METRON - *International Journal of Statistics*
2006, vol. LXIV, n. 3, pp. 357-375

WILLEM ALBERS – WILBERT C. M. KALLENBERG

Alternative Shewhart-type charts for grouped observations

Summary - Classical Shewhart control charts based on groups of observations use the average as statistic to decide whether a signal should be produced or not. Assuming a known process distribution here several other statistics are investigated as well. Moreover, individual charts are compared with charts with larger group size. The chart using the minimum of each group of observations is a promising new proposal. The behavior of the minimum-chart is competitive to that of the average-chart when distributions are known. Moreover, when the process distribution is unknown, for this chart (nonparametric) estimation of an extreme quantile is reduced to estimation of an ordinary quantile of the underlying distribution itself, which is far easier to perform accurately.

Key Words - Statistical Process Control; Phase II control limits; Order statistics.

1. Introduction

During a continuously operating process the mean of the production process is monitored using a Shewhart chart based on (groups of) incoming measurements. For example, a packaging operation screws lids on bottles continuously as they move through the process. An implementation of Shewhart control charts may have samples of 4 bottles randomly selected in 10 minutes intervals. Even if we start with samples of size 1 we may decide to form groups of for instance size 4, that is, we may postpone the decision until 4 observations have arrived. For ease of presentation we concentrate on the one-sided case where only an upper limit is needed; the two-sided case can be treated in a similar fashion and will lead to similar results. We observe a group of independent observations $X_1, \dots, X_m$ (with $m = 1, \dots, 5$, thus including the individual chart as well), which are collectively either in-control (IC) or out-of-control (OoC), see also below for further explanation about this collectivity. Under IC X_i is distributed as X, having distribution function (df) F and under OoC X_i is distributed as $X + d$ with $d > 0$. Let $w(X_1, \dots, X_m)$ denote a statistic which is applied to

Received May 2005 and revised August 2006.

mutually exclusive rational subgroups of m observations to create the sequence $W_{1m} = w(X_1, \ldots, X_m)$, $W_{2m} = w(X_{m+1}, \ldots, X_{2m})$, $W_{3m} = w(X_{2m+1}, \ldots, X_{3m})$ and so on. An alarm is produced when W_{km} exceeds an upper limit, written as $\mathrm{UL}(w, m)$ to emphasize the dependency on the statistic and rational group size.

In the individual case ($m = 1$), it is trivially clear that a signal will arise if the new observation is too large. But in the grouped case, first the question has to be dealt with which statistic w based on the m observations, should actually be used. Under normality, the answer is straightforward: the sample mean is optimal and easy to work with. In fact, in a few simple steps the case $m > 1$ is reduced to the case $m = 1$. Beyond the normal model, the picture is quite different, however. The sample mean is not necessarily optimal, and it also is not particularly easy to deal with (remember that $m \leq 5$, so the central limit theorem is not of much use here, especially not as the interest under IC is focused on the tails of the distribution).

Hence, the subject of the present paper will be the study of a variety of possible statistics for use in grouped control charts. One aspect will obviously be how efficient a particular choice is: given a certain average run length (ARL) under IC, how large is the probability of detection during OoC offered by the choice made? Another criterion will be its ease of application, in particular in view of needed estimation steps. Moreover, note that in fact two types of comparisons play a role. In the first place, for each fixed value of m, various statistics can be compared. But each given type of statistic can also be compared for varying m. Even the normal case is not quite trivial in this respect and still leads to some interesting insights. The point is of course that we are not dealing with a single given OoC-situation, implying that the optimal choice of m will vary according to the shift d under consideration.

The first step is to figure out answers on the above questions for a known (but not necessarily normal) underlying distribution. Once a more or less clear picture has been obtained about which statistics have which properties under which conditions, a second step should be taken. This will entail the estimation of the parameters and/or distributions involved, the study of the estimation effects incurred and the derivation of possible corrections for errors which are considered to be intolerably large. In the present paper we shall address the first step, and *thus work under the assumption of a known underlying distribution*. But estimation is nevertheless present in the background, since the statistics involved should be related e.g. to corresponding nonparametric control charts and the possibilities for estimation and its consequences should be taken into account. (Therefore, we do not use for instance a likelihood ratio approach, which for known distributions is optimal, but gives insuperable problems when estimation comes in.) The latter concerns only a first impression. The full second step, concerning the estimation aspects, will be dealt with in a forthcoming paper.

A short remark about the estimation step should be made here, in particular for nonparametric charts. To get suitable nonparametric charts for less than 500 Phase I distributions it seems necessary to use groups of observations, see Albers, Kallenberg and Nurdiati (2006) and Albers and Kallenberg (2006). The advantage of grouping is that we do not have to decide immediately when a Phase II observation arrives, but that we may postpone until some more observations appear. Clearly, $2, 3$ or 4 observations tell us more than only 1. Taking for instance the minimum of 3 observations as yardstick, we have to deal with its upper 0.003-quantile to get ARL equal to 1000 (since $1000 = 3/0.003$) and that corresponds to the 0.144-quantile of the original distribution. Hence, estimating of the extreme 0.001-quantile for the individual chart is replaced by estimating the 0.144-quantile here, which can be done very well with 50 Phase I observations, say.

The disadvantage of working with a group of $m > 1$ Phase II observations is that we cannot stop at an earlier time than after m steps. Moreover, the next possible stop is at time $2m$ etc. We have jumps of size m in stopping the process. Moreover, when the process gets OoC we may encounter a group of observations with a mix of IC- and OoC-observations. Therefore, we should take m rather small: $2, 3, 4$ or 5. When restricting to those m's, it is clear that in case of IC, where the ARL equals 1000, a difference in RL of m observations is completely negligible and hence the discrete character of RL and the possible mix of IC- and OoC-observations do not matter. The OoC situation is slightly different, but we concentrate on changes in the process with ARL for the individual chart between 10 and 200, see Section 2. Hence, also here we do not bother too much about differences in detecting of a few more observations. For that reason we have assumed that $X_1, \dots, X_m$ are collectively either in-control or out-of-control.

The paper is structured as follows. In Section 2 the general approach is given. Here we introduce the set-up, the notation involved and we present the various statistics for use in grouped observations and the general form of their control limits. The comparison in case of normality between the individual chart and the group chart based on averages is already presented in Section 2 to make the connection between the introduction of statistics (in Section 2) and the evaluation (mainly in Section 3) in the context of a problem where there exists strong intuition for the result. Among the statistics considered is the minimum of the m observations. As far as we know the chart based on this choice is new. Howell (1949) proposed a Shewhart-type control chart using the maximum for the upper control limit and the minimum for the lower control limit (see also Sarkadi and Vincze (1974) and Amin and Wolff (1994); in the context of EWMA control charts we refer to Amin *et al.* (1999)). In our opinion (see Section 2.3) this should be done the other way around: the minimum when dealing with the upper control limit and the maximum for the lower control limit. Next, in Section 3, we compare the several methods and the size m of the groups of observations. We start with the normal distribution followed by some symmetric distributions (random normal mixture) which are not too far from normality and end with a skew distribution, representing a distribution farther away from normality. It turns out that the new chart based on the minimum is very promising. A brief summary of the conclusions is given in Section 4.

2. General approach

For a group of m observations we define the ARL by

$$\text{ARL} = \frac{m}{\text{FAR}},$$

where FAR, the false alarm rate, is given by

$$\text{FAR} = P\left(W_m > \text{UL}(w, m)\right). \tag{1}$$

To compare the charts for different values of m in a fair way we match the ARL's under IC. The prescribed ARL will be denoted by $1/p$, leading to

$$\text{FAR} = mp \tag{2}$$

under IC. For instance, with $p = 0.001$ we arrive at ARL $= 1000$ and FAR $= 0.001m$.

For any df H we will write $\overline{H} = 1 - H$, and H^{-1} and $\overline{H}^{-1}$ for the respective inverse functions. (Observe that the inverse is defined unambiguously for continuous and increasing H; for the remaining cases a choice has to be specified, but we will tacitly assume in this paper that all df's are continuous and increasing.) Writing $F_{w,m}$ for the df of $w(X_1, \ldots, X_m)$ in the IC case, we have, see (1) and (2),

$$\text{UL}(w, m) = \overline{F}_{w,m}^{-1}(mp). \tag{3}$$

The performance of several statistics $w(X_1, \ldots, X_m)$ (and several values of m) will be investigated by their ARL under OoC: the smaller the ARL, the better the chart.

Remark 2.1. The variability of the run-length distribution is an important issue. For instance, when estimating the parameters μ and σ in the individual ($m = 1$) normal chart, we get a positive bias under in-control for the ARL. This could suggest that even when we know μ and σ it would be better to estimate them! However, the bias is due to occurrence of extremely long runs, which are not relevant in practice. For more details see e.g. Remark 2.1 (and further remarks) in Albers and Kallenberg (2004).

In the situation where parameters are known, the run-length distribution is simply given by the distribution of mZ, where Z has a geometric distribution with parameter FAR. Hence, ARL $= m/\text{FAR}$. Since the geometric distribution is completely determined by its parameter, for comparison there is no essential difference to take the standard deviation of the run-length distribution, being $m\sqrt{1-\text{FAR}}/\text{FAR} = \text{ARL}\sqrt{1-m/\text{ARL}}$, or the median of the run-length distribution, being $m\left[(-\log 2)\,/\log(1-\text{FAR})\right] = m\left[(-\log 2)\,/\log(1-(m/\text{ARL}))\right]$. These are simply monotone increasing functions of ARL. Therefore, we take ARL as our metric.

We take the shift d such that the ARL for $m = 1$ runs from 10 to 200. Such changes in the process are the most interesting ones when dealing with Shewhart charts: for smaller changes the Shewhart chart is less natural (cf. discussions about the relative merits of Shewhart and CUSUM charts), larger changes are detected anyhow in a few steps. Hence, this natural restriction of $d = d(F)$ is given by

$$P\left(X + d > \overline{F}^{-1}(p)\right) \in \left(\frac{1}{200}, \frac{1}{10}\right), \tag{4}$$

with X having df F.

In the following subsections various choices of the statistics $w\left(X_1, \ldots, X_m\right)$ are presented.

2.1. AVE

To begin with we consider the obvious choice (at least under normality), which is the average-chart (AVE), based on

$$w\left(X_1, \ldots, X_m\right) = m^{1/2}\overline{X}\,.$$

Hence, in view of (3) the AVE-chart reads as

$$m^{1/2}\overline{X} > \overline{F}^{-1}_{m^{1/2}\overline{X}}(mp)\,, \tag{5}$$

where $\overline{F}^{-1}_{m^{1/2}\overline{X}}(mp)$ denotes the upper $(mp)^{th}$-quantile of $m^{1/2}\overline{X}$ under IC.

When normality holds this clearly is the optimal choice, but also in a nonparametric context it is of interest. Let us briefly digress into the estimation case. Using averages while F is unknown brings us into the area of normal permutation tests. Some efforts of this type were mentioned in Chakraborti *et al.* (2001). For example, Alloway and Raghavachari (1991) use a procedure based on the Hodges-Lehmann estimator and work with Walsh averages.

However, as pointed out by Chakraborti *et al.* (2001) and by Pappanastos and Adams (1996), the resulting charts are in fact not truly nonparametric or distributionfree. Their actual IC run length distribution involved does depend on the underlying distribution of the observations. As is seen from (5) the $(mp)^{th}$-quantile of $m^{1/2}\overline{X}$ should be estimated in a nonparametric way when F is completely unknown. A possible approach is to use the $\binom{n}{m}$ groups of m observations obtained from $X_1, \ldots, X_n$ and to take a suitable order statistic among these $\binom{n}{m}$ "observations". A thorough analysis of the problem was performed by Albers and Kallenberg (2005), focusing on the tail behavior of the empirical df of convolutions. It turns out that going to $m > 1$ does not really help that much: the estimation step will still require uncomfortably large values of n. Next we return to the case of known df's.

2.1.1. Choice of m under normality

Let u_p be the upper p-quantile of the standard normal distribution and Φ its df. Under IC X_i has df $F = \Phi$. The upper limit for AVE is directly obtained from (5), since under IC $m^{1/2}\overline{X}$ has a $N(0, 1)$-distribution and hence $\text{UL}(\text{AVE}, m) = u_{mp}$. We take the shift d according to (4) and hence the ARL of IND (the individual chart with $m = 1$) runs from 10 to 200. For the normal distribution the ARL under OoC of AVE is given by

$$\text{ARL}(\text{AVE}, m, d) = \frac{m}{\overline{\Phi}\left(u_{mp} - m^{1/2}d\right)} . \tag{6}$$

In figure 1 the difference between the group chart and the individual chart is shown. On the horizontal axis the ARL of the individual chart is given. We write for it $\text{ARLIND}(d)$ or $\text{ARL}(\text{AVE}, 1, d)$. This is a monotone function of the shift d, see (6) with $m = 1$. On the vertical axis $\text{ARLIND}(d) - \text{ARL}(\text{AVE}, m, d) = \text{ARL}(\text{AVE}, 1, d) - \text{ARL}(\text{AVE}, m, d)$ is plotted for $m = 2, 3, 4, 5$ as indicated shortly in the figure by $\text{IND} - \text{AVE}(m)$ for $m = 2, 3, 4, 5$.

We prefer to plot $\text{ARLIND}(d)$ on the horizontal axis instead of d, since d itself is not important, but its effect on detecting a shift when we compare several values of m. By this presentation it is easily seen what the gain is, for example, for $m = 3$ when the individual chart has ARL $= 40$. This is more easily interpreted than knowing what the gain in ARL is for $m = 4$ and $d = 0.9$, for example. Note that a gain of 10 in ARL is more important when the ARL of the individual chart equals 20 than when it is 200. To quickly see such aspects we have presented in our figures not d, but $\text{ARLIND}(d)$ on the horizontal axis.

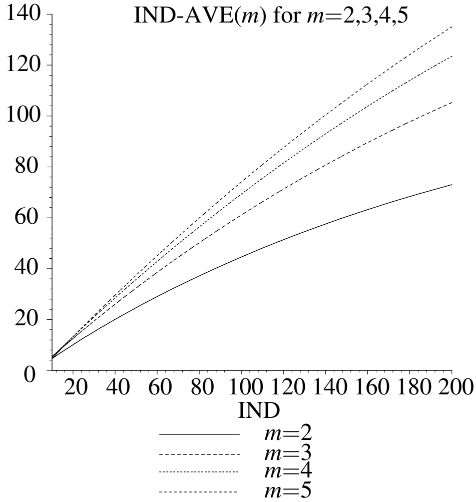

Figure 1. Average-chart under normality.

For a large ARL of IND, that is for small d, a larger value of m gives a non-negligible improvement, but for a shift $d \geq 1$ (corresponding to ARLIND ≤ 55) the differences between $m = 3, 4, 5$ are not very large. From (6) it is evident that the individual chart will beat the m-chart ($m \geq 2$) definitely as soon as $1/\overline{\Phi}\left(u_p - d\right) \leq m$, i.e. when $d \geq \widetilde{d}_1 = u_p - u_{1/m}$. For $p = 0.001$, we obtain $\widetilde{d}_1 = 3.09, 2.66, 2.42$ and 2.25 for $m = 2, 3, 4$ and 5, respectively. Note that these values are quite close to the actual d_1 for which ARL(AVE, m, d) = ARL(AVE, $1, d$) : $d_1 = 2.97, 2.63, 2.40$ and 2.24 for $m = 2, 3, 4$ and 5, respectively. Hence, the individual chart beats the various AVE's for large d, but not really much sooner than in the obvious case where ARL(AVE, $1, d$) $\leq m$.

We conclude that under normality the AVE-chart based on a group of $m = 2, \ldots, 5$ performs better than the individual chart.

2.2. UNI

A second statistic is defined by

$$w\left(X_1, \ldots, X_m\right) = \sum_{i=1}^{m} F(X_i)\,. \tag{7}$$

As $F(X_i)$ is uniformly distributed under IC, we call this UNI. Replacing F, when it is unknown, in $F(X_i)$ by F_n, the empirical distribution function of the Phase I observations $Y_1, \ldots, Y_n$, we get that $nF_n(X_i) = R(X_i) - 1$, where $R(X_i)$ is the rank of X_i among $Y_1, \ldots, Y_n, X_i$. Therefore, UNI is the limiting form of the sum of ranks based statistic. Hence, the statistic in (7) produces a Wilcoxon-type of approach in the estimated version, and as such offers a likely and possibly attractive alternative to the standard parametric approach AVE. In the review on nonparametric charts by Chakraborti *et al.* (2001), rank based charts of this nature by e.g. Bakir and Reynolds (1979) and Hackl and Ledolter (1991, 1992) are mentioned. However, here we shall refrain from going into the estimation aspects and concentrate on the performance under known F. Hence, ranks will remain in the background and uniform-charts are the ones we focus on. So, from now on we again assume F to be known.

Let $U_1, \ldots, U_m$ be a sample from the uniform distribution on (0, 1). Then $\sum_{i=1}^{m} F(X_i)$ has under IC the same df as $\sum_{i=1}^{m} U_i$. To determine the upper limit $\mathrm{UL}(w, m)$ for this case, we begin by observing that for $c \le 1$ we simply have

$$P\left(\sum_{i=1}^{m} U_i > m - c\right) = P\left(\sum_{i=1}^{m} U_i < c\right) = \frac{c^m}{m!} .$$

The latter can be easily seen by induction, using

$$P\left(\sum_{i=1}^{m} U_i < c\right) = \int_0^c P\left(\sum_{i=1}^{m-1} U_i < c - u\right) du .$$

Hence, $\overline{F}_{w,m}^{-1}(mp) = m - c$ with $c = \{m!(mp)\}^{1/m}$, which c will indeed be ≤ 1 for $m \le 5$ and $p = 0.001$. (Of course, for larger c the result can also be readily obtained, but for ease of presentation we concentrate on this most simple case.) Therefore, see (3), $\mathrm{UL}(\mathrm{UNI}, m) = m - \{m!(mp)\}^{1/m}$ and the UNI-chart reads as

$$\sum_{i=1}^{m} F(X_i) > m - \{m!(mp)\}^{1/m} \quad \text{for } m \le 5 \text{ and } p = 0.001 . \tag{8}$$

Note that for $m = 1$ this leads to $F(X_1) > 1 - p$, which indeed agrees with the requirement $X_1 > \overline{F}^{-1}(p)$, used in the individual chart.

2.3. MAX, MIN *and* MIX

Instead of ranks we may also consider order statistics. Under OoC we have a shift to the right and hence in general larger values than under IC. Hence, the first idea might be to take MAX, the largest of $X_1, \ldots, X_m$, that is

$$w(X_1, \ldots, X_m) = \max(X_1, \ldots, X_m) .$$

But this first idea is not so brilliant. In fact, MAX is very close to using m times the individual chart. For instance, when $m = 2$, the difference between $P(\max(X_1, X_2) > x)$ and $P(X_1 > x) + P(X_2 > x)$ is only

$P(X_1 > x, X_2 > x)$ which is very small, since already $P(X_1 > x)$ is small. The advantage of taking a group is hardly used: when we have under IC a group with a large value, as a rule this large value will be the only one in the group (the other ones being "regular") and hence it is just like taking m individual charts.

Under IC we have

$$\overline{F}_{\mathrm{MAX},m}(x) = 1 - P(\max(X_1, \ldots, X_m) \le x) = 1 - \{F(x)\}^m$$

and hence we obtain as upper limit, cf. (3),

$$\mathrm{UL}(\mathrm{MAX}, m) = \overline{F}^{-1}_{\mathrm{MAX},m}(mp) = \overline{F}^{-1}(1 - \{1 - mp\}^{1/m}) . \tag{9}$$

Therefore, the MAX-chart is given by

$$\max(X_1, \ldots, X_m) > \overline{F}^{-1}(1 - \{1 - mp\}^{1/m}) . \tag{10}$$

Indeed, for small p, it holds that $1 - \{1 - mp\}^{1/m} \approx p$, which shows that approximately MAX does nothing but finish the series of m observations in which the individual chart has given a signal.

Taking MIN, the smallest of $X_1, \ldots, X_m$, that is

$$w(X_1, \ldots, X_m) = \min(X_1, \ldots, X_m)$$

looks more promising. When under OoC a shift occurs, we have this shift also in the next observations. When using MIN we take advantage of this effect that in a group the observations intensify each other. That is, already if m observations are pretty large and not necessarily extremely large, this is enough evidence to give an alarm. Here really the group is used. Because under IC

$$\overline{F}_{\mathrm{MIN},m}(x) = P(\min(X_1, \ldots, X_m) > x) = \left\{\overline{F}(x)\right\}^m$$

we get as upper limit, cf. (3),

$$\mathrm{UL}(\mathrm{MIN}, m) = \overline{F}^{-1}_{\mathrm{MIN},m}(mp) = \overline{F}^{-1}(\{mp\}^{1/m}) . \tag{11}$$

Hence, the MIN-chart is given by

$$\min(X_1, \ldots, X_m) > \overline{F}^{-1}(\{mp\}^{1/m}) . \tag{12}$$

A short remark should be made about the two-sided version of the minimum-chart. It uses the maximum when dealing with the lower control limit, thus giving a signal when $\min(X_1, \ldots, X_m) > \mathrm{UL}$ or $\max(X_1, \ldots, X_m) < \mathrm{LL}$. Note that the previously proposed charts in literature are just the other way around, thus being maximum-charts. It will be seen in the next section that during OoC the minimum-chart amply outperforms the individual chart, while the in literature proposed maximum-chart is even slightly worse than the individual chart.

As concerns estimation, (11) immediately shows that MIN reduces estimation of the far tail to estimation of the ordinary tail, because $\{mp\}^{1/m}$ is much larger than p, see also the Introduction, where we encountered $0.144 = (0.003)^{\frac{1}{3}}$ instead of 0.001.

Apart from MAX and MIN we consider a mixture of the two, called MIX, where an alarm is produced if for some probability s and some $0 \le \gamma \le (1-mp)^{1/m}$

$$\min(X_1, \ldots, X_m) > \overline{F}^{-1}(s) \text{ and } \max(X_1, \ldots, X_m) > \overline{F}^{-1}((1-\gamma)\, s)\,. \quad (13)$$

It is immediate to see that during IC the event in (13) has probability $s^m(1-\gamma^m)$. If, as before, the comparison is made fair again by setting this equal to mp, it follows that

$$s = \left(\frac{mp}{1-\gamma^m}\right)^{1/m} \text{ or } \gamma = \left(1 - \frac{mp}{s^m}\right)^{1/m}. \quad (14)$$

Letting γ increase from 0 to $(1-mp)^{1/m}$ it easily follows from (14) that we go from MIN to MAX, cf. also (11) and (9).

The alarm rate of MIX during OoC is given by

$$P\left(\min\left(X_1^{(0)}, \ldots, X_m^{(0)}\right) + d > \overline{F}^{-1}(s), \max\left(X_1^{(0)}, \ldots, X_m^{(0)}\right) + d > \overline{F}^{-1}((1-\gamma)\, s)\right),$$

where $X_1^{(0)}, \ldots, X_m^{(0)}$ still have df F. Direct calculation gives

$$\begin{aligned} &P\left(\min\left(X_1^{(0)}, \ldots, X_m^{(0)}\right) + d > \overline{F}^{-1}(s), \max\left(X_1^{(0)}, \ldots, X_m^{(0)}\right) + d > \overline{F}^{-1}((1-\gamma)\, s)\right) \\ &= \left\{\overline{F}\left(\overline{F}^{-1}(s) - d\right)\right\}^m - \left\{\overline{F}\left(\overline{F}^{-1}(s) - d\right) - \overline{F}\left(\overline{F}^{-1}((1-\gamma)\, s) - d\right)\right\}^m \end{aligned}$$

and hence the ARL of MIX during OoC gives

$$\begin{aligned} &\mathrm{ARL}(\mathrm{MIX}, m, d) \\ &= \frac{m}{\left\{\overline{F}\left(\overline{F}^{-1}(s) - d\right)\right\}^m - \left\{\overline{F}\left(\overline{F}^{-1}(s) - d\right) - \overline{F}\left(\overline{F}^{-1}((1-\gamma)\, s) - d\right)\right\}^m} \end{aligned}$$

with s or γ given by (14). In particular we get for MIN

$$\mathrm{ARL}(\mathrm{MIN}, m, d) = \frac{m}{\left\{\overline{F}\left(\overline{F}^{-1}((mp)^{1/m}) - d\right)\right\}^m}$$

and for MAX

$$\mathrm{ARL}(\mathrm{MAX}, m, d) = \frac{m}{1 - \left\{F\left(\overline{F}^{-1}(1 - (1 - mp)^{1/m}) - d\right)\right\}^{m}}\,.$$

As far as we know the MIX control charts are new and hence in particular the control chart based on the minimum of a group of observations is new. Remarkably, when paying attention to the minimum or the maximum in literature, the first idea of taking the maximum when dealing with an upper control limit is exploited (Howell (1949), Amin and Wolff (1994)), while the minimum is far more promising.

2.4. Example

To conclude this section, we summarize the above by means of an explicit example when $F = \Phi$. We take $p = 0.001$. Hence for $m = 2$ we subsequently have that a signal occurs for

- the individual chart if X_1 exceeds $u_{0.001} = 3.09$ and otherwise (or again) if $X_2 \geq 3.09$,
- AVE if $X_1 + X_2$ exceeds $u_{0.002}\sqrt{2} = 4.07$,
- MIN if both X_1 and X_2 exceed $u_{0.045} = 1.70$,
- MAX if X_1 and/or X_2 exceeds $u_{0.001} = 3.09$,
- MIX (with $\gamma = 1 - 1/8$) if both X_1 and X_2 exceed $u_{0.092} = 1.33$ and at least one of these exceeds $u_{0.012} = 2.27$,
- UNI if $\Phi(X_1) + \Phi(X_2)$ exceeds $2 - (0.004)^{1/2} = 1.94$.

For instance, if we have observations $X_1 = 1.73$ and $X_2 = 2.30$, the individual chart, AVE and MAX will give no signal, while MIN, MIX and UNI give a signal.

3. Comparison

In this section we compare the various methods AVE, see (5), UNI, see (8), MAX, see (10), MIN, see (12), MIX, see (13) and (14), with for each of them the values $m = 1, 2, 3, 4, 5$. For all these charts, the case $m = 1$ produces the same result: the individual chart (IND), giving a signal when $X_1 > \overline{F}^{-1}(p)$. To fix ideas, we take ARLIND $= 1000$ under IC and hence we choose $p = 0.001$, unless stated otherwise. The shifts d are in principle according to (4). The IC behavior is put on a par by matching the ARL's, see (2). Therefore, the comparison concerns the ARL's during OoC: the smaller they are, the better the method with the restriction that small differences are not taken into account.

We present (mostly) the difference in ARL between a certain method at some m and the individual chart IND ($m = 1$). In case of MIX we take $\gamma = 1-(4m)^{-1}$, which turns out to be the nearly best value under normality. With regard to the df's we consider the normal distribution, some members of the family of random normal mixtures and the Gamma(4, 1)-distribution. The random mixtures under consideration are still symmetric and most of them not that far from normality. The Gamma(4, 1)-distribution is skew and farther away from the normal distribution. For all the distributions we take without loss of generality the expectation of the observations under IC equal to 0 and the variance equal to 1.

To quantify the "distance" of the distributions considered here from normality in the context of control charts we calculate the so called model error (ME) for $p = 0.001$. This ME is the error (in terms of probability of a false alarm) for the individual chart when acting as if normality still holds, that is applying the individual normal chart under the given df. In formula $\mathrm{ME} = \overline{F}(u_p) - p = \overline{F}(u_{0.001}) - 0.001$. Note that when normality holds, indeed we get ME $= 0$. For the distributions from Section 3.2 the ME's are given in Table 1.

TABLE 1: *Model error for the distributions from Section 3.2 for* $p = 0.001$. *The unit in the table is* 0.001.

distribution	normal mixture				Gamma
	$\eta = 0.25$		$\eta = 0.5$		
	$\kappa = 2$	$\kappa = 3$	$\kappa = 2$	$\kappa = 3$	
ME	4.1	8.3	2.6	4.3	8.0

Note that this means that when applying simply the individual normal chart for these distributions, FAR is 3.6 to 9.3 times as large as it should be! One could hope that for somewhat larger m the normal upper limit $\mathrm{UL(AVE}, m) = u_{mp}$ can be applied. The argument would be that the standardized mean of $X_1, \dots, X_m$ can be approximated well by the normal distribution. Moreover, a slightly less extreme quantile than for the individual chart is required. The larger m, the better this should work. Unfortunately, this is for $m = 3$ and even for $m = 5$ often too optimistic. We illustrate this by the Gamma distribution of Section 3. First, we take $m = 3$. The 0.003-quantile of the standard normal distribution equals 2.748. For the Gamma distribution with parameters 4 and 1 we get $P(3^{1/2}\overline{X} > 2.748) = 0.0099$, which even is 3.3 times as large as it should be! Next consider $m = 5$. The 0.005-quantile of the standard normal distribution equals 2.576. For the Gamma distribution with parameters 4 and 1 we get $P(5^{1/2}\overline{X} > 2.576) = 0.0115$, which is 2.3 times as large as it should be.

We conclude that application of the average-chart with normal control limits may lead to large relative errors under IC when normality does not hold, even for $m = 3$ or 5.

3.1. Normal distribution

In Section 2.1 we have already discussed the AVE-chart for different values of m. Next we consider MIN for $m = 1, \ldots, 5$ (see figure 2). The same pattern is seen as in figure 1: for small d a substantial gain can be obtained when using larger values of m, but for $d \geq 1$ the differences between $m = 3, 4, 5$ are small. The values of d for which the individual chart is better are somewhat lower than when dealing with AVE, but still they are so large, that no important gain can be obtained. Figures 1 and 2 together show that even for normally distributed observations MIN actually performs quite well, in particular if we compare it with the individual chart. For example, at $d = 1$ the ARL of the individual chart equals 54.6; it is improved with 26.7 by taking MIN with $m = 3$, yielding ARL $= 27.9$; the further improvement when using AVE with $m = 3$ is much less: 8.5, giving ARL $= 19.4$.

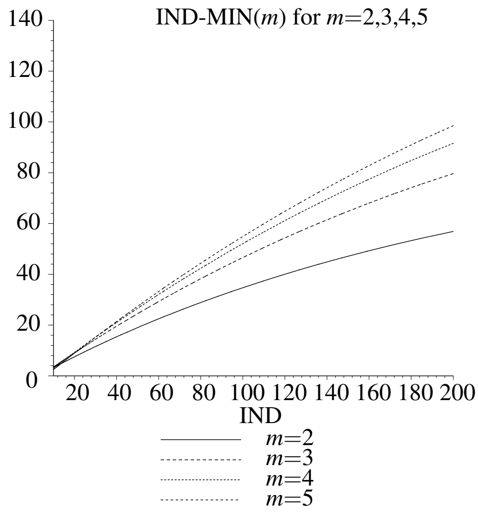

Figure 2. Minimum-chart under normality.

As stated before, MAX is very close to taking m individual charts. In fact, MAX is even slightly worse than IND: for $0.5 \leq d \leq 2$ and $2 \leq m \leq 5$ it turns out that $\text{ARLIND}(d) - \text{ARL}(\text{MAX}, m, d) = \text{ARL}(\text{MAX}, 1, d) - \text{ARL}(\text{MAX}, m, d)$ lies between -2.3 and -0.4. Hence, MAX is no serious option.

Taking MIX with $\gamma = 1 - (4m)^{-1}$, a slight improvement w.r.t. MIN is obtained as is seen in figure 3. For instance, the difference between MIX and MIN at $d = 0.8$, where the individual chart has ARL = 91, goes from 5.1 for $m = 2$ to 6.5 for $m = 5$. Therefore, it seems that this type of improvement does not outweigh the increased complexity of the resulting chart.

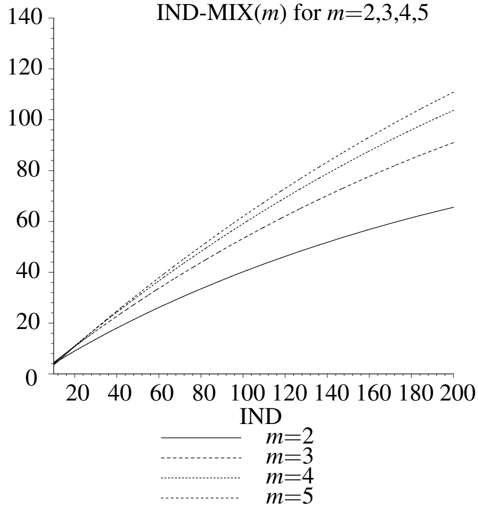

Figure 3. MIX-chart under normality.

Next we consider the UNI-chart. The probability of a signal for this case equals

$$P\left(\sum_{i=1}^{m}\left\{\Phi(X_i^{(0)}+d)\right\} > m - \{m!(mp)\}^{1/m}\right)$$

with $X_1^{(0)}, \ldots, X_m^{(0)}$ a sample from the standard normal distribution, cf. also (8). The resulting expressions are much less explicit than the corresponding ones for AVE and MIN. Hence in this sense UNI is somewhat less attractive to work with. Moreover, for larger m we are more dealing with the shift $E\Phi(X_i^{(0)} + d) - E\Phi(X_i^{(0)})$ than with the shift d of X_i itself. Figure 4 shows the behavior for AVE, MIN and UNI when $m = 2$. It is seen that UNI appears to lie between AVE and MIN: it also loses a bit compared to the optimal AVE, but even less than MIN. This agrees with the intuition according to which $\Phi(X_1) + \Phi(X_2)$ is somewhat closer to $X_1 + X_2$ than $\min(X_1, X_2)$.

In view of the results under normality we restrict ourselves for the remaining distributions to AVE and MIN. We do not further consider UNI for two reasons: firstly, because it seems to be intermediate between AVE and MIN. Secondly, the ARL of UNI is quite complicated; straightforward application of Maple results in serious numerical problems.

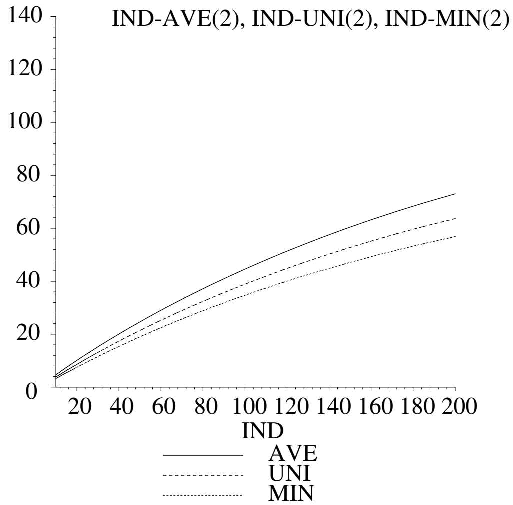

Figure 4. Difference between individual chart and $m = 2$-chart for AVE, UNI and MIN.

3.2. Random normal mixture

In this section we consider the random normal mixture given by

$$F(x) = (1 - \eta)\Phi\left(\frac{x}{\sigma_1}\right) + \eta\Phi\left(\frac{x}{\sigma_2}\right),$$

where the σ_i are such that the variance equals 1, that is $(1 - \eta)\sigma_1^2 + \eta\sigma_2^2 = 1$. Obviously, these distributions are still symmetric around 0, but the tail behavior differs from that of the normal distribution. The normal df occurs when $\kappa = \sigma_2/\sigma_1 = 1$. We will consider here $\kappa = 2, 3$ together with $\eta = 0.25$ and 0.50. Figures 5-8 show the differences between the individual chart ($m = 1$) and the average- and minimum-chart with $m = 2$ under the normal mixtures.

For $\eta = 0.25$ the minimum-chart is superior to the average-chart both for $\kappa = 2$ and 3. This illustrates that the superiority of AVE in the normal case – where it is simply optimal – can indeed be lost if κ moves away from 1. Since we are dealing with the very far tails, for rather unbalanced variances σ_1^2 and σ_2^2 the normal distribution with the larger variance will soon dominate in the mixture. As a consequence the tail behavior of the mixture will be close to that of the normal distribution. For larger values of η like 0.75 and $\kappa = 2, 3$ (not presented here) this is clearly seen. Indeed, in figures 7 and 8 AVE starts to beat MIN again.

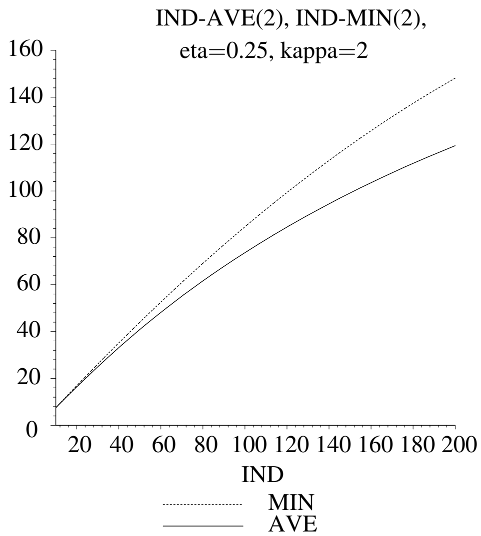

Figure 5. AVE and MIN under normal mixture with $\eta = 0.25$ and $\kappa = 2$.

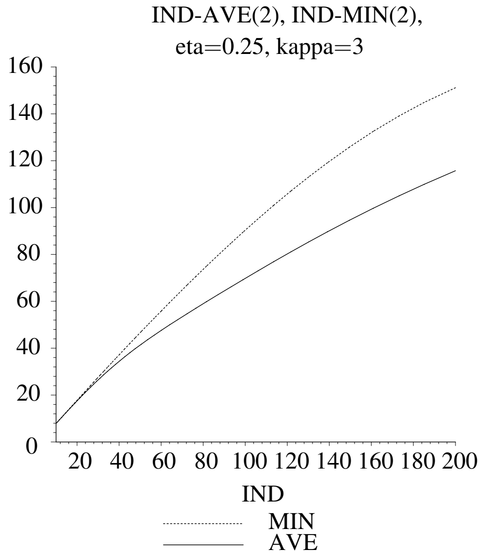

Figure 6. AVE and MIN under normal mixture with $\eta = 0.25$ and $\kappa = 3$.

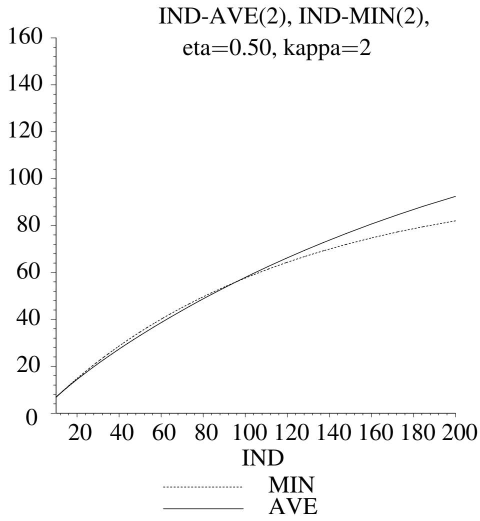

Figure 7. AVE and MIN under normal mixture with $\eta = 0.50$ and $\kappa = 2$.

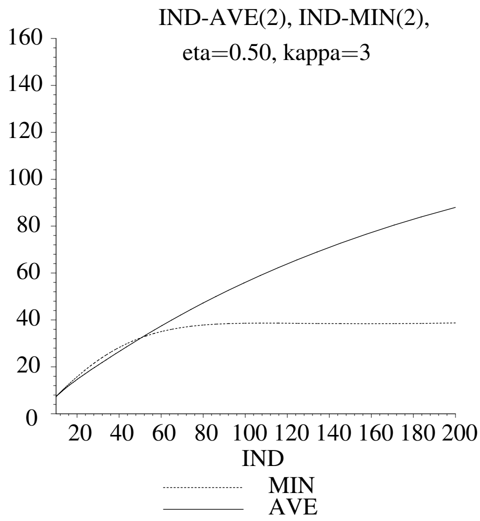

Figure 8. AVE and MIN under normal mixture with $\eta = 0.50$ and $\kappa = 3$.

3.3. Gamma

Here we consider as an example of a skew distribution the Gamma distribution with parameters 4 and 1 having density $\frac{1}{6}x^3e^{-x}$. Its coefficient of skewness equals 1. In figure 9 the difference of the ARL's of AVE and MIN are plotted against the ARL of AVE.

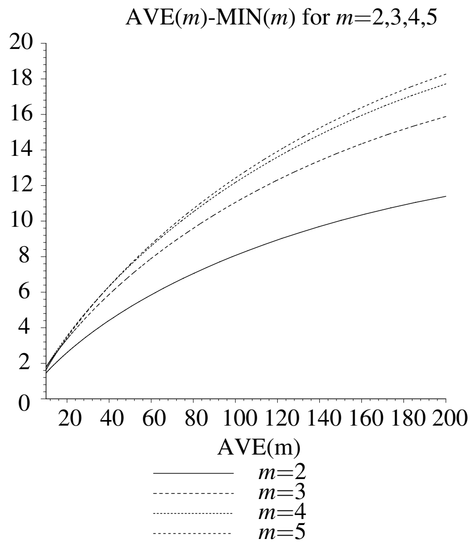

Figure 9. Difference between the ARL's of AVE and MIN.

It is seen that MIN is somewhat better than AVE. Both of them are much better than the individual chart. For instance the ARL of the individual chart at $d = 1$ equals 213.2, the ARL of the MIN-chart at $d = 1$ equals 79.6, 41.1, 26.2, 19.3 for $m = 2, 3, 4, 5$, respectively, while the AVE-chart at $d = 1$ gives 87.1, 47.8, 31.4, 23.3 for $m = 2, 3, 4, 5$, respectively.

4. Conclusions

The results of this paper lead to the following conclusions. The first 3 conclusions concern control charts with F known, the fourth conclusion deals with unknown F.

1. The chart based on a group of $m = 2, \ldots, 5$ in general performs better than the individual chart.
2. The charts based on the average or on the minimum of the group of observations together with the so called uniform-chart, are the most promising ones.
3. Application of the average-chart with normal control limits may lead to large relative errors under IC when normality does not hold, even for $m = 3$ or 5.

4. Accurate nonparametric estimation for the minimum-chart requires much less Phase I observations and is quite straightforward: instead of estimation of a very extreme quantile a rather modest quantile of the distribution has to be estimated. Nonparametric estimation for the average-chart does not lead to a reduction of the required number of Phase I observations and is far more complicated. Also for the uniform-chart accurate nonparametric estimation is far more complicated.

REFERENCES

ALBERS, W. and KALLENBERG, W. C. M. (2004) Estimation in Shewhart Control Charts: Effects and Corrections, *Metrika*, 59, 207–234.

ALBERS, W. and KALLENBERG, W. C. M. (2005) Tail behavior of the empirical distribution function of convolutions, *Math. Methods of Statist.*, 14, 133–162.

ALBERS, W. and KALLENBERG, W. C. M. (2006) Self adapting control charts, *Statist. Neerlandica*, 60, 292–308

ALBERS, W., KALLENBERG, W. C. M., and NURDIATI, S. (2006) Data driven choice of control charts, *J. Statist. Plann. Inference*, 136, 909–941.

ALLOWAY, J. A. and RAGHAVACHARI, M. (1991) Control chart based on the Hodges-Lehmann estimator, *J. Qual. Technol.*, 23, 336–347.

AMIN, R. and WOLFF, H. (1994) The behavior of $\overline{X}$-, R-, and extreme value control charts in the case of mixture alternatives, *Sequential Anal.*, 13, 35–51.

AMIN, R., WOLFF, H., BESENFELDER, W., and BAXLEY, R. JR. (1999) EWMA Charts for the smallest and largest observations, *J. Qual. Technol.*, 31, 189–206.

BAKIR, S. T. and REYNOLDS, M. R. JR. (1979) A nonparametric procedure for process control based on within-group ranking, *Technometrics*, 21, 175–183.

CHAKRABORTI, S., VAN DER LAAN, P., and BAKIR, S. T. (2001) Nonparametric control charts: an overview and some results, *J. Qual. Technol.*, 33, 304–315.

HACKL, P. and LEDOLTER, J. (1991) A control chart based on ranks, *J. Qual. Technol.*, 23, 117–124.

HACKL, P. and LEDOLTER, J. (1992) A new nonparametric quality control technique, *Commun. Statist. Simul. Comput.*, 21, 423–443.

HOWELL, J. M. (1949) Control chart for largest and smallest values, *Ann. Math. Statist.*, 20, 305–309.

PAPPANASTOS, E. A. and ADAMS, B. M. (1996) Alternative designs of the Hodges-Lehmann control chart, *J. Quality Technol*, 28, 213–223.

SARKADI, K. and VINCZE, I. (1974) *Mathematical Methods of Statistical Quality Control*, Academic Press, New York.

WILLEM ALBERS
Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede (The Netherlands)
w.albers@math.utwente.nl

WILBERT C. M. KALLENBERG
Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede (The Netherlands)
w.c.m.kallenberg@math.utwente.nl