

Nonparametric statistical process control : an overview and some results

Citation for published version (APA):

Chakraborti, S., Laan, van der, P., & Bakir, S. T. (1999). *Nonparametric statistical process control : an overview and some results*. (Memorandum COSOR; Vol. 9908). Technische Universiteit Eindhoven.

Document status and date:

Published: 01/01/1999

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.



Eindhoven University
of Technology

Department of Mathematics and Computing Sciences

Memorandum COSOR 99-08

Nonparametric statistical process control: an overview and some results

S. Chakraborti
P. van der Laan
S.T. Bakir

Eindhoven, June 1999
The Netherlands

Nonparametric Statistical Process Control: An Overview and Some Results

S. Chakraborti¹
Management Science and Statistics
University of Alabama
Tuscaloosa, AL, U.S.A.

P. van der Laan²
Mathematics and Computing Science
Eindhoven University of Technology
Eindhoven, The Netherlands

S. T. Bakir
College of Business
Alabama State University
Montgomery, AL, U.S.A.

ABSTRACT

An overview of the literature on some nonparametric or distribution-free quality control procedures is presented for univariate data. A nonparametric control chart is defined along with some general motivations and formulations. Various advantages of these charts are highlighted while some disadvantages of the more traditional, distribution-based, control charts are pointed out. Specific observations are made in the course of the review of articles and constructive criticism is offered, so that opportunities for further research can be identified. Connections to some areas of active research are made, such as sequential analysis, that are of relevance to process control. It is hoped that this article would lead to a wider acceptance of distribution-free control charts among the practitioners and would serve as an impetus to future research and development in this area.

KEY WORDS AND PHRASES: *Control charts, Shewhart, CUSUM, EWMA, Sequential, Detection and Change-Point Methods, Signs, Ranks, Distribution-free.*

1. Introduction

One of the primary objectives of statistical process control is to distinguish between two sources of variation in a given process, those which cannot be economically identified and corrected (chance causes) and those which can be (assignable causes). When a process operates only under chance causes, it is said to be in a state of statistical control (hereafter in-control). Control charts help researchers identify and eliminate assignable causes so that the state of statistical control can be

ensured. Intuitively, in the event there is a change in the process, a control chart should detect it as quickly as possible and give an out-of-control signal. Clearly, the quicker the detection and the signal, the more efficient is the chart. The number of samples or subgroups that need to be collected before the first out-of-control signal is given by a chart is a random variable, called the run length. The distribution of the run length is often used to characterize the efficacy or the performance of a chart. A popular measure of performance is the expected value or the first moment (about 0) of the run length distribution, called the average

¹ On leave at the Department of Biostatistics, University of North Carolina, CB#7400, Chapel Hill, NC 27599. Dr. Chakraborti's work was supported in part by a 1998 summer research award from the College of Commerce and Business Administration, The University of Alabama and by a NATO Collaborative Research Grant CRG 920287. e-mail:schakrab@cba.ua.edu.

² Dr. van der Laan's research was supported in part by the NATO Collaborative Research Grant CRG 920287.

run length (ARL), although several authors currently suggest examining other characteristics of the run length distribution, such as the second central moment or the variance. By definition, the run length is a positive integer valued random variable so that the ARL loses much of its attractiveness as a typical summary, if the distribution is skewed (as is often the case). As a consequence, other measures, such as the median, are sometimes considered. It is desirable (often stipulated) that the ARL be large when the process is in-control, whereas the exact opposite should be the case when the process is out-of-control. The false alarm rate is the probability that a chart signals a process change when in fact there hasn't been any, that is, when the process is in-control. This is synonymous with the probability of a type-I error in a hypothesis testing context. Two control charts are often compared, prospectively, such that their in-control ARL's are roughly the same. Again, this parallels comparing two statistical tests on the basis of power, against some alternative hypothesis, when they are roughly of the same size.

In the context of process control, often, the pattern of chance causes is assumed to follow some parametric distribution. The most common assumption in the literature is that the chance distribution is normal. The statistical properties of the usually employed control charts are exact only if this normality assumption is satisfied. However, the underlying process is not normal in many applications, and as a result the statistical properties of the standard charts could be highly affected in such situations. On this point see, for example, Shewhart (1939; p. 12, 54), Ferrell (1953), Tukey (1960; p. 458), Langenberg and Iglewicz (1986), Jacobs (1990), Alloway and Raghavachari (1991) and Yourstone and Zimmer (1992). In addition, normal-like but heavier-tailed distributions also

occur in practice; for more details refer to Noble (1951), Tukey (1960; p. 458), Lehmann (1983; p. 365) and Gunter (1989). These authors and others, including practitioners, provide ample justifications for the application of distribution-free or nonparametric techniques in statistical process control. For clarification, it should be noted that the term nonparametric is not intended to imply that there are no parameters involved, in fact quite to the contrary. This is not always clear, particularly to the practitioners. In this paper both terms, distribution-free (hereafter DF) and nonparametric (hereafter NP), will be used to emphasize the fact that they are the same.

In spite of the weight of the evidence, however, development and implementation of NP methods have not been commonplace in industrial process control. There might be a multitude of reasons behind this. Practitioners sometimes have felt that the central limit theorem would "come to the rescue" and somehow render the charts "correct." While this might be true for some control charts based on averages of certain statistics from processes that are "well-behaved," it is far from being true in general. More importantly, in the problem where control charts are to be applied to individual observations (see for example, Montgomery, 1991) the central limit theorem can not be invoked (since the sample size is one). It has been shown that in this case the standard charts lack distribution-robustness (Lucas and Croisier (1982), Rocke (1989)). Other reasons for the apparent lack of interest might have included past unavailability of adequate "in the field" computing facilities and the perception that one has to sacrifice "efficiency" when using these "simple" techniques based often on counting and ranking. The former is no longer a problem in today's computer age and the latter isn't necessarily true as has been well documented in the statistical testing and estimation literature. In fact, it has been known, for a long time, that for many heavy-tailed distributions, common NP methods outperform their parametric counterparts. Moreover, when the underlying

distribution is truly normal, the efficiency of some NP methods, relative to the corresponding (optimal) normal theory methods can be as high as 0.955. Finally, to be fair, it should be noted that a large part of the developments in nonparametric methodology have taken place in the classical confines of statistical estimation and hypothesis testing and not much effort has been made to understand the problems of “practical statistical process control.”

The main advantage of NP procedures is, of course, the flexibility that one doesn't need to assume any parametric probability distribution for the underlying process, at least as far as establishing and implementing the procedures are concerned. Obviously, this could be a big plus in the field of process control, particularly in start-up situations, where not much data might be available to use a parametric (for example, normal theory) procedure. Also, NP control charts are likely to share the robustness properties of NP tests and confidence intervals and are, therefore, likely to be less impacted by outliers.

A formal definition of a NP control chart is given as follows.

Definition: Let ϑ be the class of continuous cumulative distribution functions. A control chart is NP or DF over the class ϑ , if the in-control run length distribution is the same for every member of ϑ .

To summarize, advantages of NP control charts include:

1. Simplicity.
2. No need to assume a particular distribution for the underlying process to set up a chart.
3. The in-control run length distribution is the same for all members of ϑ . The same is true for the false alarm rate. Thus

different NP charts can be compared more easily.

4. More robust and outlier resistant.
5. More efficient in detecting changes, when the true distribution is markedly non-normal, particularly with heavier tails.
6. No need to estimate the variance to set up charts for the location parameter.
7. Useful in start-up and short-run applications, allowing implementation earlier in the product life cycle.

In this paper we provide a framework for NP statistical process control (hereafter NSPC), so that the objectives as well as the problems are more easily understood. Within this framework, an overview of the literature, mainly on univariate methods, is presented. Not all papers on the subject could be included in this review since in order for the paper to be of a reasonable length, a choice had to be made so that some of the important advances can be surveyed. In course of the review, some constructive criticism is offered wherever applicable, so that opportunities for further research can be identified. It is hoped that these observations would generate more questions, comments and discussions so that the advantages (and the disadvantages) of these simple methods can be better understood and more fully appreciated. Note that we consider only the so-called “variables control charts” since most NP procedures require a continuous population to be DF, at least for finite sample sizes. Finally, although multivariate process control problems are important in their own right, very few multivariate NSPC techniques are currently available, and these will be covered elsewhere.

2. Terminology and Problems

An important problem in the quality literature is the problem of tracking a process mean. More generally however, one can consider tracking the center or a location (or a shift) parameter. For example, the location parameter could be the mean or the median or some percentile of the distribution. The latter are especially attractive when the

distribution is skewed. Also, many processes are implicitly assumed to follow either a pure location (or a shift) model, say of the form $F(x - \theta)$ where θ is an unknown (location) parameter (corresponding to say a normal distribution with unknown mean and known variance), or a pure scale model, say of the form $F(x/\tau)$ where $\tau (>0)$ is the unknown (scale) parameter (corresponding to say a normal distribution with unknown variance and known mean), and $F \in \mathcal{F}$, is some underlying continuous cumulative distribution function (cdf). Sometimes one might be interested in the location-scale model: $F\{(x - \theta)/\tau\}$, where both θ and τ are unknown parameters (corresponding to say a normal distribution with both mean and variance unknown). Under these (often implicit) model assumptions, the problem is to track θ , or τ (or both), based on random samples taken (usually) at equally spaced time points. As noted earlier, in the usual control charting problems F is assumed to be the cdf of the standard normal distribution, in the NP setting, for variables data, it is enough to say that F is some arbitrary continuous cdf. Although the location-scale model seems to be a natural model to consider paralleling the normal case with mean and variance both unknown, most of what is currently available in the NSPC literature deals only with either the pure location or the pure scale model.

The starting point for designing a control chart is usually a "control statistic", which is often an estimator of the parameter of interest (see e.g., Montgomery, 1991; page 105). Traditional control statistic for the mean is the sample mean (\bar{X}), whereas for the process variation one uses the sample variance S^2 (or the sample standard deviation S) or the sample range R . One problem with these statistics is that although in the normal

case their distributions are either well known or derivable, in general these are not DF (i.e. the end result depends on the fact that the original distribution is normal) unless the sample size is large. In fact, the lack of distribution-robustness (even for moderate sample sizes) is a concern, particularly for the S and the R charts. Thus, unless the process is known (say normal) or the sample sizes are quite large, the false alarm rates for the standard parametric charts can be (unacceptably) high.

While constructing NP charts, it seems natural, as a first step, to consider replacing these parametric control statistics with other reasonable statistics that are DF and study analogs of the parametric charting methods. This will allow computation of control limits etc. that are valid for a whole class of distributions. It turns out, however, that in the NP (or robust) charting setting, the well-known estimators are often not DF for finite sample sizes. Accordingly, one then has to resort to NP tests (often there is a correspondence between the tests and the estimators) and adapt those for the control charting problem. This is what has been mostly done so far in the literature and some of the contributions based on this idea will be reviewed in the next section.

Recall that the most common quality control charting methods include the Shewhart, the cumulative sum (CUSUM) and the exponentially weighted moving average (EWMA) with various proposed refinements. When tracking the process mean, the control statistic used in these charts is the sample mean (although (robust) variations have been considered) whereas for tracking the process variation the choice is usually between the sample standard deviation or the sample range. The relative advantages and disadvantages of these charts are well documented (see, e.g., Montgomery, 1991).

3. A Review of Literature

While Shewhart-type charts are the most widely used because of their simplicity, CUSUM procedures are

quite natural in view of the sequential nature of the process control problem. In the normal theory (parametric) setting Page (1954) proposed CUSUM charts based on the sample mean. In the NP setting Reynolds (1975) studied charts based on “signed sequential ranks” of observations. McGilchrist and Woodyer (1975) considered a CUSUM technique that allows for DF tests and applied it to the problem of detecting a change in the median of a rainfall distribution. However, this is a problem in hydrology and not in process control.

Bakir and Reynolds (1979) (hereafter BR) proposed a CUSUM chart based on the Wilcoxon signed-rank (WSR) statistic to track the shift of θ (a location parameter) from an in-control known value θ_0 (assumed equal to 0, without any loss of generality). We discuss the BR paper in some detail since the same basic ideas can be and has been used in the literature with other DF statistics.

The WSR test is a well known (see for example, Gibbons and Chakraborti, 1992) NP test and a competitor to the classical one-sample t test, for testing hypotheses or setting-up a confidence interval about the location parameter θ of a continuous distribution symmetric about θ . Typically in control charting, $m = 20$ to 25 random samples (groups) are taken, sequentially from the process, each of size $g = 4$ to 5 observations. Let (X_{i1}, \dots, X_{ig}) , $i = 1, 2, \dots, m$, denote the i th random sample. The BR procedure is based on the idea of ranking observations within the i th sample or group. The idea of “within group ranking” has been employed earlier by Wilcoxon, Rhodes and Bradley (1963) and Van der Laan (1966) to develop NP sequential two-sample tests. Let R_{ij} be the rank of $|X_{ij}|$ among $(|X_{i1}|, \dots, |X_{ig}|)$, $j = 1, 2, \dots, g$, $i = 1, 2, \dots$ and let $SR_i = \sum_{j=1}^g \text{sign}(X_{ij})R_{ij}$,

where $\text{sign}(x)$ is 1, 0 or -1 , according as x is $>$, $=$, or $<$ 0. The statistic SR_i is linearly related to the more well-known WSR statistic V_g , the sum of the ranks of the positive observations, through the relation $SR_i = 2V_g - g(g+1)/2$. Thus the “in-control” probability distribution (mass function) of SR_i can be obtained from the “null” distribution of V_g . Assuming that none of the X_{ij} is equal to 0, (an event with probability 0 in \mathfrak{D}), the latter has been tabulated by several authors, the table by Wilcoxon, Katti and Wilcox (1972) being one of the most extensive. The typical CUSUM chart for the mean is based on the cumulative sum of the sample means. The BR grouped signed-rank (GSR) procedure uses the SR statistics with a CUSUM type stopping rule. Clearly, the procedure is DF since the in-control ($\theta=0$) distribution of the V_g and hence the SR_i statistics don't depend on the underlying distribution for all continuous symmetric distributions. The one-sided procedure for detecting a positive deviation in θ , from the in-control value $\theta_0 = 0$, signals at the first n for which

$$\sum_{i=1}^n (SR_i - k) - \min_{0 \leq m \leq n} \sum_{i=1}^m (SR_i - k) \geq h. \quad (1)$$

The corresponding procedure for detecting a negative shift in the mean signals at the first n at which

$$\max_{0 \leq m \leq n} \sum_{i=1}^m (SR_i + k) - \sum_{i=1}^n (SR_i + k) \geq h. \quad (2)$$

A two-sided symmetric procedure signals at the first n for which either of the two inequalities is satisfied.

The two parameters of the CUSUM chart are the reference value k and the decision value h . One criterion for the optimal choice of (k, h) is that the combination minimizes the ARL of the procedure when the process mean has shifted, subject to the condition that the in-control ARL be a specified value. It can be shown that for varying large values of n , the behavior of the cumulative sum process can be approximated by a Brownian motion process. Hence, as in Reynolds (1975), the optimal value of k is approximately equal to $\theta/2$, where $n\theta = n\theta(\Delta)$ is the mean of the

cumulative sum, $\sum_{i=1}^n SR_i$, corresponding to the shift

Δ . The expression for $\theta(\Delta)$ is obtained from the mean of the WSR statistic based on g observations. Tables are given for the optimal values of k corresponding to different shifts in location, when the parent distribution is uniform, normal, double-exponential and Cauchy, respectively. These are normal-like distributions with different tail properties. It was recommended that the optimal k values obtained for the normal distribution be used in practice unless very heavy tails are indicated. Using this value of k , the value of h is then chosen to achieve the desired one-sided in-control ARL value. Tables are given for the one-sided ARL values for various combinations of h , k and g . For example, when observations are collected in groups of size 6, using $h=10$ and $k=11$ yield an in-control ARL = 301.01. Comparisons are made, on the basis of the "exact" one-sided ARL, with the Shewhart chart and the usual CUSUM chart under various positive shifts when the process is normally distributed. For non-normal distributions such as the uniform, the double exponential and the Cauchy, comparisons were made on the basis of simulated one-sided ARL values for various positive shifts. The overall conclusion is that when observations are naturally collected in groups, the GSR-CUSUM chart is only slightly less efficient than the usual CUSUM chart based on the sample mean when the process is normally distributed, whereas for non-normal distributions the GSR-CUSUM chart can be considerably more efficient. A suitable group size (n) for this NP procedure was suggested to be between $n=5$ and 10, depending on the shift-size and the desired in-control ARL. This recommendation is nearly the same as the group size recommended for the normal theory based procedures.

Arnold (1985) presented a NP test procedure. He selected the sign test because its power is easily obtained. It is assumed that the production speed is constant and equal to v items per time unit. The control chart works as follows: Every T time units of production a sample of n items is taken; let x_i denote the value of the i -th item in the sample. Since it is assumed that the characteristic variable X is continuously distributed, we have that $x_i \neq z$ for a certain value z , e.g., 0 for all i with $i = 1, 2, \dots, n$ (almost surely). If the number K of the x_i with $x_i < z$ is at most c or at least $n-c$ a search and, if necessary, a repair is undertaken. Otherwise the process is continued. Comparison of control chart is made considering several economic parameters.

Park and Reynolds (1987) developed NP procedures for monitoring the location parameter of a continuous process when the control value for the parameter is not specified. These procedures are based on the so-called linear placement statistics, introduced earlier by Orban and Wolfe (1982) for comparing current samples with a standard sample taken when the process was operating properly. The linear placement statistics are used in versions of Shewhart and CUSUM charts. Asymptotic approximations to the run length distributions are obtained.

McDonald (1990) considered a CUSUM procedure based on what are called "sequential ranks". The sequential rank R_i of an observation X_i is defined as $R_i = 1 + \sum_{j=1}^{i-1} I(X_j < X_i)$, where $I(\cdot)$ is the usual indicator function, and a CUSUM chart is based on $U_i = R_i/(i+1)$, $i=1, 2, \dots$. When the process is in-control, the U_i are independent random variables, uniformly distributed on $\{1/(i+1), 2/(i+1), \dots, i/(i+1)\}$. Thus for a one-sided chart constants k (> 0 ; the reference value) and h (> 0 ; the signal level) are fixed and one computes $T_i = \max\{T_{i-1} + U_i - k, 0\}$ for $i=1, 2, \dots$, where $T_0 = 0$. An out of control signal is given at the first i where $T_i \geq h$. When the process is in-control, the ARL of this scheme depends only on h and k and not on the underlying cdf F . Note that this procedure is not a direct

analog of the usual CUSUM based on the sample means. The approach taken here is to determine, numerically, for a given reference value k , the appropriate signal level h corresponding to a desired average run length. It may be noted that this procedure tracks the actual sequence of random variables X_i , $i=1,2,\dots$, through the cumulative sequential ranks.

Alloway and Raghavachari (1991) (hereafter AR) considered a Shewhart-type chart for the median of a continuous symmetric population, based on a DF confidence interval for θ , calculated using the Hodges-Lehmann (HL) estimator. Let m subgroups, each of size n , be available. The HL estimator for the point of symmetry of a continuous symmetric distribution is defined as follows. For the i th random sample, define $M = n(n+1)/2$ "Walsh averages" $W_{ir} = (X_{ij} + X_{ih})/2$, $r=1,2,\dots,M$; $1 \leq j \leq h = 1,2,\dots,n$. Then the HL estimator of θ_i is $\tilde{\theta}_i$, the median of the Walsh averages, i.e., $\tilde{\theta}_i = W_{i(\frac{M+1}{2})}$ if M is odd and $\tilde{\theta}_i = \frac{1}{2}(W_{i(\frac{M}{2})} + W_{i(\frac{M+1}{2})})$ if M is even. It is known that if the underlying distribution is normal, in large samples, the efficiency of $\tilde{\theta}$ relative to \bar{X} is 0.955. This means that although the sample mean is the most efficient estimator of the population mean when the distribution is normal, the HL estimator is almost as efficient for moderate to large sample sizes. Of course, the advantage with the HL estimator is that the normality assumption is not required and it is robust against outliers.

If $W_{i(1)}, W_{i(2)}, \dots, W_{i(M)}$ (our notation is slightly different from AR) are the M ordered Walsh averages for the i th sample, then a $100(1-\alpha)\%$ DF confidence interval for θ is given by two ordered Walsh averages, $W_{i(a_i)}$ and $W_{i(M-a_i+1)}$

such that $P(W_{i(a_i)} \leq \theta \leq W_{i(M-a_i+1)}) \geq 1-\alpha$. Using the connection with the WSR statistic, tables have been constructed for finding the a_i (see for example, Gibbons and Chakraborti (1992) and Wilcoxon, Katti and Wilcox (1972)). The steps for calculating the AR control chart are as follows. First find the $100(1-\alpha)\%$ confidence intervals: $(W_{1(a_1)}, W_{1(M-a_1+1)}), \dots, (W_{m(a_m)}, W_{m(M-a_m+1)})$, for the median θ from each of the m groups. The control lines are defined by LCL = median of the m lower confidence limits, UCL = median of the m upper confidence limits and CL = average of the m HL estimators. One plots $\tilde{\theta}_i$ versus $i=1,2,\dots$ and compares against the control lines. The sample size n is recommended to be at least 10 so that the type I error probability is comparable to a 3-sigma Shewhart \bar{X} chart. Performance of the proposed chart was examined in a simulation study. As it might be expected, this approach compares favorably with the \bar{X} chart for the normal distribution and is better in case of heavy-tailed symmetric distributions.

In spite of the intuitive appeal, the design of the AR charts appears to be flawed from a practical point of view since it is not clear what the type I error probability or the in-control ARL for this chart is. As noted by Pappanastos and Adams (1996), and reviewed later in this section, the problem seems to be that the AR control limits don't seem to be directly based on the in-control distribution of the control statistic $\tilde{\theta}_i$. It is also not clear whether the AR chart was to be used retrospectively or prospectively. More will be said about this later including possible modifications.

Hackl and Ledolter (1991) (hereafter HL) considered NP control chart procedures for individual observations that use the so-called "standardized ranks" of the observations relative to an in-control distribution. The standardized rank R_i of an observation X_i is defined as $R_i = 2[F_0(X_i) - 1/2]$, where F_0 is the cdf of the in-control distribution. In the known F_0 case, the R_i 's can be computed

directly; in the unknown case, the standardized rank R_i is redefined as $R_i^\# = 2g^{-1}[R_i^* - (g+1)/2]$, where a random sample (a historic or a reference sample) of size $g-1$, say $(Y_1, Y_2, \dots, Y_{g-1})$, is assumed to be available when the process is in-control and R_i^* is the rank of X_i with respect to the reference sample, so that, $R_i^* = 1 + \sum_{j=1}^{g-1} I(X_i > Y_j)$. Taking the reference sample as fixed (that is, conditionally, given the reference sample), it can be shown that the standardized ranks $R_i^\#$ are independent and identically distributed. The difference is that whereas the ranks R_i follow a continuous uniform distribution over $[-1, 1]$, the ranks $R_i^\#$ follow a discrete uniform distribution over the g mass-points $\{1/g-1, 3/g-1, \dots, 1-3/g, 1-1/g\}$. The proposed control chart is based on an EWMA of the ranks R_i (or $R_i^\#$): $T_i = (1-\lambda)T_{i-1} + \lambda R_i$, $i = 1, 2, \dots$, where T_0 is usually set to 0 and λ is a smoothing parameter (in $(0, 1)$) usually recommended to be between 0.1 and 0.3. Against a two-sided shift alternative, the process is declared out of control if at some i (observation number or time), $|T_i| > h$, where $h > 0$, is a suitably chosen control limit. Thus the main idea here is to define ranks of the accumulating observations in some suitable way and apply the usual EWMA method on these ranks. In simulation studies it is observed that the proposal is resistant to outliers and performs well if one is concerned about sudden shifts in the mean. In the same spirit, Hackl and Ledolter (1992) considered a chart based on the sequential rank of an observation. In contrast with BR (1979) however, their sequential rank of an observation is defined as its rank among the most recent group of g observations. The control statistic used is an EWMA of the sequential ranks. From simulation results, HL suggest that this chart is also outlier resistant and performs well if one is concerned about slowly trending process levels.

Amin and Searcy (1991) considered a NP EWMA chart based on the control statistic $Z_i = \lambda Y_i + (1-\lambda) Z_{i-1}$, where $Y_i = SR_i$ is the Wilcoxon group signed-rank (GSR) statistic introduced earlier by BR (1979). The starting value Z_0 is taken to be the process target value. The process is considered to be out-of-control whenever some Z_i either falls above the UCL or below the LCL. The control limits are given by $\mu_0 \pm L$. Properties of the GSR-EWMA were evaluated and compared on the basis of ARL by simulation. Distributions were taken to be normal, uniform, double-exponential, Gamma and the Cauchy. The control limits for both the standard \bar{X} -EWMA and the GSR-EWMA were obtained such that the "frequency of points falling outside the control limits were approximately equal for both procedures when the process is in-control." It is suggested that a control chart for variability is used along with the GSR procedure. The authors also examined the effect of autocorrelation. The performance of the GSR-EWMA relative to the \bar{X} -EWMA was shown to be similar to that of the GSR-CUSUM (studied by BR) relative to the \bar{X} -CUSUM. It is seen that the ARL properties of the proposed GSR-EWMA is insensitive to the choice of λ values. Enhancements such as addition of warning limits improve performance of the chart. Autocorrelation doesn't seem to affect the ARL as much it affects the ARL of an \bar{X} -EWMA chart. Overall, the GSR-EWMA methods provides a nice alternative NP charting procedure.

Yaschin (1992) discussed the run length distribution of a CUSUM control scheme when the underlying distribution is unknown. He suggested a NP analysis of the run length and some associated characteristics simply replacing the true underlying distribution by the empirical distribution of a reference sample. Properties of the resulting estimators were considered and simulation results were presented.

Amin, Reynolds and Bakir (1995) (hereafter ARB) presented NP charts for the process median (or the mean)

and the process variability. These procedures are based on what might be called within group sign statistics, used instead of the average, in the usual Shewhart and CUSUM charts. The sign test is the simplest of NP tests (see for example, Gibbons and Chakraborti, 1992) that can be used to test for the median (or a specified quantile) of any continuous population. This test doesn't require that the distribution is symmetric and therefore is applicable in a variety of situations. ARB used the statistics $SN_i = \sum_{j=1}^n \text{sign}(X_{ij} - \theta_0)$, where X_{ij} is the j th observation from the i th group of size n . The SN_i are linearly related to the usual sign statistics say K_i (the total number of positive signs among $X_{ij} - \theta_0$), through the relation $2K_i = SN_i + n$, so that the probability distribution of SN_i can be found from that of K_i ; the latter being Binomial $(n, 1/2)$ when the process is in-control (median = θ_0). These authors also considered Shewhart-type (zone) charts with warning limits and runs rules and provide formulas for the ARL of the combined chart. For example, the ARL of a one-sided (positive direction) chart with warning limit at w_2 ($0 \leq w_2 < a_2$) and control limit at a_2 is given by

$$L^+(\theta) = \frac{1 - p_1^r}{1 - p_1 - p_0(1 - p_1^r)}, \quad (3)$$

where $p_0 = P(SN_i < w_2|\theta)$, $p_1 = P(w_2 \leq SN_i < a_2|\theta)$ and a signal is given if r consecutive points fall in $[w_2, a_2)$ or any point falls outside a_2 . A table is provided for values of L^+ for various a_2 , w_2 and r values when $n=10$. As a practical note, some observations can be tied with the specified median. If the number of ties is small (relative to n) simply drop the tied cases and reduce n accordingly. On the other hand if the number of ties is large, more sophisticated analysis might be possible. The authors also considered a NP chart to monitor

process variability by adapting the two-sample interquartile range test. Clearly, and as the authors pointed out, there needs to be much further work done on this topic. For CUSUM charts the authors use the same type of rule as in (1) or (2) with SN_i used in place of SR_i and calculate the ARL as before using a Markov chain approach, where the transition probabilities are calculated via a binomial distribution. Optimal values of k and h are determined similarly and tables are given for $n=10$ and various distributions. As in the case of the WSR statistics, it is observed that using k values for the normal distribution does not lead to large errors. Finally, the Shewhart \bar{X} chart and the Shewhart sign chart (with and without warning limits) are compared, on the basis of ARL (both one and two-sided) for various shift sizes and underlying distributions like the normal, the double-exponential and the Gamma. The in-control ARL (say ARL_0) of the charts is kept at some constant value. It is seen that generally speaking, when the distribution is either asymmetric or symmetric with heavy tails, the NP (sign statistic based) charts are more efficient while the reverse is true for the normal and the normal-like distributions with light tails. These authors also compared the proposed NP chart for variability to the chart based on S^2 and suggested that the chart based on S^2 is more efficient, but of course the chart based on S^2 is not NP. Finally, one-sided CUSUM charts using the sample means and the sign statistics are compared. It is seen that the CUSUM chart using the SN_i is more efficient than the Shewhart charts, with or without warning limits. The overall conclusion is that the NP charts provide a useful alternative to the standard charts when normality is in doubt.

Pappanastos and Adams (1996) (hereafter PA) noted that a problem with the AR charts is their inability to maintain ARL_0 at any practically reasonable value. For example, using simulations with $n=10$ and $m=30$ and under normality, it was found that the ARL_0 of the AR chart is 20,820.89, when the anticipated ARL_0 is just 500. Also, when different distributions such as the uniform or the

double-exponential were used, the ARL_0 values of the AR chart varied widely. This contradicts the fact that the AR charts are claimed to be DF. PA thus conclude "If the Hodges-Lehmann control chart were truly NP, one would expect the same ARL_0 's for different distributions." These authors go on to say "The discrepancy in the actual and anticipated average run lengths is due to the fact that the control limits for the Hodges-Lehmann control chart are not based on the distribution of the plotted statistic (i.e., the Hodges-Lehmann estimator)." However, as we have noted earlier, using even the "correct" in-control distribution of the HL estimator to set the control limits wouldn't help in this respect, because the in-control distribution of the HL estimator is not DF.

PA considered two alternative forms of the AR charts as "robust" alternatives to the \bar{X} chart. The alternative design schemes allow the user to construct a control chart with a specified ARL_0 while maintaining the advantages of the AR control chart. But as we have just noted, using the AR charts with the HL estimator as the control statistic is inherently problematic. Also, it is not clear what is exactly meant by a robust alternative. In any case, PA explored (i) plotting the HL estimators against control limits based on the asymptotic variance of the estimator and (ii) plotting a multiple of the HL estimator against control limits based on the medians of the m smallest and m largest sample observations. Using simulations, the authors recommend using the limits $\theta^* \pm c_2^* \sqrt{\bar{s}^2}$, when the process is normally distributed, where $\theta^* = \text{median}(\tilde{\theta}_i, i=1,2,\dots,m)$, \bar{s}^2 is the average of the m subgroup variances, and c_2^* is some constant chosen to achieve a specific ARL_0 . A chart is provided for finding the constant for various ARL_0 , when $m=30$ and $n=3(1)10$. This, however, seems to

be pointless since it is not clear if anyone is going to use a NP chart (such as the one based on the HL estimator) in practice if the underlying distribution is known to be normal. The same caveat applies to the authors' second modification. The interesting question in all of this, however, is how to incorporate or use a confidence interval (or perhaps some other interval estimation techniques) into defining a control charting scheme. Such a question has ramifications both for parametric and NP statistical process control.

Willemain and Runger (1996) (hereafter WR) considered designing control charts for individual observations using a so-called "empirical reference distribution." They assume that a large reference sample is available and argued that "With sufficient historical data, regardless of the distribution, control limits can be selected as particular order statistics of the observed distribution of the variables to be charted." They go on to say, "In general, we favor the approach of developing control limits from an empirical reference distribution based on process data acquired during normal operating conditions instead of strict reliance on a normality assumption." The proposed Shewhart-type control limits are given by two order statistics of a reference sample of size m , the k^{th} smallest and the $(b+k)^{\text{th}}$ smallest, where $0 \leq k \leq m$ and $1 \leq b \leq m+1-k$. Individual observations are then collected, one at a time, and are compared to these limits. It is shown that the conditional probability P (given the reference sample order statistics) that a future independent observation will fall within the control limits, when the process is in-control, is a beta random variable with parameters b and $m-b+1$. From this, the (unconditional) distribution of the in-control run length is derived analytically, which is related to a hypergeometric distribution, with a right tail longer than that of the geometric. This yields, for example, the mean (i.e., $ARL_0 = m/(m-b)$) and the variance of the in-control run length distribution. However, it is not completely clear how to determine the chart parameters k and b . For example, the ARL_0 can be used to choose the constant b as follows. Take

$m = 1,000$, so that to achieve $ARL_0 = 370$, one would set $1000/(1000-b) = 370$ and solve for b , which yields $b = 997$. However the constant k still needs to be found, which would have to be one of 1, 2 or 3. WR seem to suggest using a symmetric two-sided chart, which would mean $k = 2$.

WR also studied the “off-target” ARL of their chart and provided a table for comparisons with the one-sided normal theory Shewhart chart. For two-sided charts simulations are used to estimate the $E(ARL)$ and a table is provided for a comparison of exact and empirical estimates of off-target ARL using $m=10,000$ observations from a standard normal distribution. In conclusion the authors state that “the results were good, although additional research may be able to improve upon the simple estimators...”

The idea of using a reference sample to set up control limits is also utilized in Janacek and Meikle (1997) (hereafter JM). They presented a DF control chart for the median of a future sample from the same process. Note that the control chart here is for a particular order statistic from a future sample (and not for the population mean or the population median). This extends the work of WR cited above. Assume that a reference sample of size m , say X_1, X_2, \dots, X_m , is available with when the process is in-control with a cdf F_0 . Whether or not the process is in-control is judged by taking a sequence of test samples of size n and comparing each test sample with the reference sample. Ideally the aim is to detect a change in the distribution, say from F_0 to F_1 , but as in practice, detecting a shift in the location of F_0 is of interest. The procedure is to compare the test sample medians M_i with the limits given by two order statistics of the reference sample, $LCL = X_{(j)}$ and $UCL = X_{(m-j+1)}$, where the constant j is determined so that the probability

$P(X_{(j)} < M_i < X_{(m-j+1)} | F_1 = F_0) \geq 1 - \alpha$, for all $i = 1, 2, \dots$, where α is the specified false alarm rate. JM has tabulated this probability for $j = 1(1)10$, when $m = 25(5)80$ and $n = 5(2)9$ and also when $m=55(5)80$ and $n = 11(2)15$. For example, for $m=70, n=5$, $P(X_{(3)} < M_i < X_{(68)} | F_0 = F_1)$ is calculated to be 0.99716 (so that the actual false alarm rate is 0.00284). Thus, taking $LCL = X_{(3)}$ and $UCL = X_{(68)}$ is roughly comparable to a traditional 3-sigma Shewhart \bar{X} chart in this situation.

In summary, when a reference sample is available from an in-control process, it can be used prospectively, to check whether or not the process is in-control. This can be done by either (i) estimating (predicting) some attribute of a future sample (say the 90th percentile, or the inter-quartile range, for example) or (ii) by estimating some attribute of the future distribution (the mean or the median, for example). Along the lines of (i), and generalizing the works of JM and WR, Chakraborti and Van der Laan (1998) considered estimating the j th order statistic (i.e., the $100*(j/n)$ th sample percentile) in a future sample, based on a class of two-sample NP statistics, called precedence statistics. They also examined the performance of their chart in terms of the ARL. Computational aspects and recommendations for the implementation are also given. More work needs to be done in this context, particularly using other two-sample NP statistics, which are known to possess “optimal” power properties.

Ledolter and Swersey (1997) discussed pre-control, an alternative to statistical control charts for monitoring processes. Pre-control and standard control charts are compared. They find that pre-control has some value, especially in machining operations where the lot sizes are small, and in situations where one deals with very capable processes. But, in general, their conclusion is that pre-control is not an adequate substitute for control charts.

Finally, we briefly describe some other problems where NP methods have been proposed in the literature. Some of these are active areas of research, especially among

the more theoretically inclined researchers. However, the problems are very much relevant in the process control setting. We list these under other methods.

4. Other Methods

Since the subgroups are almost always collected sequentially over time (at some equally spaced time points), it seems natural to consider some sequential statistical methods for process control problems. In the classical (Wald) sequential setup, subgroup size is 1 and the number of observations required to reach a decision is a random variable. The 'optimal' procedure is chosen in such a way that subject to given bounds on the type I and the type II errors, the expected number of observations to reach a decision is a minimum. To this end, adapting from Sen (1991; page 235) one possible formulation of the problem is as follows. Let T_n be a class of (control) statistics. In order to test if the process is in-control (versus that the process is not in-control) based on T_n , start with an initial sample of size n_0 and define a stopping variable

$N =$ least positive integer $n (\geq n_0)$ such that T_n gives a signal
 $= \infty$, if no such n exists.

Thus we continue drawing observations, starting with n_0 , then n_0+1, n_0+2 , and so on, until for the first time (for some $n = n^* = n_0+K$), T_{n^*} gives a signal (that the process is not in-control); then $N = n^*$. If no such n exists then process is allowed to continue under the assumption that the process is in-control. Sequential statistical methods have been successfully used in medical experimental settings and various procedures have been developed in view of the applications. It would be useful to examine how these methods, adapted if and as

necessary, can be applied in process control problems.

In the literature on sequential testing and estimation, problems have been discussed that are called "change-point", or more generally, "detection" problems. Bhattacharyya and Frierson (1981) considered the following problem. Let X_1, X_2, \dots, X_N be a sequence of independent random variables whose distribution changes from F to G after the first $[N\theta]$ observations, where θ is an unknown parameter. This is one version of a change-point problem. The object is to detect the unknown change-point quickly without too many false alarms and without making any parametric model assumptions on F or G . A NP control chart based on the (partially) weighted sums of sequential ranks is proposed and the asymptotic behavior of the cumulative sums of sequential ranks, under the assumption that a small change in distribution takes place after a large number of observations, is studied.

Zacks (1991) presented an overview of detection and change-point problems and considered some applications of the proposed methods. The reader is referred to this paper for an introduction to the various problems and proposed solutions, along with references to the literature; the discussion on applying CUSUM procedures in change-point problems is particularly interesting in the process control setting. In addition, we cite three more references: Huskova and Sen (1989), Siegmund (1994) and Siegmund and Venkataraman (1995), where more recent works and further references can be found.

5. Concluding remarks

As noted in section 1, in some applications the location-scale model is the more relevant model from a practical point of view. For this situation it seems worthwhile to consider a "combined" control chart, combining, say, a (two-sample) location statistic with a (two-sample) scale statistic. For different NP location and scale tests, see for example, Gibbons and Chakraborti, (1992). However, one possible drawback of a combined chart is that when a signal is given, it is not always easy to isolate the

reason, i.e., it is not easy to diagnose if there has been a shift only in, the location, or the scale, or both.

Also, as noted before, one could explore the issue of “optimal” NP charting by using, for example, “optimal” NP tests. Of course, one needs to define what is meant by an optimal control chart. In the same spirit, one also needs to define what might be called the “efficiency” of a NP chart over a parametric (say classical normal theory) based chart and study the advantages of one chart over the other. To this end, one could examine “local” properties of the ARL of a chart under, for example, “contiguous” shift alternatives. Some of these analyses would entail asymptotics, where the sample size and/or the number of samples might be large. Clearly, more research is needed in these directions.

Since the choice of a control chart depends on the type of the underlying process distribution, it seems useful to explore what might be called “adaptive control charts.” Here one could use a preliminary reference sample to gauge, for example, the skewness and the kurtosis of the population, and based on such estimates one could choose an “optimal” NP control charting method. For an introduction to adaptive statistical procedures, see Hogg (1974).

References

- Alloway, J. A. and Raghavachari, M. (1991). Control chart based on the Hodges-Lehmann estimator. *J. of Quality Technology* **23**, 336-347.
- Amin, R. and Searcy, A. J. (1991). A nonparametric exponentially weighted moving average control scheme. *Commun. Statist.-Simul. Comput.* **20**, 1049-1072.
- Amin, R. W., Reynolds, M. R. Jr., and Bakir, S. T. (1995). Nonparametric quality control charts based on the sign statistic. *Commun. Statist.-Theory Meth.* **24**, 1597-1623.
- Arnold, B. (1985). The sign test in current control. *Statistische Hefte* **26**, 253-262.
- Bakir, S. T. and Reynolds, M. R. Jr. (1979). A nonparametric procedure for process control based on within-group ranking. *Technometrics* **21**, 175-183.
- Bhattacharyya, P. K. and Frierson, D. (1981). A nonparametric control chart for detecting small disorders. *Annals of Statistics* **9**, 544-554.
- Chakraborti, S. and Van der Laan, P. (1998). Nonparametric control charts based on precedence statistics. In preparation.
- Ferrell, E. B. (1953). Control charts using midranges and medians. *Industrial Quality Control* **9**, 30-34.
- Gibbons, J. D. and Chakraborti, S. (1992). *Nonparametric Statistical Inference*, Third Edition, Marcel Dekker, New York.
- Gunter, B. H. (1989). The use and abuse of C_{ph} . Part 2. *Quality Progress* **22**, 3, 108-109.
- Hackl, P. and Ledolter, J. (1991). A control chart based on ranks. *J. of Quality Technology* **23**, 117-124.
- Hackl, P. and Ledolter, J. (1992). A new nonparametric quality control technique. *Commun. Statist.-Simula.* **21**, 423-443.
- Hogg, R. V. (1974): Adaptive robust procedures: a partial review and some suggestions for future applications and theory. *J. Am. Statist. Assoc.* **69**, 909-927.
- Huskova, M. and Sen, P. K. (1989): Nonparametric tests for shift and change in regression at an unknown time point. In *Statistical Analysis and Forecasting of Economic Structural Change* (P. Hackl, ed.), Springer-Verlag, New York.
- Jacobs, D. C. (1990). Statistical process control: watch for nonnormal distributions. *Chemical Engineering Progress* **86**, 19-27.
- Janacek, G. J. and Meikle, S. E. (1997). Control charts based on medians. *The Statistician* **46**, 19-31.
- Langenberg, P. and Iglewicz, B. (1986). Trimmed mean \bar{X} and R charts. *J. of Quality Technology* **18**, 152-161.
- Ledolter, J. and Swersey, A. (1997). An evaluation of pre-control. *J. of Quality Technology* **29**, 163-171.
- Lehmann, E. L. (1983). *Theory of point estimation*. John Wiley, New York.
- Lucas, J. M. and Croisier, R. B. (1982). Robust CUSUM: A robustness study for CUSUM quality control schemes. *Commun. Statist.-Theory and Methods* **11**, 2669-2687.
- McDonald, D. (1990). A CUSUM procedure based on sequential ranks. *Naval Research Logistics* **37**, 627-646.
- McGilchrist, C. A. and Woodyer, K. D. (1975). Note on a distribution-free CUSUM technique. *Technometrics* **17**, 321-325.
- Montgomery, D. C. (1991): *Statistical Quality Control*. John Wiley, New York.
- Noble, C. E. (1951). Variations in conventional control charts. *Industrial Quality Control* **8**, 17-22.
- Orban, J. and Wolfe, D. A. (1982). A class of distribution-free two-sample tests based on placements. *J. Am. Statist. Assoc.* **77**, 666-670.
- Page, E. S. (1954). Continuous inspection schemes. *Biometrika* **41**, 100-114.
- Pappanastos, E. A. and Adams, B. M. (1996). Alternative designs of the Hodges-Lehmann control chart. *J. of Quality Technology* **28**, 213-223.
- Park, C. and Reynolds, M. R., Jr. (1987). Nonparametric procedures for monitoring a location parameter based on linear placement statistics. *Sequential Analysis* **6**, 303-323.
- Reynolds, M. R. Jr. (1975). A sequential signed-rank test for symmetry. *Annals of Statistics*. **3**, 382-400.
- Rocke, D. M. (1989). Robust control charts. *Technometrics* **31**, 173-184.

- Sen, P. K. (1991). *Sequential Nonparametrics*. John Wiley, New York.
- Shewhart, W. A. (1939). *Statistical methods from the viewpoint of quality control*. Republished in 1986 by Dover Publications, New York, NY.
- Siegmund, D. (1994). A retrospective of Wald's sequential analysis-its relation to change-point detection and sequential clinical trials. In *Statistical Decision Theory and Related Topics, V* (S. S. Gupta and J. O. Berger, eds.), 19-33.
- Siegmund, D. and Venkataraman, E. S. (1995). Using the generalized likelihood ratio statistic for sequential detection of a change-point. *Annals of Statistics* **23**, 255-271.
- Tukey, J. W. (1960). *A survey of sampling from contaminated distributions*. Contributions to Probability at Statistics, Essays in Honor of Harold Hotelling. (I. Olkin et al., eds.) Stanford University Press, Stanford, CA.
- Van der Laan, P. (1966). A sequential distribution-free two-sample grouped test with three possible decisions. *Statistica Neerlandica* **20**, 31-41.
- Wilcoxon, F., Katti, S. K. and Wilcox, R. A. (1972). Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test. *Selected Tables in Mathematical Statistics*, Vol. I., American Mathematical Society, Providence, R. I.
- Wilcoxon, F., Rhodes, L. J. and Bradley, R. A. (1963). Two sequential two-sample grouped rank tests with applications to screening experiments. *Biometrics* **19**, 58-84.
- Willemain, T. R. and Runger, G. C. (1996). Designing control charts using an empirical reference distribution. *J. of Quality Technology* **28**, 31-38.
- Yaschin, E. (1992). Analysis of Cusum and other Markov-type control schemes by using empirical distributions. *Technometrics* **34**, 54-63.
- Yourstone, S. A. and Zimmer, W. J. (1992). Non-normality and the design of control charts for averages. *Decision Sciences* **23**, 1099-1113.
- Zacks, S. (1991). Detection and change-point problems. In *Handbook of Sequential Analysis* (B. K. Ghosh and P. K. Sen, eds.), Marcel Dekker, New York.