Faculty of Mathematical Sciences

University of Twente
The Netherlands

P.O. Box 217
7500 AE Enschede
The Netherlands
Phone: +31-53-4893400
Fax: +31-53-4893114
Email: memo@math.utwente.nl
www.math.utwente.nl/publications

Memorandum No. 1651

Empirical nonparametric control charts: estimation effects and corrections

W. Albers and W.C.M. Kallenberg

October, 2002

ISSN 0169-2690

Empirical nonparametric control charts: estimation effects and corrections

Willem Albers* and Wilbert C.M. Kallenberg

Department of Applied Mathematics
University of Twente
P.O. Box 217, 7500 AE Enschede
The Netherlands

Abstract Due to the extreme quantiles involved, standard control charts are very sensitive to the effects of parameter estimation and nonnormality. More general parametric charts have been devised to deal with the latter complication and corrections have been derived to compensate for the estimation step, both under normal and parametric models. The resulting procedures offer a satisfactory solution over a broad range of underlying distributions. However, situations do occur where even such a larger model is inadequate and nothing remains but to consider nonparametric charts. In principle these form ideal solutions, but the problem is that huge sample sizes are required for the estimation step. Otherwise the resulting stochastic error is so large that the chart is very unstable, a disadvantage which seems to outweigh the advantage of avoiding the model error from the parametric case. Here we analyze under what conditions nonparametric charts actually become feasible alternatives for their parametric counterparts. In particular, corrected versions are suggested for which a possible change point is reached at sample sizes which are markedly less huge (but still larger than the customary range). These corrections serve to control the behavior during in-control (markedly wrong outcomes of the estimates only occur sufficiently rarely). The price for this protection clearly will be some loss of detection power during out-of-control. A change point comes in view as soon as this loss can be made sufficiently small.

Keywords and phrases: Statistical Process Control, Phase II control limits, exceedance probability, empirical quantiles.

2000 Mathematics Subject Classification: 62G15, 62P30

---

*Corresponding author: tel: +31534893816; fax: +31534893069; w.albers@math.utwente.nl.
This research was supported by the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs.

1 Introduction

Suppose we control the mean of a production process using a Shewhart $\overline{X}$-chart: an upper and lower limit are prescribed and an out-of-control signal is given as soon as an observation falls outside the interval determined by these two limits. Standard practice is to assume normality of the distribution involved and to estimate its parameters using so-called Phase I observations. The resulting values are plugged in and it is hoped that the estimated chart behaves well. Several authors have pointed out that often this unfortunately is not the case: both the estimation step and the normality assumption can lead to serious errors. See e.g. Woodall and Montgomery (1999) (p. 379), Ghosh et al. (1981), Quesenberry (1993), Chen (1997), Chakraborti (2000) and Chakraborti et al. (2001). In a series of papers we have analyzed the resulting picture in a systematic manner. The main source of trouble is the fact that p, the probability of getting an out-of-control signal while the process is actually in control, typically is chosen to be very small, like $p = 0.001$. Hence the quantiles involved in the estimation are quite extreme and large relative errors will result, unless the number n of Phase I observations is uncharacteristically large. In Albers and Kallenberg (2000, 2001) (to be denoted for short as AK (2000, 2001) in the sequel) we have analyzed these estimation effects and proposed corrections to get the behavior of the charts under control again. Such remedies work quite well as long as the normality assumption is reasonable. However, quite often there is ample reason to worry about this aspect as well. Again the extremeness of the quantiles involved complicates matters: normality may hold fairly well in the central part of the distribution, but in the tails the relative errors tend to become very large. In Albers, Kallenberg and Nurdiati (2002a, 2002b) (henceforth denoted by AKN (2002a, 2002b)) the second problem has been tackled, using larger parametric models, containing the normal family as a member.

A natural question is why one should stick with such a larger parametric family: the true distribution may still not be in it and the resulting model error may remain unacceptably large. Would it not be better to adopt a fully nonparametric approach? On the one hand, the answer is yes: indeed all problems eventually vanish in a nonparametric chart. A model error is simply not present here, and the stochastic error becomes arbitrarily small as n increases. But on the other hand, for common sample sizes encountered in practice, this is not a lot of help. If we e.g. want to estimate the 0.999-quantile, it is clear that with $n = 100$ we will not get anywhere with a nonparametric estimate. Consequently, it does make sense to look seriously at a larger parametric model, which may successfully bridge the gap between assuming everything (i.e. the distribution simply is normal) or nothing (i.e. the distribution can be anything). Nevertheless, it may happen that after rejection of the normal model, the larger parametric model turns out to be inadequate as well: the actual underlying distribution then is too far away from the model and thus a too large model error does result.

In the latter situation we are back once again at the empirical nonparametric approach and nothing remains but to assume that flexibility with respect to n and/or p is allowed. How much larger should at least one of these quantities be taken before the estimation effects become more mild and such charts start to behave? If the increases involved in n or

p are indeed tremendous, does it help to look for suitable corrections? In the parametric case such adaptations were seen to be useful in bringing the corresponding charts under control again. But the larger the parametric model, the larger such corrections are for given n and p. Will corrections be feasible as well for the even larger fully nonparametric case? Or will the n and p involved still be unreasonably large before such corrections become sufficiently small in the sense that the out-of-control behavior of the corresponding chart is not affected too much? To answer these and similar questions, we need to study the behavior of nonparametric control charts, and this will precisely be the topic of the present paper.

We will restrict attention to the obvious choice based on the empirical quantile function, as this will already provide a clear picture of what can be expected in general. Some previous work on closely related charts can be found in Willemain and Runger (1996) and in Ion et al. (2002). For a recent overview of nonparametric charts in general, see e.g. Chakraborti et al. (2001). Incidentally, as these authors point out, several of the procedures that have been proposed are in fact not truly nonparametric or distributionfree. They are based on a nonparametric estimator, like the Hodges-Lehmann estimator, rather than on $\overline{X}$, but the actual in-control run length distribution involved does depend on the underlying distribution of the observations.

In studying the behavior of the charts, the first thing to note is that due to the estimation the usual performance characteristics of the chart become random. Hence p itself is replaced by a stochastic counterpart P, and the average run length (ARL) $1/p$ likewise by $1/P$. For each characteristic the relative error (e.g. $P/p - 1$) can be studied with respect to aspects such as expectation, standard deviation and exceedance probability, each of which in its own way helps to characterize the behavior of the estimated charts. In section 2 we introduce the nonpararametric chart, after which the next section is devoted to its expectation and bias. Adaptations are suggested to remove the latter, thus improving the performance of the charts in the long run. Some attention is devoted to the out-of-control behavior as well. However, the resulting procedure still is far from satisfactory, which is mainly due to the large variability involved, as is demonstrated in section 4.

This leads in section 5 to a more rigorous approach towards controlling the performance of the chart, using exceedance probabilities. The idea is as follows: in each given application of the control chart procedure, the Phase I observations produce an estimated chart, and thus a value of P, which determines the subsequent behavior of the chart in that application. Instead of merely looking at the average performance of the procedure (i.e. over a long series of different applications), we can also consider the probability with which values of P occur that are too unpleasant in a given sense. For example, these are values of P which are likely to produce a low run length, even if the underlying process is in control. Due to the variability of the nonparametric chart, such probabilities of unpleasant values will typically be unacceptably large for common n and p.

For this situation as well, corrections can be derived to bring the corresponding exceedance probabilities under control. Again the impact of such adaptations on the out-of-control behavior is investigated: it is one thing to largely avoid premature stopping during the in-control period, but this should not be achieved by stopping in general much later once the process has gone out of control. Not surprisingly, the effect here is more pronounced than in section 3, as the changes involved are of a larger magnitude. Hence a considerable price has to be paid in terms of out-of-control performance, in order to guarantee a satisfactory behavior during the in-control phase, unless again of course n or p are sufficiently large. Anyhow, the results obtained allow to strike a proper balance between the following three negative aspects: (i) unacceptable behavior during in-control, (ii) reduced detection power during out-of-control and (iii) using higher n or p than originally intended. Typically, one will figure out how much needs to be sacrificed with respect to (iii) in order to arrive at acceptable results with respect to (i) and (ii). Some examples nicely illustrate what can be expected. For convenience, a short point-by-point description of the algorithm involved is presented in section 6.

Summarizing, we analyze when and how nonparametric charts can be used in a sensible way in situations between two obvious extremes. The first extreme meaning that n is astronomic (like $n = 184830$ on p. 34 in Willemain and Runger (1996)), which makes any problem disappear, any correction superfluous and any price to be paid negligible. The second extreme meaning that n is in the customary range of a few hundreds, while p is one (or a few) tenth(s) of a percent, and we are simply out on a limb. The situation in between is much less clear-cut and the clarity which is obtained here helps to make a balanced choice between sticking with a (large) parametric model or switching to a nonparametric approach. The former suffers from the disadvantage of a nonvanishing model error, whereas the latter in view of its higher variability requires larger, and thus more costly, corrections to keep the in-control behavior under control. A next step will be to use the data in deciding whether to use a parametric control chart or a nonparametric one. This latter topic is treated in Albers, Kallenberg and Nurdiati (2002c).

2 The nonparametric chart

Let X be a random variable (rv) with a continuous distribution function (df) F. For a given, very small p (typically $0.001 \leq p \leq 0.01$), we need an upper limit UL such that $P(X > UL) = p$. (We concentrate on the one-sided case; the two-sided case can be treated in a similar fashion and will lead to completely analogous results.) For any df H we write $\overline{H} = 1 - H$ and H^{-1} and $\overline{H}^{-1}$ for the respective inverse functions. (Note that the inverse is defined unambiguously for continuous and increasing H; for the remaining cases a choice has to be specified.) For known F we thus simply use $UL = F^{-1}(1-p) = \overline{F}^{-1}(p)$. Usually, however, F is unknown and some estimation has to be invoked, using a sample $X_1, ..., X_n$ from F (the Phase I observations) as a starting point. Here we shall consider the case of individual measurements, i.e. the group size $m = 1$. The situation where $m > 1$ is essentially more complicated and will be dealt with in a separate paper. In the normal case, $F(x) = \Phi((x-\mu)/\sigma)$, in which Φ stands for the standard normal df, and thus $UL = \mu + \sigma u_p$, with $u_p = \overline{\Phi}^{-1}(p)$ (e.g. $u_{0.001} = 3.09$ and $u_{0.00135} = 3$). The corresponding $\widehat{UL} = \widehat{\mu} + \widehat{\sigma} u_p$, with $\widehat{\mu}$ and $\widehat{\sigma}$ e.g. the sample mean and sample standard deviation, respectively (see AK (2000, 2001)). Corrected versions of the chart are obtained by replacing u_p by $u_p + c$, for suitably chosen, small correction terms c. A larger parametric family is obtained by taking e.g. $F(x) = K_\gamma((x-\mu)/\sigma))$, with K_γ a member of some suitable family of df's. Clearly this leads to $\widehat{UL} = \widehat{\mu} + \widehat{\sigma}\overline{K}_{\widehat{\gamma}}^{-1}(p)$ (see AKN (2002a, 2002b)). Corrections for this type of chart are again obtained by adding a suitable correction term c to the standardized quantile $\overline{K}_{\widehat{\gamma}}^{-1}(p)$.

In a sense, the nonparametric approach is more transparent than these parametric attempts. Just let $F_n(x) = n^{-1}\#\{X_i \le x\}$ be the empirical df and F_n^{-1} the corresponding quantile function, i.e. $F_n^{-1}(t) = \inf\{x|F_n(x) \ge t\}$. Then it follows that $F_n^{-1}(t)$ equals $X_{(i)}$ for $(i-1)/n < t \le i/n$, where $X_{(1)} < ... < X_{(n)}$ are the order statistics corresponding to $X_1, ..., X_n$. Hence, letting $\overline{F}_n^{-1}(t) = F_n^{-1}(1-t)$, we get

$$\widehat{UL} = \overline{F}_n^{-1}(p) = X_{(n-r)} \tag{2.1}$$

where $r = [np]$, with $[y]$ the largest integer $\le y$. Note that for ordinary p and n, like $p = 0.001$ and $n = 100$, we will have $r = 0$, and thus $\widehat{UL} = X_{(n)}$. See Willemain and Runger (1996) and Ion et al. (2002) for these or closely related charts.

Just as in our previous papers on the parametric case, we need to analyze and possibly correct the behavior of the estimated chart based on $\widehat{UL}$ from (2.1). Let X_{n+1} be another rv from F, then we observe that the role of p in the case of known F will now be played by

$$P = P(X_{n+1} > \widehat{UL}|(X_1, ..., X_n)) = \overline{F}(X_{(n-r)}). \tag{2.2}$$

(In what follows we will also write $P = P(X_{n+1} > \widehat{UL})$, without explicitly stating that we work conditionally on $(X_1, ..., X_n)$). Let $U_{(1)} < ... < U_{(n)}$ denote order statistics for a sample of size n from the uniform df on (0,1), then it is immediate from (2.2) that $P \cong U_{(r+1)}$, with '$\cong$' denoting 'distributed as'.

To judge the performance of the estimated chart, we propose to consider the relative error

$$W = \frac{g(P)}{g(p)} - 1, \tag{2.3}$$

for some suitable function g. An obvious choice is the identity, leading to $W_1 = P/p - 1$, but $g(p) = 1/p$, producing $W_2 = p/P - 1$, is also quite interesting, as it corresponds to the average of the run length RL, given by $ARL = 1/p$. A third choice which is sometimes used (see e.g. Does and Schriever (1992) or Roes (1999), p. 102, 103) follows from $g(p) = 1-(1-p)^k = P(RL \le k)$, where typically $k = [\tilde{\gamma}/p]$ for some small $\tilde{\gamma}$ like 0.1 or 0.2. This possibility we will consider in section 5.

As announced in the introduction, quantities like EW, σ_W and $P(W > \varepsilon)$, for some, usually small, ε, can now be studied. For what values of p and n (or r) are these quantities sufficiently close to 0? And, for given n and p, what kind of modification is needed to make each of these quantities behave as desired after all? The following general observation can already be made: the fact that $P \cong U_{(r+1)}$ strongly suggests that it will suffice to watch the product np, rather than both n and p separately. For, as long as $r = [np] = 0$, no satisfactory results seem feasible, whereas for really large values of np no problems remain. The point is to make explicit what happens in between. Will e.g. r equal 3 or 4 do? Or do we need even 10 and larger? In the next section we shall begin by studying the behavior of EW.

3 Expectation and bias

Let $\delta = np - r$ (and thus $0 \leq \delta < 1$), then it follows from $P \cong U_{(r+1)}$ that

$$EP = \frac{r+1}{n+1} > \frac{r+\delta}{n} = p \tag{3.1}$$

for $0 \leq \delta < (n-r)/(n+1)$. Hence P has a positive bias, unless δ is very nearly 1. It is immediate that $W_1 = P/p - 1$ satisfies

$$EW_1 = \frac{\frac{n-r}{n+1} - \delta}{r+\delta} = \frac{\frac{1-\delta}{p} - 1}{n+1}. \tag{3.2}$$

Note the discrete character of this relative error: if we let n increase for some given p, the expression in (3.2) gradually decreases, jumping upwards whenever np becomes integer again (i.e. when $\delta = 0$) to $(n-r)/\{(n+1)r\} = (1-p)/\{(n+1)p\}$. This maximal value behaves like $(1-p)/r \approx 1/r$ (supposing of course that $r > 0$).

Hence it is straightforward to analyze for which combinations of n and p the chart will start to behave with respect to (3.2). To begin with, as $n \to \infty$, so does $r = [np]$ for given p, and the bias eventually becomes negligible. Thus, for $p = 0.001$, the relative change is at most 1% as soon as $n \geq 10^5$ (e.g. $n = 184830$), but on the other hand, we require $n \geq 1000$ before r is even positive. The situation in between these two extremes is also still quite transparent. For example, for $n \geq 5000$ we have $r \geq 5$, and then the relative error will be at most 20%, which might for example be reasonable for practical purposes. This is still a very large sample size; an alternative of course is to raise the value of p to e.g. $p = 0.01$. The n required to get the same result then obviously reduces to $n = 500$, which is still not really small. Consequently, even with respect to a rather mild criterion, only concerning the bias involved, reasonable behavior seems to require considerably higher values of n and p than the customary ones.

For the ARL similar conclusions can be drawn. As $E(1/U_{(r+1)}) = n/r$ (and thus $E(1/U_{(1)}) = \infty$, cf. Willemain and Runger (1996), where it is also noted that r has to be at least 1 here), we observe that typically

$$E\left(\frac{1}{P}\right) = \frac{n}{r} \geq \frac{n}{r+\delta} = \frac{1}{p}. \tag{3.3}$$

(The phenomenon of both P and $1/P$ being positively biased is already well-known from the parametric case, see e.g. AK (2000), p. 6 or Quesenberry (1993), p. 245.) Clearly, for

$W_2 = p/P - 1$ we have $EW_2 = \delta/r < r^{-1}$, and precisely the same comments as for W_1 can be given.

Since the behavior for small r indeed turns out to be unsatisfactory, the next question is what types of corrections can be proposed to improve matters. One remedy is to invoke the aforementioned flexibility required with respect to n and/or p in a very simple way: just alter p such that e.g. $p = (r+1)/(n+1)$ for some $r \geq 0$ to arrive at $EP = p$. Or, alternatively, set $p = r/n$ for some $r \geq 1$ to obtain $E(1/P) = 1/p$. But do realize that this may seem to achieve more than it actually does. To give a simple example, first let $r = 1$ and choose $p = 2/(n+1)$. Then $EP = p$, but $E(1/P) = n = \{2n/(n+1)\}(1/p)$, i.e. $EW_2 = (n-1)/(n+1) \approx 1$. Hence for $p = 0.001$ and e.g. $n = 1999$, the expected relative error in the ARL is still about 1. If for the present value of $p = 0.001$ we increase n by just 1 to 2000, the picture revolves: then r becomes 2, suddenly EW_2 drops to 0, but EW_1 becomes almost 1/2. Hence the chart is very unstable due to its discrete character, and moreover the errors involved are also very large. Once again, the phenomenon is due to the fact that n is large but r is not. If r is large as well, the problem neatly dissolves: for $p = (r+1)/(n+1)$, we have $EW_1 = 0$ but $EW_2 = (n-r)/\{r(n+1)\}$, while for $p = r/n$ it is precisely the other way around.

For parametric charts the picture is quite different. Obviously, the complications due to discreteness are not present there. In addition, relative errors of a similar magnitude as in the example above do occur as well, but for much smaller sample sizes. First consider the uncorrected normal chart based simply on $\widehat{\mu} + \widehat{\sigma} u_p$. In this case we e.g. observe from Table 1 in AK (2000) that for $p = 0.001$ we would obtain $EW_2 = 1$ for $n \approx 65$ and $EW_1 = 1/2$ for $n \approx 70$. Hence for such sample sizes, corrections are called for here as well. (In fact, these are derived in AK (2000) and shown to work well as soon as $n \geq 40$, but let us not digress any further in this direction.) For this same p and $n = 500$, the relative errors in the uncorrected chart have already gone down to about 7% for EW_1 and about 8% for EW_2, and the use of corrections has become superfluous. This conforms with the common recommendation to take at least 300 observations for estimating the parameters under normality.

If we go to a larger parametric model and replace u_p by some $\overline{K}_{\widehat{\gamma}}^{-1}(p)$, as described in section 2, we can readily obtain an example from AKN (2002a). For the particular choice considered there, which is demonstrated to work well over a broad range of underlying distributions, we observe from Table 3 that if normality happens to be true after all, the uncorrected chart leads to an EW_1 of 117%, 43% and 21% for $n = 100$, 250 and 500, respectively. Hence corrections are certainly needed. (Again, in this particular paper these are derived and it is demonstrated that the corrections are having the desired effect, but we will not dwell on that point here.) From Table 1 in that same paper we see that the uncorrected normal chart would have produced 36%, 13% and 7% (the latter case we already encountered above) for these same sample sizes. Clearly, this is quite a bit better than what is achieved with the more general parametric chart. But the price for this gain is also immediately evident from the two tables mentioned: if we move towards non-normal distributions, EW_1 varies wildly for the normal chart (e.g. $EW_1 > 11$ occurs for the cases considered at $n = 100$ and as this is due to the model error rather than the stochastic error, there is not much improvement for larger n). On the other hand, its parametric counterpart keeps the damage much more limited (e.g. EW_1 is at most 2.5 at $n = 100$). Hence the premium paid for the latter chart is rather large, but nevertheless well-spent in view of the erratic behavior of the model error for the normal chart outside the normal model.

Note that this discussion of parametric charts helps to better understand what happens in the nonparametric case. In that situation the model error is eliminated by estimating the distribution in a nonparametric way, rather than by just adding a single third parameter to the normal model. When, as we saw, even this latter extension already leads to considerably higher values of the relative errors involved, it becomes less surprising that in the nonparametric case the growth is so tremendous that excessive values of n and p are needed before this approach becomes of potential use. Of course, it should be remarked that we have concentrated on the maximal errors that occur. In the nonparametric case, a minimal error equal to zero can be achieved by just taking a lucky combination of n and p, whereas in the parametric case corrections are always needed until n and p are sufficiently large.

Next we shall consider a second way of introducing corrections, which will turn out to be of use also in section 5. In the above the bias was removed by adapting p to the n at hand. If we want to stick to a given p, unbiasedness can be achieved in a relatively simple way by randomizing between consecutive order statistics. Let V be independent of $(X_1, \ldots, X_n, X_{n+1}, \ldots)$, with $P(V = 1) = 1 - P(V = 0) = \lambda$ and replace $\widehat{UL}$ from (2.1) for example by

$$\widehat{UL} = (1 - V)X_{(n+1-r)} + VX_{(n-r)}, \tag{3.4}$$

where we use the convention $X_{(n+1)} = \infty$. Hence for $r = 0$, a signal only results with probability λ in case $X_{n+1} > X_{(n)}$. As $P = P(X_{n+1} > \widehat{UL}) = (1 - V)P(X_{n+1} > X_{(n+1-r)}) + VP(X_{n+1} > X_{(n-r)}) \cong (1 - V)U_{(r)} + VU_{(r+1)}$ (we now work conditionally on $(X_1, \ldots, X_n, V)$, cf. the remark following (2.2)), it is immediate that $EP = (r+\lambda)/(n+1)$. This equals $p = (r + \delta)/n$ for

$$\lambda = \frac{r + (n + 1)\delta}{n} = p + \delta. \tag{3.5}$$

Consequently, unless $\delta > 1 - p$, replacement of $X_{(n-r)}$ by the next higher order statistic with suitable probability will produce $EP = p$. Singling out the case $p \leq 1/(n + 1)$, in which $r = 0$ again, we see that in the event $\{X_{n+1} > X_{(n)}\}$, which has probability $1/(n+1)$, a signal is produced with probability $\lambda = (n + 1)p$ only, thus bringing down the expected signal rate to p. Incidentally, note that this example shows that a deterministic mixture $(1 - \lambda)X_{(n+1-r)} + \lambda X_{(n-r)}$, which might look as a more natural counterpart of (3.4), runs into problems as $X_{(n+1)} = \infty$.

We can also combine $X_{(n-r)}$ with the next lower order statistic, i.e. replace $X_{(n+1-r)}$ by $X_{(n-1-r)}$ in (3.4). Then $EP = p$ can be achieved for $\delta > 1 - p$, but more importantly, this combination serves to obtain $E(1/P) = 1/p$. Actually, as $1/P \cong 1/\{(1-V)U_{(r+2)} + VU_{(r+1)}\} = (1-V)/U_{(r+2)} + V/U_{(r+1)}$, we have for $r \geq 1$ that $E(1/P) = n\{(1-\lambda)/(r+1) + \lambda/r\}$, which equals $1/p = n/(r+\delta)$ for

$$\lambda = \frac{r}{(r+\delta)}(1-\delta) \tag{3.6}$$

with again $r \geq 1$. For $r = 0$, relation (3.6) still works, in the sense that it indeed produces the only feasible value $\lambda = 0$, but now $E(1/P) = E(1/U_{(2)}) = n < 1/p$, which seems unavoidable since $E(1/U_{(1)}) = \infty$.

Hence once more the extremes are clear: for $r = 0$ the corrected chart becomes outright awkward, while for large r it is already intuitively evident that taking a random mixture of two adjacent order statistics, rather than just one of them, will make little difference. In order to further illustrate the situation between these two ends of the scale, we shall now address the interesting question of what happens in the out-of-control situation, with particular attention for the effect of the corrections studied above. Consequently X_{n+1} now comes from a shifted df $F(x-\Delta)$, where Δ typically is such that $p_1 = \overline{F}(\overline{F}^{-1}(p) - \Delta)$ may still be small, but not extremely so, like p. If we take the choice $\widehat{UL} = X_{(n-r)}$ from (2.1) as our starting point again, it follows that P from (2.2) transforms into $P \cong \overline{F}(\overline{F}^{-1}(U_{(r+1)}) - \Delta)$, and thus EP can be approximated by $\tilde{p}_1 = \overline{F}(\overline{F}^{-1}(q) - \Delta)$, where $q = (r+1)/(n+1)$. As this $\tilde{p}_1$, just like p_1, is not extremely small, the relative error involved will be reasonably small compared to the relative error during the in-control situation.

Moreover, it is also easy to see what happens if the standard $X_{(n-r)}$ is replaced by either $X_{(n+1-r)}$ or $X_{(n-1-r)}$. The change in EP in going from $X_{(n-r)}$ to $X_{(n\pm1-r)}$ will approximately equal $-/+\ w$, with

$$w = \frac{f\Big(\overline{F}^{-1}(q) - \Delta\Big)}{(n+1)f\Big(\overline{F}^{-1}(q)\Big)}, \tag{3.7}$$

where $f = F'$ is the density involved. Clearly, if the replacement only happens with probability $(1-\lambda)$ (cf. (3.5) and (3.6)), the effect will have size $(1-\lambda)w$, and the corresponding relative error will be $(1-\lambda)w/\tilde{p}_1$ (approximately in both cases!). This will become reasonably small somewhat sooner than in the in-control situation, where the relative errors were seen to behave like r^{-1}. On the other hand, here as well, it is easy to see that when changing $X_{(n-r)}$ into $X_{(n+1-r)}$ things go wrong as r gets close to or even equal to zero, as in that situation $f(\overline{F}^{-1}(q))$ becomes very small and moreover (3.7) no longer provides a reasonable approximation. As $X_{(n+1-r)} = \infty$ for $r = 0$, we simply have that $\overline{F}(X_{(n+1-r)} - \Delta) = 0$ and thus EP is reduced, both under out-of-control and in-control, by the factor $\lambda = (n+1)p$.

To be a bit more explicit, as well as to provide some illustration, we shall briefly compare these results to the situation in the parametric case. Keeping things as simple as possible, just let $F = \Phi$ and use $\widehat{UL} = \widehat{\mu} + \widehat{\sigma}u_p$. (cf. section 2). With X_{n+1} coming from $\Phi(x-\Delta)$, we then obtain that $EP \approx p_1 = \overline{\Phi}(u_p - \Delta)$, where again the relative error committed is reasonably small because this p_1, just like the one above in the nonparametric case, is supposed not to be (very) small. Hence for this specific choice of F, think of values of Δ between 1 and 3. If we now replace u_p by $u_p + c$ (cf. once more section 2), the change in EP will to first order equal $-c\phi(u_p - \Delta)$ (cf. e.g. (4.4) from AK (2000)). Consequently, the relative change caused by using c will approximately be $-ck(u_p - \Delta)$, where

$$k(x) = \frac{\phi(x)}{\overline{\Phi}(x)}. \tag{3.8}$$

The function $k(x)$ can be approximated for $0 \leq x \leq 3.09 = u_{0.001}$ by $4(1+x)/5$ (see also (17) and (18) from AK (2001)). Finally, the correction needed to get $EP = p$ during in control for the normal case equals $c = u_p(u_p^2 + 2)/(4n)$ (see (3.6) in AK (2000)), which in combination with (3.8) shows that the size of the relative error during out-of-control due to this correction approximately equals

$$u_p(u_p^2 + 2)\frac{1 + u_p - \Delta}{(5n)}. \tag{3.9}$$

For $p = 0.001$ this boils down to $7.15(4.09 - \Delta)/n$, which clearly becomes reasonably small quite soon. E.g. take $\Delta = 1.81$, which leads to a value $\overline{\Phi}(3.09 - 1.81) = 0.10$ for p_1. For this case the relative change is about $16/n$.

This parametric example is of some interest on its own, but of course it is primarily meant to provide an explicit comparison to the nonparametric situation. Thus, if there we also let $F = \Phi$, the expression from (3.7) translates into $w = \phi(u_q - \Delta)/\{(n+1)\phi(u_q)\}$. As moreover (3.8) implies that $k(u_q) = (n+1)\phi(u_q)/(r+1)$, it readily follows that the size of the nonparametric relative change simply reduces to approximately $(1-\lambda)w/\tilde{p}_1 = (1-\lambda)k(u_q - \Delta)/\{(r+1)k(u_q)\}$. The final step is to approximate this expression in its turn by

$$(1-\lambda)\frac{1 + u_q - \Delta}{(r+1)(1+u_q)}. \tag{3.10}$$

The result in (3.10) is very simple and transparent. The general observation above, according to which the relative error 'will become reasonably small somewhat sooner than in the in-control situation, where the relative errors were seen to behave like r^{-1}' can now be made explicit for this special case: $(r+1)^{-1}$ is reduced by the factor $\{1 + u_q - \Delta\}/(1 + u_q)$, which will vary between 1/4 and 1/2 for customary values of n, p and Δ. As concerns the factor $(1-\lambda)$, it is immediate from (3.5) that this will keep returning to (almost) 1 as n increases for given p. To give a numerical example, for the standard $p = 0.001$, let $n = 5000$, and thus obtain $r = 5$, $\delta = 0$, $1 - \lambda = 1 - p \approx 1, u_{6/5001} = 3.04$, which together produces $(1 - \Delta/4.04)/6$ as the outcome of (3.10). For $\Delta = 1.81$, which is the value used in the parametric example above, this boils down to 9.2%, which sounds acceptable. But note that this requires n to be as large as 5000; in comparison, the parametric value $16/n$ produces such an outcome already for $n \approx 175$. Indeed, comparing (3.9) and (3.10) shows that both relative errors have a factor $1 + u_s - \Delta$ in common (with s either equal to p or to q), but apart from that they differ markedly. The parametric c is small and behaves like n^{-1}, whereas the effect of shifting to the next order statistic introduces a factor $1/\{(n+1)f(\overline{F}^{-1}(q))\}$, which behaves like r^{-1}.

It is of course quite interesting to study what happens to the ARL as well. But to avoid repetition, we once more point out that, although the analysis above has been given for P, it is easily verified that the size of the relative error is (again approximately) the same for $1/P$. Just note that $|g'(p)/g(p)| = 1/p$ for both $g(p) = p$ and $g(p) = 1/p$ (cf. (2.3)). Hence again completely similar conclusions hold for this case.

4 Variation

In the previous section we have seen that estimation unfortunately introduces bias which is definitely not negligible for the usual small values of r. Adaptations of the chart can easily be devised to remove such bias, but this obviously does not solve the problem to complete satisfaction. Removing bias from P means increasing bias in $1/P$ and vice versa. Moreover, the effect under out-of-control is also considerable as long as r is small. The source of these troubles is the fact that the estimated nonparametric chart seems to contain too much variation to be very useful for small r. Apparently the balance is lost: the model error from the parametric models has been eliminated, but the resulting stochastic error more than spoils the benefit.

To make this feeling explicit, we shall now consider the standard deviation σ_W of the relative error W from (2.3) for the cases $W_1 = P/p - 1$ and $W_2 = p/P - 1$. As $\mathrm{var}(U_{(r+1)}) = (r+1)(n-r)/\{(n+2)(n+1)^2\}$ while $\mathrm{var}(1/U_{(r+1)}) = n(n-r)/\{(r-1)r^2\}$, it follows readily that

$$\sigma^2_{W_1} = \frac{n^2(n-r)(r+1)}{(n+1)^2(n+2)(r+\delta)^2}, \qquad \sigma^2_{W_2} = \frac{(n-r)(r+\delta)^2}{n(r-1)r^2}. \tag{4.1}$$

(Note that for W_2 we even need $r \geq 2$ now; cf. Willemain and Runger (1996)). From (4.1) a similar pattern is observed as in the previous section, where making $EW_1 = 0$ led to $EW_2 = (n-r)/\{(n+1)r\}$ and vice versa. The variances behave like r^{-1} or $(pn)^{-1}$, which is fine for n large and p (not too) small, or n extremely large and p very small, but not otherwise.

Next we consider parametric charts and show that the situation here is somewhat different. Suppose that instead of $\widehat{UL} = X_{(n-r)}$ from (2.1) we use some parametric $\widehat{UL}$. Then $P = \overline{F}(\widehat{UL})$ in this case leads to

$$\sigma^2_{W_1} \approx \tau^2 \frac{f^2(\overline{F}^{-1}(p))}{p^2} \tag{4.2}$$

with $\tau^2 = \mathrm{var}\ (\widehat{UL})$ of order n^{-1}. Now typically $f(\overline{F}^{-1}(p))/p$ may grow somewhat as p becomes small, but not as fast as $p^{-1/2}$. Hence $\sigma^2_{W_1}$ from (4.2) will behave better than

$(pn)^{-1}$. As $|g'(p)/g(p)|$ is the same for both $g(p) = p$ and $g(p) = 1/p$ the same holds for $\sigma^2_{W_2}$.

To be more explicit again consider the example $\widehat{UL} = \widehat{\mu} + \widehat{\sigma} u_p$ from the normal case. From AK (2001) we have that $\tau^2 \approx (u_p^2 + 2)/(2n)$ while the fact that $\overline{\Phi}(x) \approx \phi(x)/x$ for x large shows that $\phi(u_p)/p \approx u_p$, thus producing $\sigma^2_{W_1} \approx u_p^2(u_p^2 + 2)/(2n)$. If we let $p = 0.00135$, we get $u_p = 3$ and thus $\sigma_{W_1} \approx 7n^{-1/2}$, while in the nonparametric case $\sigma_{W_1} \approx (pn)^{-1/2} \approx 27n^{-1/2}$. Indeed these results nicely agree with the simulations from Ion et al. (2002) (see Tables 4 and 8), which also exhibit this difference in behavior between the two types of chart. Yet another comparison is obtained by Table 1 from Willemain and Runger (1996), where the ARL case is illustrated and hence σ_{W2} from (4.1) is used. Using $p = 0.00270$, and thus $1/p = 370$, the standard deviation of the ARL is brought down all the way to 16.6, requiring $n = 184830$ and consequently $r = 499$. For $n = 1482$, however, r is merely 4 and the standard deviation equals 214.

5 Exceedance probabilities

To summarize the situation up to now, we know how to remove the bias in P (or $1/P$), but the variation around the obtained correct expectation p (or $1/p$) remains quite large. Hence in specific applications of the procedure we must reckon with the occurrence of considerably 'wrong' values. If we want to control the frequency with which such accidents occur, exceedance probabilities are the instrument to use.

Hence for the relative error W from (2.3) we shall now move from EW and σ_W (for $\varepsilon > 0$) to $P(W > \varepsilon)$ for increasing g and $P(W < -\varepsilon)$ for decreasing g. For $W_1 = P/p - 1$ this simply produces $P(P > p(1+\varepsilon))$, while for $W_2 = p/P - 1$ we obtain $P(P > p/(1-\varepsilon))$. For ε small, this is virtually the same, as $|g'(p)/g(p)| = 1/p$ in either case. In fact, for increasing g we have that (2.3) becomes $P(P > p(1 + \tilde{\varepsilon}))$, with

$$\tilde{\varepsilon} = \frac{g^{-1}(g(p)(1+\varepsilon)) - p}{p} \approx \varepsilon \frac{g(p)}{pg'(p)}. \tag{5.1}$$

In the third choice mentioned in section 2, we have $g(p) = P(RL \le k) = 1 - (1-p)^k$ with typically $k = [\tilde{\gamma}/p]$ for some small $\tilde{\gamma}$ like 0.1 or 0.2. From (5.1) it follows that here $\tilde{\varepsilon} = \varepsilon(1-p)\{(1-p)^k - 1\}/(kp) \approx \varepsilon(e^{kp} - 1)/(kp) \approx (1 + \tilde{\gamma}/2)\varepsilon$. Consequently, for each of these three criteria it makes sense to study $P(P > p(1+\varepsilon))$: just use ε itself, or $\varepsilon/(1-\varepsilon)$, or $(1 + \tilde{\gamma}/2)\varepsilon$, the latter being an approximate solution, valid for ε small.

Again we start with the standard nonparametric chart based on $\widehat{UL} = X_{(n-r)}$ (cf. (2.1)). As $P \cong U_{(r+1)}$, we simply obtain that

$$P(U_{(r+1)} > p(1+\varepsilon)) = B(n, p(1+\varepsilon), r) \approx Po(np(1+\varepsilon), r), \tag{5.2}$$

where $B(n, \tilde{p}, k)$ stands for the cumulative binomial probability $P(Y \le k)$, with Y $bin(n, \tilde{p})$ and $Po(\lambda, k)$ likewise for the cumulative Poisson probability $P(\widetilde{Y} \le k)$, where $\widetilde{Y}$ is Poisson with parameter λ. Here we use the well-known relation $P(U_{(k)} > \tilde{p}) = P(Y < k)$, as well as the fact that as n is typically large, while $r = [np]$ is not, a Poisson approximation will work extremely well in (5.2).

To begin with, we check that the behavior at the two extremes is again as expected. Indeed, as $n \to \infty$ for fixed p, all is well: $P(U_{(r+1)} > p(1+\varepsilon)) \to 0$. According to Hoeffding (1963), we in fact have that $B(n, p(1+\varepsilon), r) \le \exp\{-2(np(1+\varepsilon) - r)^2/n\} \le \exp(-2np^2\varepsilon^2)$ and thus the convergence is even exponentially fast. But clearly, in most practical applications $np^2\varepsilon^2$ will be very small, rendering the bound correct but useless. In fact, not only the bound will become large, but the same will hold for the actual probability involved as well. To illustrate the opposite end of the scale, let us now assume that $n < p^{-1}$ and thus $r = 0$. Then the exceedance probability in (5.2) equals $\{1 - p(1+\varepsilon)\}^n \approx \exp\{-np(1+\varepsilon)\}$, which attains some prescribed value α for $n = (\log \alpha^{-1})/(p(1+\varepsilon))$. Letting $p = 0.001$ again and choosing $\varepsilon = 9$ (hence P is off by a factor at least 10), we obtain $n = 100 \log \alpha^{-1}$, which e.g. produces $n = 300$ for $\alpha = 0.05$. In the normal case (see AK (2001), p. 7), by comparison we have $n \approx 9.90u_\alpha^2$, leading to $n = 27$ for $\alpha = 0.05$. Hence really huge errors occur in the parametric case as well, but only at relatively small sample sizes. In the nonparametric case, this will still be the case for much larger sample sizes. Note that just as in section 3 was done for the expectation case, we could use larger parametric families to produce an example (this time from AKN (2002b)) to show that in the exceedance case as well such an extended family nicely fits between the normal and nonparametric ends of the scale. However, for brevity we merely point out this possibility and refrain from actually carrying it out.

The case between the two extremes is still quite easy to analyze. It is less trivial than in the expectation case from section 3, where the maximal bias was seen to behave simply as r^{-1}. But the Poisson probability from (5.2) also readily shows what to expect in a given configuration. To give an example, let as before $p = 0.001$ and $n = 5000$, and thus $r = 5$, then we deal with $Po(5(1+\varepsilon), 5)$, which equals 0.45, 0.19 and 0.067 for ε equal 0.2, 0.6 and 1.0, respectively. Even if we raise n to 10000, the results still do not look very satisfactory: we see that $Po(10(1+\varepsilon), 10)$ equals 0.35, 0.18 and 0.077 for ε equal to 0.2, 0.4 and 0.6, respectively. Hence, as expected, with respect to the present criterion, the behavior of the uncorrected chart remains unsatisfactory for even larger n than in section 3.

Consequently, there is ample reason to look for suitable corrections. Fortunately, from section 3 it is immediately clear how to modify $\widehat{UL} = X_{(n-r)}$ such that the associated exceedance probability satisfies some desired upper bound α. Instead of (3.4) we will now use

$$\widehat{UL} = (1-V)X_{(n+k+1-r)} + VX_{(n+k-r)}, \tag{5.3}$$

where k is the smallest integer for which $Po(np(1+\varepsilon), r-1-k) \le \alpha$ and $\lambda = P(V=1)$ is subsequently chosen such that $P(X_{n+1} > \widehat{UL}) = \alpha$. Indeed, the resulting chart, which randomizes between $X_{(n+k-r)}$ and $X_{(n+k+1-r)}$, guarantees that relative errors in P in excess of ε will only occur in a fraction α of cases. Hence in effect we copy the solution from section 3, but instead of merely using an adjacent order statistic, we also shift k steps: a superior form of protection requires a stronger correction. Clearly, the price to be paid for such protection will be far from negligible, unless again r is quite large. For example, note that during the in control situation EP will now decrease from $(r+1)/(n+1)$ to $(r+\lambda-k)/(n+1)$. Hence rather than removing the positive bias there, we replace it by a typically even larger negative one, in order to be able to satisfy the exceedance criterion.

To study the effect of the correction defined through (5.3), let us first look at the extremes again. As $n \to \infty$ for fixed p, a normal approximation e.g. shows that $Po(np(1+\varepsilon), r) \to 0$ and thus k in (5.3) will eventually become 0 (actually, it might even become negative, but by then no risk of dealing with realistic n and p exists anymore, so we will not bother about this possibility). Turning to the opposite end, a simple example illustrates that for $r=0$ things are still awkward. Let $p=0.001, n=800$ and $\varepsilon=0.1$, then we obtain $Po(0.88,0) = \exp(-0.88) = 0.42$. Hence if we choose $\alpha=0.2$, it means $\widehat{UL} = X_{(n)}$ with probability $\lambda = 0.48$ and $\widehat{UL} = \infty$ otherwise. Hence this adaptation clearly also decreases the out-of-control expected $1/P$ by about a factor 2. In comparison again, for the normal case in AK (2001) (see section 4) and the same p, ε and α as used here, the factor by which the ARL is increased equals about 1.3. And that while $n=100$, rather than the $n=800$ used here! Clearly the difference is again tremendous.

For the analysis of what happens between the extremes $r \to \infty$ and $r=0$, we shall consider the impact of the correction (5.3) on the out-of-control probability. In analogy to section 3 we immediately obtain that approximately a relative error

$$\frac{(k+1-\lambda)w}{\tilde{p}_1} \tag{5.4}$$

will result from applying such a correction, with w as in (3.7), λ as defined through (5.3) and once more $\tilde{p}_1 = \overline{F}(\overline{F}^{-1}((r+1)/(n+1)) - \Delta)$. Clearly, due to the presence of k, the expression in (5.4) will become small even slower than its counterpart from section 3. In the latter only a factor $(1-\lambda)$ occurs, which obviously is at most 1. Some feeling for what actually happens is again obtained by considering a numerical example (for simplicity we shall as much as possible continue with the one from section 3). Take the usual $p=0.001$, let $n=5000$, and thus $r=5$, and fix ε at 0.2. Then we obtain for $Po(6,z)$ the values 0.45, 0.29, and 0.15 for $z=5,4,$ and 3, respectively. Hence if we choose $\alpha=0.2$, it follows that $k=1$ and $\lambda=0.36$: the uncorrected $\widehat{UL} = X_{(4995)}$ is replaced through (5.3) by $X_{(4996)}$ with probability 0.36 and by $X_{(4997)}$ otherwise. In this way, the realized P differs from the intended p by more than 20% in at most 20% of the times this procedure is applied. The price to be paid for this protection will according to (5.4) be given by roughly $1.64w/\tilde{p}_1$. The behavior of $w/\tilde{p}_1$ was already illustrated in section 3. For example, the 9.2% obtained there will now become 15.1%, which still sounds reasonable.

Of course, the n involved is again large. In particular, it is much larger than would be needed in the normal case. In order to illustrate this through a specific comparison, we first quote from section 3 (cf. (3.8) and (3.9)) that a correction c will produce with respect to the uncorrected case a relative error with size approximately $4c(1+u_p-\Delta)/5$. Next we obtain from AK (2001) (see e.g. (10)) that for the present purpose the relevant correction, which thus keeps the exceedance probability below α, approximately satisfies $c = \tau u_\alpha - \varepsilon/u_p$, where (cf. 4.2)) $\tau^2 = var(\widehat{UL}) \approx (u_p^2+2)/(2n)$ for the normal case $\widehat{UL} = \hat{\mu} + \hat{\sigma} u_p$. For the values considered above, $p = 0.001$, $\alpha = 0.2$ and $\varepsilon = 0.2$, we find that $c = 2.02n^{-1/2} - 0.065$, and therefore a size $(1.62n^{-1/2} - 0.052)(4.09 - \Delta)$ for the relative error results. Letting again $\Delta = 1.81$, this will wind up at the 15.1% obtained above for $n \approx 190$. Indeed, this is well below $n = 5000$. An intermediate example involving a larger parametric family is easily derived from the one presented after (5.4) in AKN (2002b). Since $\widehat{UL}$ there equals $\hat{\mu} + \hat{\sigma}\overline{K}_{\hat{\gamma}}^{-1}(p)$, a third parameter γ needs to be estimated, which thus leads for each given n to an increase of $\tau^2 = var\widehat{UL}$ in comparison to the normal case. Following (5.4) in AKN (2002b) it is calculated that $n\tau^2$ actually grows by a factor 4.32 when $p = 0.001$ and $\gamma = 0$. Consequently, the n required also grows by this factor: for $\Delta = 1.81$, the desired 15.1% is now reached for $n \approx 820$, a value which is indeed intermediate between the normal 190 and the nonparametric 5000.

It is nice to see from the example above that by using a suitable correction it is indeed possible to have acceptable behavior both while being in-control ('big errors occur rarely') as well as while being out-of-control ('detection power remains close to the uncorrected value'). Clearly, this cannot be achieved for customary n and p, for which r typically equals zero. Substantially larger values of n and/or p are needed than for normal or parametric charts, as the examples have demonstrated. But on the other hand, the pessimism according to which r itself should be really large, is fortunately also not realistic. Values like $r = 5$ seem to offer reasonable opportunities, and it matters little whether this is achieved by increasing n, or p, or both.

6 Summary

For convenience, we close with a straightforward summarized description of the proposed procedure. No comments are given on derivation, motivation, alternative approaches, etc., as all of these have been amply discussed above.

Suppose starting values are given for p, the signal probability while in-control, and n, the number of Phase I observations available for estimation. The basic chart simply uses as an upper limit $\widehat{UL}$ the order statistic $X_{(n-r)}$, where $r = [np]$ (here $[y]$ denotes the largest integer $\le y$). To keep the in-control behavior under control, for example require that $P(P > p(1+\varepsilon)) \le \alpha$. (An alternative is to focus on the ARL and use $P(1/P < (1-\varepsilon)/p)$, but this is equivalent: just replace ε in the first criterion by $\varepsilon/(1-\varepsilon)$). Choose values for ε and α. Let $Po(\tilde{\lambda}, k)$ denote the cumulative Poisson probability $P(\tilde{Y} \le k)$, where $\tilde{Y}$ is Poisson with parameter $\tilde{\lambda}$. If $Po(np(1+\varepsilon), r) \le \alpha$, the requirement is met and the chart can be used without further correction.

If not, somewhat larger values of ε and α could be settled for. But if this is not allowed or does not work either, a correction is applied. Find the smallest integer k for which $Po(np(1+\varepsilon), r-1-k) \le \alpha$. (Note that for $r = 0$ usually k will be 0 as well.) Let λ be such that $(1-\lambda)Po(np(1+\varepsilon), r-k-1) + \lambda Po(np(1+\varepsilon), r-k) = \alpha$. Then replace $\widehat{UL}$ by $X_{(n+k-r)}$ with probability λ and by $X_{(n+k+1-r)}$ with probability $(1-\lambda)$. The requirement is met and the behavior during in-control thus is acceptable.

Next check whether the price for the protection afforded by this correction, is acceptable during out-of-control as well. Take the normal case as a yardstick and suppose the mean has shifted by Δ standard deviations to the right. Let u_q denote the standard normal upper quantile for any q. Then the effect of the correction will lead to a relative change that approximately equals $(k+1-\lambda)\{1+u_q-\Delta\}/\{(r+1)(1+u_q)\}$, with $q=(r+1)/(n+1)$. (For very small r, like $r=0$, the approximation is not reliable, but then the relative errors involved are anyhow too large to be acceptable.) If this expression is suitably small, the corrected chart can indeed be applied.

If not, n and/or p need to be raised. In view of the simplicity of the relative error approximation considered, it is easy to figure out by some trial and error what kind of values will produce a result that is acceptable. Finally, the corrected chart based on the updated n and/or p, is the one to be applied.

References

Albers, W. and Kallenberg, W.C.M. (2000). Estimation in Shewhart control charts: effects and corrections. Technical Report 1559, University of Twente.

Albers, W. and Kallenberg, W.C.M. (2001). Are estimated control charts in control? Technical Report 1569, University of Twente.

Albers, W. and Kallenberg, W.C.M. (2000). Estimation in Shewhart control charts: effects and corrections. Technical Report 1559, University of Twente.

Albers, W. and Kallenberg, W.C.M. (2001). Are estimated control charts in control? Technical Report 1569, University of Twente.

Albers, W., Kallenberg, W.C.M. and Nurdiati, S. (2002a). Parametric control charts. Technical Report 1623, University of Twente.

Albers, W., Kallenberg, W.C.M. and Nurdiati, S. (2002b). Exceedance probabilities for parametric control charts. Technical Report 1650, University of Twente.

Albers, W., Kallenberg, W.C.M. and Nurdiati, S. (2002c). Data driven choice of control charts. In preparation, University of Twente.

Chakraborti, S. (2000). Run length, average run length and false alarm rate of Shewhart $\overline{X}$ chart: exact derivations by conditioning. *Commun. Statist. Simul. Comput.* **29**, 61-81.

Chakraborti, S., van der Laan, P. and Bakir, S.T. (2001). Nonparametric Statisical Process control: an overview and some results. *J. Qual. Technol.* **33**, 304-315.

Chen, G. (1997). The mean and standard deviation of the run length of $\overline{X}$ charts when control limits are estimated. *Statistica Sinica* **7**, 789-798.

Does, R.J.M.M. and Schriever, B.F. (1992). Variables control chart limits and tests for special causes. *Statist. Neerl.* **46**, 229-245.

Ghosh, B.K., Reynolds, M.R.Jr. and Hui,Y.V. (1981). Shewhart $\overline{X}$-charts with estimated process variance. *Commun. Statist. Theory Methods* **10**, 1797-1822.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Ass.* **58**, 13-29.

Ion, R. A., Does, R.J.M.M. and Klaassen, C.A.J.(2000). A comparison of Shewhart control charts based on normality, nonparametrics, and extreme-value theory. Report 00-8, University of Amsterdam

Quesenberry , C.P. (1993). The effect of the sample size on estimated limits for $\overline{X}$ and X control charts. *J. Qual. Technol.* **25**, 237-247.

Roes, C. (1999). Shewhart-type Charts in Statistical Process Control. Ph.D.-thesis, university of Amsterdam.

Willemain, T.R. and Runger, G.C. (1996). Designing control charts using an empirical reference distribution. *J. Qual. Technol.* **28**, 31-38.

Woodall, W.H. and Montgomery, D.C. (1999). Research issues and ideas in statistical process control. *J. Qual. Technol.* **31**, 376-386.