteen combinations of $m$ and $n$. The table also shows the computing times on a 1.7GHz Pentium PC with 128MB of internal RAM.

Table 2: $ARL_0$ approximations and computing times

<table>
<thead>
<tr>
<th>$m$</th>
<th>$n$</th>
<th>$U_{true}$</th>
<th>Exact $ARL_0$ time (sec.)</th>
<th>LR-formula $ARL_0$ time (sec.)</th>
<th>Normal $ARL_0$ time (sec.)</th>
<th>Fixed reference sample $ARL_0$ time (sec.)</th>
<th>$1/(false\ alarm\ rate)$ $ARL_0$ time (sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>5</td>
<td>217</td>
<td>486</td>
<td>54</td>
<td>506</td>
<td>36</td>
<td>403</td>
</tr>
<tr>
<td>10</td>
<td>89</td>
<td>305</td>
<td>486</td>
<td>395</td>
<td>506</td>
<td>34</td>
<td>327</td>
</tr>
<tr>
<td>25</td>
<td>857</td>
<td>488</td>
<td>4850</td>
<td>491</td>
<td>491</td>
<td>31</td>
<td>425</td>
</tr>
<tr>
<td>100</td>
<td>5</td>
<td>435</td>
<td>496</td>
<td>220</td>
<td>505</td>
<td>48</td>
<td>219</td>
</tr>
<tr>
<td>10</td>
<td>776</td>
<td>505</td>
<td>1920</td>
<td>506</td>
<td>506</td>
<td>47</td>
<td>339</td>
</tr>
<tr>
<td>25</td>
<td>1707</td>
<td><strong>496</strong></td>
<td>26168</td>
<td>503</td>
<td>503</td>
<td>48</td>
<td>422</td>
</tr>
<tr>
<td>500</td>
<td>5</td>
<td>2172</td>
<td>491</td>
<td>10633</td>
<td>496</td>
<td>207</td>
<td>226</td>
</tr>
<tr>
<td>10</td>
<td>3872</td>
<td><strong>73516</strong></td>
<td><strong>373516</strong></td>
<td><strong>2673</strong></td>
<td><strong>513</strong></td>
<td><strong>179</strong></td>
<td><strong>367</strong></td>
</tr>
<tr>
<td>25</td>
<td>8484</td>
<td><strong>7.59*10^3</strong></td>
<td><strong>494</strong></td>
<td><strong>176</strong></td>
<td><strong>445</strong></td>
<td><strong>1.6</strong></td>
<td><strong>578</strong></td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>4347</td>
<td><strong>31766</strong></td>
<td><strong>500</strong></td>
<td><strong>356</strong></td>
<td><strong>235</strong></td>
<td><strong>2.1</strong></td>
</tr>
<tr>
<td>10</td>
<td>7732</td>
<td><strong>3.42*10^3</strong></td>
<td><strong>499</strong></td>
<td><strong>373</strong></td>
<td><strong>355</strong></td>
<td><strong>2.4</strong></td>
<td><strong>516</strong></td>
</tr>
<tr>
<td>25</td>
<td>16942</td>
<td><strong>3.15*10^6</strong></td>
<td><strong>500</strong></td>
<td><strong>348</strong></td>
<td><strong>442</strong></td>
<td><strong>1.7</strong></td>
<td><strong>548</strong></td>
</tr>
<tr>
<td>2000</td>
<td>5</td>
<td>8691</td>
<td><strong>1.71*10^3</strong></td>
<td><strong>503</strong></td>
<td><strong>713</strong></td>
<td><strong>234</strong></td>
<td><strong>2.1</strong></td>
</tr>
<tr>
<td>10</td>
<td>15460</td>
<td><strong>1.44*10^6</strong></td>
<td><strong>504</strong></td>
<td><strong>659</strong></td>
<td><strong>354</strong></td>
<td><strong>1.9</strong></td>
<td><strong>513</strong></td>
</tr>
<tr>
<td>25</td>
<td>33855</td>
<td><strong>1.29*10^7</strong></td>
<td><strong>509</strong></td>
<td><strong>676</strong></td>
<td><strong>446</strong></td>
<td><strong>2.1</strong></td>
<td><strong>531</strong></td>
</tr>
</tbody>
</table>

** $ARL_0$ could not reliably be estimated within reasonable time; computing time for $K = 1000$ is obtained by multiplying computing time for $K = 1$ by 1000 (sampling algorithm is linear in K).**

Several observations can be made from a study of the entries of Table 2. First it is seen that the “gold standards” or the exact computations are very time-consuming for most values of $m$ and $n$. However, when available, they naturally form the basis of our comparisons of various approximations. It is seen that the normal approximation is not very accurate, but the LR-approximation is, particularly for $m \geq 100$, $n \leq 10$ and for $m = 500$, $n = 5$. Since the LR-approximation is known to become more accurate when the sample sizes increase, one may safely apply the LR-formula also when $m \geq 50$ and $n \geq 5$ in order to implement the proposed chart. It may be noted that when $m$ increases, the computing times with the LR–formula also increase, although, by far, not as dramatically as the times for the exact computations. This suggests that in practice (for finding the chart constants, to be discussed next) there is still room for an alternative, quick approximation of the $ARL_0$, particularly for large values of $m$. To this end, we propose the following. Observe that formula (5) requires Monte Carlo simulation using
draws of the reference sample. When \( m \) is large, the empirical cdf \( F_m(x) \) converges to \( F(x) \). Hence, for large \( m \) we may obtain a reasonable approximation to \( ARL_0 \) by replacing the \( i^{th} \) reference sample observation by the \( i/(m+1)^{th} \) quantile of the uniform(0,1) distribution, \( i = 1,\ldots,m \). Since this quantile is simply \( q_i = i/(m+1) \), we approximate the \( ARL_0 \), for large \( m \), by \( 1/p_u(\bar{q}) \), where \( \bar{q} = (q_1,\ldots,q_m) \). This is method 4. The major benefit with this approximation is that we need to compute \( p_u(\bar{x}) \) only once (namely at \( \bar{x} = \bar{q} \)) instead of \( K \) times for \( K \) Monte Carlo reference samples.

A second quick approximation for the \( ARL_0 \) is given by the inverse of the false alarm rate (which of course ignores the dependence between successive charting statistics) and thus \( ARL_0 \) is approximated by \( 1/(2P_0(M_{XY} > U_{mn})) \). This is method 5. Chakraborti (2000) showed that for the Shewhart \( \bar{X} \) chart, \( 1/FAR \) is a lower bound to \( ARL_0 \) and noted that this bound can serve as a “quick and dirty” approximation to the \( ARL_0 \) for moderate to large values of \( m \). The same result can be shown to hold for the MW chart. However, even for this “simple” approximation, fast computation of the Mann-Whitney in-control tail probability, \( P_0(M_{XY} > U_{mn}) \), is not a trivial problem, particularly in our situation where \( m \) and \( n \) are highly unbalanced with \( n \) being much smaller relative to a large \( m \). To this end, we used formula (11) of Fix and Hodges (1955), based on an Edgeworth approximation, which significantly improves the normal approximation by including moments of order higher than 2. Compared to the LR-formula, we observe that both “fast” approximations are quite good for \( m \geq 1000 \) and that the fixed reference sample approximation (method 4) performs somewhat better for relatively small values of \( n \) (\( n = 5,10 \)) than for \( n = 25 \).

**Determination of Chart Constants**

So far we have outlined how to compute \( ARL_0 \) corresponding to a given value of \( U_{mn} \). In order to implement the chart, however, we are faced with the inverse problem: find (the unknown control limit) \( U_{mn} \) such that the \( ARL_0 \) equals a pre-specified value, say 500. Towards this end we use an iterative procedure based on linear interpolation. To begin the iteration, a reasonable approximation of \( U_{mn} \) is needed. This is where the fast approximations 4 and 5, particularly 5, are found to be very useful, since the Fix and Hodges approximation for the Mann-Whitney tail probability \( P_0(M_{XY} > u) = FH(u) \), say, is an explicit function of \( u \).
Therefore, to get an initial guess for $U_{mn}$, we simply equate the inverse of the false alarm rate to 500, which means solving $1/(2\times FH(u)) = 500$. Let $UCL^{(1)}$ denote the solution of this equation. Now we evaluate $ARL_0$ at $UCL^{(1)}$ by applying formula (5) and the LR-approximation for computation of $p_U(\bar{X})$, which results in a new $ARL_0$, say, $ARL_0^{(1)}$. If $ARL_0^{(1)} > 500$, we lower the value of $UCL^{(1)}$ by a certain amount $s$, say $s = 10$, to obtain a new guess, say $UCL^{(2)} = UCL^{(1)} - 10$ and we again evaluate formula (5) with this value to find $ARL_0^{(2)}$. Analogously, we use $UCL^{(2)} = UCL^{(1)} + 10$ when $ARL_0^{(1)} < 500$. Since we now have two guesses we can use linear interpolation or extrapolation (when both guesses are on the same side) to find a new guess. Note that since the MW statistic can take only integer values between 0 and $mn$, we need to search only over these values; we also use the fact that $ARL_0$ is strictly increasing in $U_{mn}$ (which can be observed from formulas (2) and (3), because $p_U(\bar{X})$ is decreasing in $U_{mn}$). The search procedure is repeated until we find the value for which the corresponding $ARL_0$ is sufficiently close to 500. This is the desired upper control limit $U_{mn}$.

The definition of “sufficiently close” depends on the user. We allow for defining a target interval of the form $500 \pm \lambda \times 500$. For example, when $\lambda = 0.02$ the iterative procedure stops at the $i^{th}$ step if $490 \leq ARL_0^{(i)} \leq 510$, which we believe is close enough. If the user specifies $\lambda = 0$, the procedure stops when it has produced two values $UCL^{(i)}$ and $UCL^{(j)} = UCL^{(i)} + 1$ for which $ARL_0^{(i)} \leq 500$ and $ARL_0^{(j)} \geq 500$. The user should then choose either $UCL^{(i)}$ or $UCL^{(j)}$. Of course, the larger $\lambda$ is, the faster a solution is found. From Table 2 we observe that the 4th approximation, labeled the fixed reference sample approximation, is somewhat better than the 5th approximation for $m \leq 500$. Therefore, before applying the linear interpolation method with the LR-formula, we use it with the fast fixed reference sample approximation to refine the Fix-Hodges approximation, when $m \leq 500$. The resulting UCL approximation is then used as an initial value for the linear interpolation method with the LR-formula. As an example, suppose we have $m = 375$ and $n = 7$. We want to find the chart constants such that $ARL_0 \approx 400$. Suppose we allow $\lambda = 0.02$, hence the search procedure would stop and yield the desired control limit.
When \( 392 \leq \hat{ARL}_0 \leq 408 \). We also have to specify the maximum Monte Carlo standard error \( D \) that we wish to allow for, which determines the number of samples \( K \), by applying criterion (9). Suppose we require \( s_{mc} \leq D = 0.015*\hat{ARL}_0 \approx 6 \). Note that \( K \) depends on \( s(ARL_0(X)) \), which is re-evaluated after each new sample is obtained, so an explicit number cannot be given \textit{a priori}. The output from our program (written in Mathematica; see Software section later) is shown in Figure 1.

\begin{verbatim}
1/(false alarm rate) approximation
1. ucl=2146 lcl=479 ARL0=400

Fixed reference sample approximation
2. ucl=2146 lcl=479 ARL0=446.761
3. ucl=2136 lcl=489 ARL0=386.729

LR approximation
4. ucl=2136 lcl=489 ARL0=380.059 smc=5.69018 5% perc=238.407 K=402
5. ucl=2146 lcl=479 ARL0=438.11 smc=6.5647 5% perc=287.53 K=319
6. ucl=2139 lcl=486 ARL0=394.496 smc=5.91419 5% perc=252.778 K=315

139.962 Second, Null
\end{verbatim}

\textbf{Figure 1.} Finding control limits for \( m = 375, n = 7 \) and target \( ARL_0 = 400 \)

As can be seen in Figure 1, the computer has executed six iterations (numbered 1 through 6) in approximately 140 seconds. The first three iterations hardly take any computing time. At each iteration, the value of the \( UCL \) (and the \( LCL \)) being used and the corresponding \( ARL_0 \) are shown. Under the LR approximation, the program also calculates the 5\(^{th} \) percentile of the conditional in-control \( ARL \) distribution and the Monte Carlo standard error \( (smc) \). It is seen that with the LR method, the first two iterations (4 and 5) produce \( ARL_0 \) values below and above the target value 400, so that linear interpolation begins at the third iteration and the new \( UCL \) is found using the two previous \( UCL \)’s and their corresponding \( ARL_0 \) values. Thus, the final chart constants are found to be 486 and 2139, with attained \( ARL_0=394.5 \). Note that in each iteration, \( K \) is chosen such that the standard error of the estimate is smaller or equal than 1.5\% of the estimate, which guarantees \( s_{mc} = 5.91 < 6 \), as stipulated. Finally, from iteration 6, the 5\(^{th} \) percentile of the conditional in-control \( ARL \) distribution is 252.78, so that using the MW chart with \( UCL = 2139 \) and \( LCL = 486 \), the user has the assurance that 95\% of (all possible) reference samples will produce a (conditional) \( ARL_0 \) greater than 252.78. This is akin to making a “confidence” statement about a control chart when parameters are estimated and should be
informative to potential users who most likely would have only one (their own) reference sample.

**Numerical Example**

Montgomery (2001, Table 5.1) gives a set of data on the inside diameters of piston rings manufactured by a forging process. Twenty-five samples, each of size five, were collected when the process was thought to be in-control. The traditional Shewhart $\bar{X}$ and R charts provide no indication of an out-of-control condition, so these “trial” limits were adopted for use in on-line process control.

For the proposed MW chart we find, using our program with $m = 125, n = 5$ and $ARL_0 = 400$, the upper control limit $U_{mn} = 540$ and the lower control limit $L_{mn} = 85$. Having found the control limits, prospective process monitoring begins. Montgomery also gives (Table 5.2) fifteen additional samples from the piston-ring manufacturing process. These “test samples” led to fifteen MW statistics calculated using Minitab: 429.0, 333.0, 142.5, 370.5, 241.5, 410.5, 393.0, 240.5, 471.0, 486.0, 340.5, 561.0, 575.5, 601.5 and 484.5. Comparing each statistic with the control limits, all but three of the test groups, 12, 13 and 14 are declared to be in-control. The control chart is shown in Figure 2.

![MW Chart](image)

**Figure 2.** MW Chart for the Piston-ring data

The conclusion from this chart is that the medians of test groups 12, 13 and 14 have shifted to the right in comparison with the median of the in-control distribution. It may be noted that the
Shewhart $\bar{X}$ chart in Montgomery (2001) for these data lead to the same conclusion with respect to the means. Of course, the advantage with our chart is that it is distribution-free with an in-control $ARL_0 = 400$ so that regardless of the underlying distribution, the in-control $ARL$ of the chart is equal to 400 and thus there is no need to worry about (non-) normality, as one must for the $\bar{X}$ chart. We also calculated the distribution-free precedence chart proposed in CVV for comparison. For this chart we found $LCL = 73.982$ and $UCL = 74.017$, with an attained $ARL_0 = 414.0$. Consequently, the precedence chart declares the 12th and the 14th groups to be out of control but not the 13th group, unlike both the MW and the Shewhart chart. This is not entirely surprising since the MW test is generally more powerful than the precedence test.

The question of ties: it may be noted that the data contain some cases where a $Y$ observation is exactly equal to an $X$ observation, a situation commonly referred to as a “tie” in the literature. As was noted before, in theory (derivation of the limits etc.) ties can not occur with positive probability (since the distributions are assumed to be continuous) but in practice ties do occur, for example, due to the limits of resolution on the measuring instrument or the recording device. This should not be a major problem unless there are an excessive number of ties in which case perhaps the whole notion of a variables control chart might be somewhat questionable. A conservative alternative is to not count any contribution from the pair of tied observations consistent with the definition of the indicator function, $I(X_i < Y_j)$, while calculating the $U$ statistics. We expect the differences to be small. For example, for the present data, these “untied” MW statistics are 405,323,134,363,232,401,382,231,460,476,332,554,570, 600 and 474 and we reach the same conclusion that was reached earlier on the basis of the “tie-broken” MW statistics obtained from MINITAB.

**Control Chart Performance**

The performance of a control chart is usually assessed in terms of its run length distribution and certain associated characteristics, such as the $ARL$. While we examine the $ARL$, we consider two other criteria for evaluating the performance of the MW chart and its parametric competitor, the Shewhart $\bar{X}$ chart. To ensure a fair comparison, first, the Shewhart $\bar{X}$ chart is used when both the mean and the variance are unknown, so these parameters are estimated from the reference sample. Second, the charts are both designed to have the same specified $ARL_0$. 
Note that for the $\bar{X}$ chart the non-robustness of the $ARL_0$ with respect to non-normal in-control distributions is a major concern. This has been recognized as a problem elsewhere (see e.g. CVV) and is perhaps one of the important reasons for considering a nonparametric chart. The distribution-free property (and the resulting robustness of the $ARL_0$) is an obvious advantage for the proposed chart, but what about its out-of-control performance? We try to address this issue here. Our main interest is in shift alternatives $G(x) = F(x - \delta)$, where $\delta$ is the unknown shift parameter. This means that the test samples are possibly from a population with say median $\delta$, whereas the reference sample is from a population with median 0 (without any loss of generality).

For the MW chart, one would expect a longer out-of-control $ARL$ against the normal shift alternatives. On the other hand, based on the hypothesis testing literature, the MW chart is expected to be more efficient (shorter out-of-control $ARL$) for heavy-tailed distributions. Note that in the fortunate situation in which the in-control distribution (e.g. normal with a specified mean and standard deviation) is completely known, one can guarantee that the in-control $ARL$ is really 500 (say). The same holds for the out-of-control $ARL$ when the user specifies the shift. However, this is not the case when parameters are unknown and a reference sample is used from the in-control process. To better understand the effect of using a reference sample on the chart’s performance, we use the concept of conditional $ARL$ introduced earlier. Let $ARL_0(X) = E\left(U | N \mid X\right)$ and $ARL_{\delta}(X) = E\left(F(x - \delta) | N \mid X\right)$, denote the in-control and out-of-control expected values of the run-length distribution, given the reference sample $X$. Clearly, $ARL_0(X)$ can differ substantially from the (unconditional) target $ARL_0$ and from user to user, each with a possibly different reference sample from the same in-control process. Therefore, it is important to judge the effect of estimation, that is, using a reference sample, on both $ARL_0(X)$ and $ARL_{\delta}(X)$, for both the MW chart and the Shewhart chart (with estimated mean and standard deviation). From this point of view, the three criteria listed below can serve as a reasonable basis for comparing the performance of the proposed chart with that of the Shewhart $\bar{X}$ chart.

1. The out-of-control (unconditional) $ARL$ for a specified shift under the location model $G(x) = F(x - \delta)$. This is a traditional chart comparison criterion. We denote this quantity by $ARL_{\delta}$. Naturally, small values of $ARL_{\delta}$ are desirable for a preferred chart.
2. The 5th percentile of the in-control distribution of the random variable $ARL_0(X)$, which is simply the in-control $ARL$ given or conditioned on the reference sample. Relatively large values of this percentile are desirable for a preferred chart, which is in the same spirit as desiring the in-control $ARL$ of a chart to be large, since that corresponds to a smaller probability of a false alarm. The conditional in-control $ARL$ percentiles provide valuable information on the in-control performance of a chart when parameters are estimated from a reference sample. The (unconditional) $ARL_0$ is an average of the conditional $ARL$ over all possible reference samples and in practice users do not have the benefit of averaging. For the proposed chart, the 5th percentile is found by drawing a reference sample $X$ from $F$ (recall that in the in-control case $F$ is taken to be the cdf of the uniform(0,1) distribution) $K$ times. For the $i$th sample, we calculate $p_F(\bar{x}_i)$ and $ARL_0(\bar{x}_i) = 1_0 p_F(\bar{x}_i)$, $i = 1, \ldots, K$, sort the list $\{ARL_0(\bar{x}_1), \ldots, ARL_0(\bar{x}_K)\}$ in ascending order, and simply read off the (0.05*$K$)th value from the bottom. Denoting this quantity by $q_{0.05}$, we can say that 95% of the values of the distribution of the conditional in-control average run length exceed $q_{0.05}$, which means that relatively large values of $q_{0.05}$ are desirable for a preferred chart. Again, as noted earlier, specifying this conditional percentile can be viewed as a “confidence” statement about the performance of a control chart when parameters are estimated.

3. The 95th percentile (or the upper 5th percentile) of the out-of-control distribution of the random variable $ARL_d(X)$, which is the conditional out-of-control $ARL$. Denoting this by $q_{0.95}$, relatively smaller values of $q_{0.95}$ are desirable for a preferred chart, since then the probability of a signal is higher in the out-of-control case. Again, the motivation behind this is from a practical point of view as explained in criterion 1; the users have the assurance that the out-of-control $ARL$ is no larger than a specified value more than 5 percent of the time.

It may be noted that others have examined percentiles of the run length distribution as a performance criterion in the literature. For example, Jones (2002) examined the 10th and the 90th percentiles of the run length distribution for an EWMA chart. We, however, consider the percentiles of the conditional distribution, for reasons stated above, in criteria 2 and 3.

Consider the second criterion first, which involves the in-control situation. For a fair comparison between the MW chart and the Shewhart $\bar{X}$ chart with estimated parameters, chart constants were determined such that the $ARL_0$ approximately equals 500 for both charts. We
kept the test sample size constant, \(n = 5\), and used several values for the reference sample size \(m\).

The results are shown in Table 3.

Table 3. Upper Control Limits and 5\(^{th}\) percentiles of the conditional in-control distribution of \(ARL_0(X)\); All cases: \(n = 5\) and \(ARL_0 = 500\)

<table>
<thead>
<tr>
<th>(m)</th>
<th>Upper Control Limit ((U_{mn})) MW</th>
<th>5(^{th}) percentile MW</th>
<th>Upper Control Chart Constant Shewhart</th>
<th>5(^{th}) percentile Shewhart</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>217</td>
<td>97</td>
<td>3.01996</td>
<td>49</td>
</tr>
<tr>
<td>75</td>
<td>326</td>
<td>146</td>
<td>3.05156</td>
<td>87</td>
</tr>
<tr>
<td>100</td>
<td>435</td>
<td>182</td>
<td>3.06535</td>
<td>112</td>
</tr>
<tr>
<td>150</td>
<td>654</td>
<td>251</td>
<td>3.07715</td>
<td>154</td>
</tr>
<tr>
<td>300</td>
<td>1304</td>
<td>284</td>
<td>3.08607</td>
<td>232</td>
</tr>
<tr>
<td>500</td>
<td>2172</td>
<td>322</td>
<td>3.08848</td>
<td>270</td>
</tr>
<tr>
<td>750</td>
<td>3258</td>
<td>360</td>
<td>3.08935</td>
<td>314</td>
</tr>
<tr>
<td>1000</td>
<td>4347</td>
<td>379</td>
<td>3.08969</td>
<td>338</td>
</tr>
<tr>
<td>1500</td>
<td>6520</td>
<td>409</td>
<td>3.08996</td>
<td>367</td>
</tr>
<tr>
<td>2000</td>
<td>8691</td>
<td>420</td>
<td>3.09007</td>
<td>376</td>
</tr>
</tbody>
</table>

For example, when \(m = 750\), while applying the MW chart with \(U_{mn} = 3258\), a user knows that 95% of the in-control \(ARL\)’s (for a large number of reference samples taken from the in-control process) will be at least 360. We believe this provides useful information in addition to saying that the in-control \(ARL\) equals 500 for \(U_{mn} = 3258\). For the \(\bar{X}\) chart with \(m = 750\) and a control chart constant of 3.089 (that guarantees \(ARL_0 = 500\)), the 5\(^{th}\) percentile is 314 so that 95% of the in-control \(ARL\)’s are at least 314. Since the in-control 5\(^{th}\) percentiles for the MW chart are larger than those of the Shewhart chart for all \(m\), we conclude that based on this criterion, the MW chart performs better than the Shewhart chart with estimated limits, particularly for \(m \leq 150\). Note that as \(m\) increases both percentiles approach the mean of the corresponding unconditional distribution, the \(ARL_0\), which is 500. Table 3 also suggests that, especially for small values of \(m\), it may be better to determine the control limits of the chart such that the 5\(^{th}\) percentile of the conditional distribution is larger than some specified large number, such as 300. The software we provide allows the user to calculate such control limits. More information about software is given later.

The first and the third criteria, which deal with the out-of-control performance of the charts, are examined in the case of three process distributions: Normal, Laplace and Gamma(2,2). The Laplace distribution is normal like but with heavier tails, which results in higher probabilities of extreme values. The Gamma(2,2) distribution is skewed and is often used...
We applied two-sided charts to the Normal and the Laplace distributions and a one-sided chart with an upper control limit to the Gamma(2,2) distribution. The test sample size $n$ is 5 and the reference sample size $m$ is 100. Control limits for both the MW chart and the Shewhart $X$ chart with estimated parameters are determined such that $ARL_0 = 500$. Using these limits the conditional out-of-control $ARL$, that is $ARL_\delta (X)$ and the 95th percentile of the distribution of $ARL_\delta (X)$ are computed for several values of the shift $\delta$, which is given in units of the standard deviation. Figures 3 through 5 show the results.

In Figure 3, we observe (the set of points: triangles and diamonds) that the 95th percentiles for the Shewhart $\bar{X}$ chart are all smaller than those for the MW chart. Thus, as one might expect, the $\bar{X}$ chart is more effective in detecting shifts than the MW chart in case of the normal alternative. However, the differences between the percentiles are small at all shifts (the largest difference is around 15) and the difference appears to vanish for shifts greater than 1. The same pattern holds for the two $ARL$'s. On the other hand, Figure 4, for the Laplace distribution, shows that the MW chart is clearly better than the Shewhart chart for all shifts, large and small. In fact, in this case, the 95th percentile of $ARL_\delta (X)$ for the MW chart is even below the average out-of-control run length (ARL) for the Shewhart chart. For the Gamma(2,2) distribution, in Figure 5, again, we see that the MW chart is better in detecting shifts, although the difference in performance is not as dramatic as in the case of the Laplace distribution. These calculations were repeated for $m = 500$ observations; the results were very similar and are therefore omitted here. We conclude that the nonparametric MW chart follows the well-known results for the MW test statistic: it is nearly as effective as the Shewhart $\bar{X}$ chart under normality, but is more effective under heavy tailed and skewed distributions. Also, note that performance of the MW chart in the case of the Laplace distribution makes it potentially useful when outliers in the data are not uncommon.

Summary and Conclusions

A nonparametric control chart is proposed for process location that maintains its in-control properties and thus can be justifiably used for any continuous process distribution. Control limits (as well as the software) for the proposed chart are provided for practical implementation. Comparisons based on 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 $\bar{X}$ chart when the process is normal, but is more effective than the Shewhart $\bar{X}$ 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) $ARL_0(X)$ distribution can be a useful chart design criterion, in addition to the traditional (unconditional) $ARL_0$. Readers can use the interactive software, available on a website, to implement the proposed chart using either criterion, with their own data. Practitioners are encouraged to use more nonparametric control charting methodology.

**Appendix: Software**

In order to support practical implementation of the methods presented in this paper two types of software resources are provided. First, a Mathematica 4.2 (Wolfram, 1999) notebook is available containing the following:

- Approximation of $ARL_0$ and percentiles of (conditional) $ARL_0(X)$ for specified chart limits using either exact computation of $p_U(\bar{x})$ or the LR-formula.
- Fast approximation of $ARL_0$ with either the reciprocal of the false alarm rate or with fixed reference sample.
- Approximation of out-of-control $ARL_0(\delta)$ and percentiles of (conditional) $ARL_0(\delta)(X)$ for shift alternatives $G(x) = F(x - \delta)$.
- Determination of MW control chart limits for specified target $ARL_0$ using the procedure as described in the section ‘Design and Implementation’.
- Determination of MW control chart limits for specified $q$th percentile of $ARL_0(X)$ using a similar procedure.
- Plotting of the MW control chart for a user-specified data set.

Second, we created a website that enables anyone interested to apply the proposed methodology. The MW control chart limits can be found for the sample sizes at hand for a specified target $ARL_0$ (or a desired ($q$th) percentile of $ARL_0(X)$). Moreover, the website allows users to import their own data set and draw the MW chart. The site can be reached via [www.win.tue.nl/~markvdw](http://www.win.tue.nl/~markvdw). The Mathematica notebook is available from the same site. The
advantage of the notebook is that it contains the complete programming code and more functions and it is very flexible in its input. However, the plus for the website is the ease of use and more importantly the fact that a Mathematica licence is not required. User instructions are available in the notebook and at the website. We sincerely hope that this arrangement will enhance the use of this distribution-free control chart.

Acknowledgement

We thank Marko Boon of the department of Mathematics and Computer Science, Eindhoven University of Technology, for assistance with the design of the Mann-Whitney control chart website. This research was supported in part by a summer research grant from the University of Alabama and by a grant from the Thomas Stieltjes Institute for Mathematics.

References


Figure 3. Performance for MW chart and Shewhart chart under Normal shift alternatives

Figure 4. Performance for MW chart and Shewhart chart under Laplace shift alternatives

Figure 5. Performance for MW chart and Shewhart chart under Gamma(2,2) shift alternatives