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. 1650

Exceedance probabilities for parametric control charts

W. Albers, W.C.M. Kallenberg and S. Nurdiati

October, 2002

ISSN 0169-2690

Exceedance probabilities for parametric control charts

Willem Albers*, Wilbert C.M. Kallenberg and Sri Nurdiati

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

Abstract Common control charts assume normality and known parameters. Quite often these assumptions are not valid and large relative errors result in the usual performance characteristics, such as the false alarm rate or the average run length. A fully nonparametric approach can form an attractive alternative but requires more Phase I observations than are usually available. Sufficiently large parametric families then provide realistic intermediate models. In this paper the performance of charts based on such families is considered. Exceedance probabilities of the resulting stochastic performance characteristics during in-control are studied. Corrections are derived to ensure that such probabilities stay within prescribed bounds. Attention is also devoted to the impact of the corrections for an out-of-control process. Simulations are presented both for illustration and to demonstrate that the approximations obtained are sufficiently accurate for use in practice.

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

2000 Mathematics Subject Classification: 62G15, 62P30

1 Introduction

Consider the following standard control chart procedure: the mean of a production process is monitored by means of a Shewhart chart. Each new value is compared with a given upper and lower limit and an out-of-control signal results if either of these two limits

*Corresponding author: tel: +31534893816; fax: +31534893069; e-mail: 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.

is exceeded. Usually normality of the distribution involved is taken for granted and it only remains to estimate its parameters using so-called Phase I observations. The outcomes are plugged into the expressions for the limits and the estimated chart is expected to behave as if it were based on known values. However, by now it is known that this unfortunately is too optimistic. In fact, quite a few authors have drawn attention to the fact that estimation can lead to serious errors. See e.g. Woodall and Montgomery (1999) (p. 379), Ghosh et al. (1981), Quesenberry (1993), Chen (1997), Chakraborti (2000), Albers and Kallenberg (2000) and Albers et al. (2002). To explain why this happens, note that the desired value p for the probability of getting a signal while in fact the in-control situation still persists, usually is extremely small. Values like $p = 0.001$ are customary and hence the estimation step will produce unexpectedly large relative errors for common sample sizes.

This situation can be repaired by applying suitable corrections to the estimated limits, as has been shown in Albers and Kallenberg (2000, 2001) (to be denoted for short as AK (2000, 2001) in the sequel). Two papers, rather than just one, resulted since (at least) two types of criterion can be used. The first is concerned with bias correction. Due to the estimation, p is replaced by a stochastic version P and typically $Eg(P)$ will differ considerably from $g(p)$ for the usual functions of interest: $g(p) = p$, $g(p) = 1/p$ (connected with the average run length (ARL)) and $g(p) = 1-(1-p)^k$ (corresponding to $P(RL \leq k)$). For each of these three choices, suitable corrections are derived in AK (2000). In this way the number n of Phase I observations which is required in order to arrive at an acceptably small bias, can be brought down from about 300 to about 40. This is gratifying, but we should realize that the bias criterion is rather mild, as it only corrects the average behavior of the chart over a long series of applications.

If instead we are more interested in what may happen for a single application, we should focus on the distribution of the random variable P around p, rather than just look at its average behavior (or that of $g(P)$). The fact that this distribution typically is asymptotically normal, clearly implies that $P(P > p)$ will tend to $\frac{1}{2}$ as n becomes large. The bias corrections mentioned above merely speed up this convergence and help to avoid that one has to start for smaller n with exceedance probabilities well above this $\frac{1}{2}$. Thus quite naturally, p will be close to the 50%-quantile of the distribution of P. Obviously, this can be felt to be much too liberal, leading to the desire to correct the estimated limits in such a way that p will be (close to) the upper α-quantile of the distribution of P for some sufficiently small α (like $\alpha = 0.1$ or $\alpha = 0.2$). In other words, it is desired that the distribution of P is shifted to the left such that $P(P > p) = \alpha$. A slightly relaxed version of this criterion is obtained by allowing the upper α-quantile in question to equal $p(1+\varepsilon)$, rather than p itself, for some small $\varepsilon \geq 0$ (like $\varepsilon = 0.1$). In this way, only in a fraction α of the applications, one is faced with a value of P which is really too large in the sense that it exceeds not only p, but even $p(1+\varepsilon)$. Using once more the functions g then finally leads to requiring $P(g(P) > g(p)(1+\varepsilon)) = \alpha$ for increasing g and $P(g(P) < g(p)(1-\varepsilon)) = \alpha$ for decreasing g. (Clearly, for $\varepsilon = 0$, g plays no role, as e.g. $P(P > p) = P(1/P < 1/p)$.)

Adaptations of this second type are obtained in AK (2001) for each of the three types of g under consideration. As this second criterion is more strict, it is not surprising that the corresponding corrections are of a larger order of magnitude than those required for the bias case. (In fact, the orders involved are $n^{-1/2}$ and n^{-1}, respectively.) Consequently, the impact on the out-of-control behavior will also be stronger when using this exceedance probability criterion. Note that this could mean a serious problem: if bringing the in-control behavior under control would result in a substantially lowered detection power once the process goes out-of-control, the price for the protection might be judged to be too high. Fortunately, however, the effects during the out-of-control stage will typically be sufficiently moderate. To understand why this is the case, note that during this stage the extremely small p from the in-control situation has increased to some p_1 which may still be small, but no longer extremely so. The change caused by applying a correction to the estimated control limits will be of the same order of magnitude for p and p_1. However, in terms of relative change, the impact on p_1 will be much more mild.

Hence according to the above, the practitioner can choose between a weak and a strong form of protection, at a low or moderate price, respectively, and it may seem that the problem has been satisfactorily solved. However, note that actually we have only repaired the effect of the unwarranted assumption that estimation effects are negligible. The other dubious assumption, according to which the distribution involved is simply normal, still stands. In fact, this second assumption is even more cumbersome. As n increases, eventually the estimation effects will decrease and the stochastic error (SE) will become negligible. The problem 'only' is that this takes much larger sample sizes than naive intuition suggests, thus making corrections typically indispensable. On the other hand, deviations from normality cause a model error (ME) that does not go away, no matter how large n is chosen. Again the extremity of the quantiles involved transforms this into a major problem: in the middle of the distribution, a normal approximation may work reasonably well, but in the far tail the relative errors caused can be unacceptably large. Incidentally, also the problems arising from assuming normality were pointed out before by several authors, see e.g. Chan et al. (1988), Pappanastos and Adams (1996) and Albers et al. (2002).

A logical next step thus is to acknowledge the possibility of a ME and to search for a compromise which keeps this ME within bounds without letting the SE explode. To see that the latter can easily happen, just go from the simple normal model to the other opposite, the fully nonparametric model. There, using the empirical quantiles, the ME indeed vanishes. But for typical configurations like $p = 0.001$ and $n = 100$, the stochastic error will clearly be overwhelming. Consequently, a family which is larger than the normal, but still parametric, is an attractive type of compromise to look for. Such families are studied in Albers, Kallenberg and Nurdiati (2002) (henceforth denoted by AKN (2002)) and there a specific choice, based on the so-called normal power family, is demonstrated to work well. For a broad class of distributions, the ME is controlled much better using this choice than under the simple normal model, where unacceptably large errors occur. The price for this improvement is of course a somewhat higher SE: here not only the mean and standard deviation need to be estimated (as in the normal model), but also the best fitting member of the available parametric family. But for such a parametric family as well, corrections can be derived to bring the SE under control. In analogy to the normal case, again two types of criterion suggest themselves: bounding the bias versus bounding the exceedance probability. The focus in AKN (2002) is on bias reduction and it is shown that with respect to this criterion accurate control limits can indeed be obtained.

In view of the above it is clear that it is of great interest to study parametric charts when the criterion is based on exceedance probabilities, and this will be the topic of the present paper. Obviously, we will amply benefit from our previous efforts and whenever possible we will refer to our earlier papers for additional motivation and details. From the four situations, normal chart with bias correction, normal chart with exceedance probability as criterion, parametric control chart with bias correction and parametric chart dealing with exceedance probability, the latter is the most ambitious one. We will now simultaneously try to control the occurrence of unpleasant values of P, as well as the occurrence of ME's which are unacceptably large. The question will be whether this still can be achieved for reasonable sample sizes without destroying the power of detection during out-of-control.

The paper is organized as follows. In section 2 the chart for the normal power family is introduced and its ME is briefly, and favorably, compared to that of the simple normal chart. Its SE on the other hand, and that of parametric families in general, is studied in section 3. In particular, corrections are derived which bring this SE under control with respect to exceedance probabilities. Such corrections are indeed larger than before on two counts: the model is larger and the criterion is more severe. Section 4 again specializes to the normal power family and presents a completely specific proposal for that situation. This proposal is subsequently investigated in a simulation study. It turns out to work quite well: without it, the exceedance probabilities are unacceptably large, whereas after correction the values obtained are indeed close to the desired α. The final section is devoted to studying the impact of the correction on the out-of-control behavior. As expected, it turns out that the effect can be substantial. Guidelines are given to check whether it is acceptable for the values of n, p and α at hand, or adaptations, such as a larger sample size, are called for. Again a simulation study is presented to support and illustrate the recommendations given.

2 A chart for the normal power family

Consider independent identically distributed random variables (rv's) $X_1, \ldots X_n, X_{n+1}$ from some distribution function (df) F. The first n of these rv's come from Phase I and form the basis for the estimation step; the last rv belongs to Phase II, the monitoring stage. Clearly, as all $(n+1)$ rv's come from the same F, we have the in-control situation as our starting point. For simplicity we shall concentrate on the one-sided case in which only an upper limit (UL) figures. The two-sided case can be dealt with in an analogous manner. Also, the monitoring variable X may be based on $m > 1$ observations, but we shall not go into that complication either and focus on the case of individual observations. If F is supposed to be $N(\mu, \sigma^2)$, the proper UL for a certain p simply equals $\mu + u_p\sigma$, where $u_p = \Phi^{-1}(1-p) = \overline{\Phi}^{-1}(p)$, in which Φ stands for the standard normal df and we use the convention that $\overline{H}$ denotes $1 - H$ for any df H. The fact that μ and σ are unknown requires these parameters to be replaced by customary estimators like $\hat{\mu} = \overline{X} = n^{-1}\Sigma X_i$

and $\hat{\sigma} = S = \{(n-1)^{-1}\Sigma(X_i - \overline{X})^2\}^{1/2}$, leading immediately to the choice

$$\widehat{UL} = \hat{\mu} + u_p\hat{\sigma}. \tag{2.1}$$

However, if normality is not taken for granted, a larger model should be selected. In AKN (2002) this issue is discussed extensively and for brevity here we just adopt the choice advocated in that paper, which means upgrading the normal family to the so-called normal power family. To be specific, instead of simply working under the model $X = \mu + \sigma Z$, in which Z has df Φ, we now suppose that $X = \mu + \sigma Z_\gamma$, where for some $\gamma > -1$,

$$Z_\gamma = c(\gamma)|Z|^{1+\gamma} sign(Z), \tag{2.2}$$

with normalizing constant $c(\gamma) = \{E|Z|^{2(1+\gamma)}\}^{-1/2} = \pi^{1/4}2^{-(1+\gamma)/2}\Gamma(\gamma + 3/2)^{-1/2}$. Clearly, the special case $\gamma = 0$ reproduces the normal case again. Let K_γ denote the df of Z_γ from (2.2), then it readily follows that $K_\gamma^{-1}(t) = c(\gamma)|\Phi^{-1}(t)|^{1+\gamma} sign(\Phi^{-1}(t))$, and thus u_p in (2.1) needs to be replaced by $c(\hat{\gamma})u_p^{1+\hat{\gamma}}$.

As concerns the choice for the required estimator $\hat{\gamma}$ of γ, we once more follow AKN (2002), where two possibilities are presented. The first is to start from the moment estimator

$$\hat{\gamma}_1^* = n^{-1}\sum_{i=1}^{n}\left\{\frac{X_i - \overline{X}}{S}\right\}^4, \tag{2.3}$$

which estimates $\gamma_1^* = EZ_\gamma^4 = h_1(\gamma)$, where $h_1(\gamma) = \pi^{1/2}\Gamma(2\gamma + 5/2)/\Gamma(\gamma + 3/2)^2$. Hence the estimator $\hat{\gamma}_1 = h_1^{-1}(\hat{\gamma}_1^*)$ can be used for γ itself. If instead of using all observations, we prefer to concentrate on the upper tail, we can proceed as follows. Let $X_{1:n}, \ldots, X_{n:n}$ denote the order statistics of $X_1, \ldots, X_n$ and write $[x]$ for the largest integer $\leq x$. Then, for some q and r with $0 < q < r < 1/2$, define

$$\hat{\gamma}_2^* = \frac{X_{[n+1-qn]:n} - \overline{X}}{X_{[n+1-rn]:n} - \overline{X}}, \tag{2.4}$$

which estimates $\gamma_2^* = \overline{K}_\gamma^{-1}(q)/\overline{K}_\gamma^{-1}(r) = (u_q/u_r)^{1+\gamma} = h_2(\gamma)$. Hence $\hat{\gamma}_2 = h_2^{-1}(\hat{\gamma}_2^*)$, with $h_2^{-1}(x) = -1 + \log(x)/\log(u_q/u_r)$, can be used as our second choice for estimating γ. Summarizing, based on the normal power model, $\widehat{UL}$ from (2.1) is replaced by

$$\widehat{UL}_i = \hat{\mu} + \hat{\sigma}c(\hat{\gamma}_i)u_p^{1+\hat{\gamma}_i}, \tag{2.5}$$

where $\hat{\gamma}_i$ for $i = 1, 2$ is given through (2.3) and (2.4).

The material above suffices to provide an explicit illustration of the ME and SE discussed in the introduction. For general $\widehat{UL}$, the stochastic counterpart P of p is given by $P(X_{n+1} > \widehat{UL}|(X_1, \ldots, X_n)) = \overline{F}(\widehat{UL})$, which can be decomposed into $p + ME + SE$, where

$$ME = \overline{F}(UL) - p, SE = \overline{F}(\widehat{UL}) - \overline{F}(UL), \tag{2.6}$$

and UL stands for the value to which $\widehat{UL}$ converges in probability under F. Now suppose for example (cf. AKN (2002)) that we have decided to simply use the normal limit from (2.1) and that in fact $X = \mu + \sigma Z_\gamma$, and therefore $\overline{F}(x) = \overline{K}_\gamma((x-\mu)/\sigma)$. In this situation ME from (2.6) reduces to

$$\overline{K}_\gamma(u_p) - p = \overline{\Phi}\left(\left\{\frac{u_p}{c(\gamma)}\right\}^{\frac{1}{1+\gamma}}\right) - p. \tag{2.7}$$

For e.g. $\gamma = 1$ and $p = 10^{-3}$, the expression in (2.7) produces $ME = 9.4 \times 10^{-3}$, showing that holding on to normality can indeed easily produce errors which are absolutely speaking small, but in a relative sense very large. Moreover, this situation is independent of n and thus does not improve if larger samples are used. On the other hand, if instead of using the simple normal $\widehat{UL}$, we would have started with $\widehat{UL}_i$ from (2.5), ME would equal 0. Of course, the normal power family is unduly favored in this particular comparison, as F belongs to it. But also for other F, it generally leads to a much better ME than in the normal family. Illustration of this point is supplied in AKN (2002). For example, for the standardized Student distribution with 6 degrees of freedom we obtain from Table 1 $ME = 3.6 \times 10^{-3}$ under the normal model, while in the normal power family Table 3 shows $ME = -\ 0.1 \times 10^{-3}$ when $\hat{\gamma}_1^*$ is used and Table 5 gives 2.1×10^{-3} when $\hat{\gamma}_2^*$ is used. For the Logistic distribution, the corresponding figures are 2.7×10^{-3}, 0.2×10^{-3} and 1.3×10^{-3}, respectively.

However, we also have to take the opposite side of the picture into account: the comparison of the various $SE's$. In this respect, clearly the normal power family is at a disadvantage. Not only μ and σ need to be estimated here (cf. (2.1)), but γ as well (cf. (2.5)). The question is whether the effects of the resulting increase of SE are sufficiently small to be outweighed by the above illustrated gain with respect to ME. In AKN (2002) this was shown to be true for the case where bias is the performance criterion. Adequate corrections were derived there, resulting in charts which combine small ME with small expected SE. But in the present paper we are interested in bounding exceedance probabilities rather than in bounding bias. As argued in the introduction, this type of criterion has a larger impact and in particular, it will necessitate larger corrections. In the next section we shall analyze how this works out.

3 Exceedance probabilities and corrections

In section 2 we have introduced estimated upper limits of the form $\widehat{UL} = \hat{\mu} + \hat{\sigma}\overline{K}_{\hat{\gamma}}^{-1}(p)$, with special emphasis on the normal power family as defined through (2.2). Note however, that the exposition goes through in general for families $\{K_\gamma\}$ with mean zero and variance one, containing some K_0 as a restricted model of special interest. In order to be able to correct the behavior of the corresponding chart, we now replace these $\widehat{UL}$ by

$$\widehat{UL}_c = \hat{\mu} + \hat{\sigma}\{\overline{K}_{\hat{\gamma}}^{-1}(p) + c_e\}, \tag{3.1}$$

which contains a correction term $c_e = c_e(\hat{\mu}, \hat{\sigma}, \hat{\gamma})$ to be determined in what follows. Obviously, by letting $c_e = 0$, we will always be able to reproduce the uncorrected charts, as $\widehat{UL}_0 = \widehat{UL}$. Let $UL = \mu + \sigma \overline{K}_{\gamma}^{-1}(p)$ again be the limit in probability of (3.1). Note that it is tacitly assumed here that c_e vanishes in the limit, which is in line with the idea that estimation effects do become negligible as $n \to \infty$.

As remarked in the previous section, P will equal $\overline{F}(\widehat{UL}_c)$. The question now is how likely it is that $g(P)$, for e.g. the three choices of g mentioned in the introduction, differs too much from its corresponding limit value $g(\overline{F}(UL))$. To be more precise, we introduce the relative error

$$W_c = \frac{g(\overline{F}(\widehat{UL}_c))}{g(\overline{F}(UL))} - 1, \tag{3.2}$$

with $W = W_0$ corresponding to the uncorrected case. Once more note that we are only dealing with SE here. Note that $ME = g(\overline{F}(UL)) - g(p)$, which hopefully has been made small by using a larger family, remains a given quantity, no matter how large a sample size we choose. For increasing g (like $g(p) = p$ or $g(p) = 1 - (1-p)^k$), we impose the following exceedance probability criterion: for certain small non-negative ε and small positive α,

$$P(W_c > \varepsilon) \leq \alpha. \tag{3.3}$$

For decreasing g (like $g(p) = 1/p$), instead consider $P(W_c < -\varepsilon)$ in (3.3). In what follows we shall, unless explicitly stated otherwise, always assume that g is increasing.

Note that (3.3) translates into $P(\widehat{UL}_c < b) \leq \alpha$, where $b = \overline{F}^{-1}(g^{-1}(\{g(\overline{F}(UL))(1 + \varepsilon)\}))$. (For decreasing g, replace the factor $1+\varepsilon$ by $1-\varepsilon$.) In the normal case, no $\hat{\gamma}$ occurs as $\gamma \equiv 0$, and (3.1) thus reduces to $\widehat{UL}_c = \hat{\mu} + \hat{\sigma}(u_p + c_e)$. Hence, using noncentral t-distributions, for this special case the exact value of c_e can be computed which produces equality in (3.3) (cf. AK (2001), (5) and (6)). Clearly, in general – or even within the normal power family – this is no longer feasible. However, this is less of a drawback than it may seem. Even in AK (2001), the emphasis is not on such exact outcomes, but rather on approximations based on asymptotics. The latter are simple and transparent and as such reveal how the estimation effects depend on the underlying parameters n, p, α and ε and the functions g, something which remains obscure when looking at mere numerical results. Hence we shall certainly apply asymptotic arguments here as well.

First we simplify matters somewhat by showing that without loss of generality we may work under $\mu = 0$ and $\sigma = 1$. Actually, let F_0 be the df of $(X_{n+1} - \mu)/\sigma$, then $b = \mu + \sigma b_0$ with $b_0 = \overline{F}_0^{-1}(g^{-1}(\{g(\overline{F}_0(\overline{K}_{\gamma}^{-1}(p)))(1+\varepsilon)\}))$. From (3.1) it follows that $P(\widehat{UL}_c < b) = P((\hat{\mu} - \mu)/\sigma + (\hat{\sigma}/\sigma)\{\overline{K}_{\hat{\gamma}}^{-1}(p) + c_e\} < b_0)$. This implies that indeed we may assume all X_i to have df F_0 and subsequently consider the relation $P(W_c > \varepsilon) = P(V + \hat{\sigma} c_e < b_0 - \overline{K}_{\gamma}^{-1}(p))$, where

$$V = \hat{\mu} + \hat{\sigma} \overline{K}_{\hat{\gamma}}^{-1}(p) - \overline{K}_{\gamma}^{-1}(p) \tag{3.4}$$

(cf. AKN (2002), (5.1)). Now typically EV and EV^2 will be of order n^{-1}, $E|V|^k$ will be $o(n^{-1})$ for $k \geq 2$ and V will be asymptotically normal. For example, this holds for the normal power model considered, with $\hat{\mu} = \overline{X}, \hat{\sigma} = S$ and $\hat{\gamma}_i$ as given through (2.3) or (2.4). These facts immediately yield that equality to first order will be achieved in (3.3) for

$$c_e = (EV^2)^{1/2} u_\alpha + b_0 - \overline{K}_\gamma^{-1}(p). \tag{3.5}$$

A further approximation step is obtained by expanding b_0 with respect to ε, which will often be small. Usually one step will suffice, leading to $b_0 - \overline{K}_\gamma^{-1}(p) \approx -\varepsilon g(\overline{F}_0(\overline{K}_\gamma^{-1}(p)))$ $/\{g'(\overline{F}_0(\overline{K}_\gamma^{-1}(p))) f_0(\overline{K}_\gamma^{-1}(p))\}$. To simplify this a bit further, let $h(p) = g(p)/\{pg'(p)\}$. For $g(p) = p$ we get $h(p) = 1$, for $g(p) = 1/p$ we find $h(p) = -1$, while for $g(p) = 1 - (1-p)^k$ we obtain $h(p) = (1-p)\{(1-p)^{-k} - 1\}/(kp) \approx (e^{kp} - 1)/(kp)$. Usually $\delta = kp$ will be small and in the last case $h(p) \approx 1 + \delta/2$. As a result we obtain the proposal

$$c_e = (EV^2)^{1/2} u_\alpha - \varepsilon \lambda \overline{F}_0(\overline{K}_\gamma^{-1}(p))/f_0(\overline{K}_\gamma^{-1}(p)), \tag{3.6}$$

where $\lambda = 1$ for the first two choices of g (in the second case $h(p) = -1$, but we also deal with $1 - \varepsilon$ there, rather than with $1 + \varepsilon$ because g is decreasing) and $\lambda \approx 1 + \delta/2$ for the third.

Note that an approximation like (3.6) is quite helpful in analyzing the impact of the various ingredients involved. To begin with, as $(EV^2)^{1/2}$ is of order $n^{-1/2}$, the corrections in the exceedance case are indeed seen to be of a larger order of magnitude than those in the bias case. The latter involve (cf. AKN (2002), section 5) EV and EV^2, which are both of order n^{-1}. Moreover, $(EV^2)^{1/2}$, and thus c_e, will be larger when more parameters are used. In the simple normal case, (3.4) reduces to $\hat{\mu} + (\hat{\sigma} - 1)u_p$ and $EV^2 \approx (u_p^2 + 2)/(2n)$ (cf. AK (2001), (9)). As concerns α, it is immediate from (3.6) that (and how) c_e increases as (3.3) is made more strict by using a smaller α. Likewise, increase of ε makes (3.3) less strict and this reflects itself in (3.6) in a decrease of c_e. Next, the effect of either looking at P itself, or at the ARL, or at $P(RL \leq k)$, is represented concisely in the factor λ. Finally, if F is indeed contained in the model used, we simply have $F_0 = K_\gamma$ and a factor $p/k_\gamma(\overline{K}_\gamma^{-1}(p))$, with k_γ the density of K_γ, results in the second term from (3.6). In the normal case, this approximately boils down to u_p^{-1} and consequently (3.6) reduces to the approximation (cf. AK (2001), (11)),

$$c_e = \left\{ \frac{u_p^2 + 2}{2n} \right\}^{1/2} u_\alpha - \frac{\varepsilon \lambda}{u_p}. \tag{3.7}$$

Before we can apply (3.5) or (3.6), it remains to estimate the unknown parts. As F_0 is supposed to be close to K_γ, we use $\overline{K}_{\hat{\gamma}}$ for $\overline{F}_0$, $k_{\hat{\gamma}}$ for f_0 and also evaluate EV^2 under $K_{\hat{\gamma}}$. Denoting the latter by $\widehat{EV^2}$, we arrive from (3.5) at

$$c_e = c_e(\hat{\gamma}) = (\widehat{EV^2})^{1/2} u_\alpha + \overline{K}_{\hat{\gamma}}^{-1}(g^{-1}(\{g(p)(1+\varepsilon)\})) - \overline{K}_{\hat{\gamma}}^{-1}(p), \tag{3.8}$$

which reduces through (3.6) to

$$c_e = c_e(\hat{\gamma}) = (\widehat{EV^2})^{1/2} u_\alpha - \frac{\varepsilon\lambda p}{k_{\hat{\gamma}}(\overline{K}_{\hat{\gamma}}^{-1}(p))}. \tag{3.9}$$

Hence with c_e from (3.8), the estimated upper limit from (3.1) will now produce approximate equality in (3.3) under the model $K_\gamma((x-\mu)/\sigma)$ (and its vicinity). Actual application in case of the special normal model already requires (see (3.7)) specification of p, ε, α and g, as well as evaluation through $(X_1, \ldots, X_n)$ of $\hat{\mu}$ and $\hat{\sigma}$. Under a more general model in addition K_γ should obviously be defined, as well as an estimator $\hat{\gamma}$ for γ, while also $\widehat{EV^2}$ has to be evaluated as a function of $\hat{\gamma}$. Especially this last step may require considerable effort. Fortunately, for the normal power model this has already been done in AKN (2002) while deriving the bias correction term, and we can readily use these results here. In the next section we shall investigate the performance of the thus obtained normal power family chart with exceedance probability correction.

4 The corrected normal power family chart

In section 3 we have uncovered the general structure of the desired corrections; here we shall demonstrate how (3.8) works out in practice for our prototype example, the normal power family. As in this case $K_\gamma^{-1}(t) = c(\gamma)|\Phi^{-1}(t)|^{1+\gamma} sign(\Phi^{-1}(t))$ (cf. (2.2)), we immediately obtain that

$$c_e = (\widehat{EV^2})^{1/2} u_\alpha + c(\hat{\gamma})\{u_{\tilde{p}}^{1+\hat{\gamma}} - u_p^{1+\hat{\gamma}}\}, \tag{4.1}$$

where $\tilde{p} = g^{-1}(\{g(p)(1+\varepsilon)\})$. As moreover $\{1/k_\gamma(\overline{K}_\gamma^{-1}(1-t))\} = 1/k_\gamma(K_\gamma^{-1}(t)) = \{K_\gamma^{-1}(t)\}' = (1+\gamma)c(\gamma)\Phi^{-1}(t)^\gamma / \phi(\Phi^{-1}(t))$ for $t > 1/2$, the expanded version transforms into $c_e = (\widehat{EV^2})^{1/2} u_\alpha - \varepsilon\lambda p(1+\hat{\gamma})c(\hat{\gamma})u_p^{\hat{\gamma}}/\phi(u_p)$. According to (3.1), the expression in (4.1) should be used in $\widehat{UL}_c = \hat{\mu} + \hat{\sigma}\{c(\hat{\gamma})u_p^{1+\hat{\gamma}} + c_e\}$ (also cf. (2.5)). The estimators to be used are again either $\hat{\gamma}_1$, defined through (2.3), or $\hat{\gamma}_2$, defined through (2.4). Indeed the main obstacle is $\widehat{EV^2}$, which has to be expressed in terms of $\hat{\gamma}$ as well. In the Appendix of AKN (2002), this laborious task has been carried out already. Suitable approximations are applied there to replace the extremely complicated expressions that arise. To be precise, in (A.9) and (A.15) from that paper the result can be found for $\hat{\gamma}_1$ and $\hat{\gamma}_2$, respectively.

Nevertheless, as those two results show, the expressions involved are still not that easy to use, among others because they contain many components which need separate definitions in their turn. Consequently, for the case where a specific choice of $\hat{\gamma}_2$ is used – which actually is the final recommendation in AKN (2002) – in section 7 of that paper a further approximation step is proposed, through which this nested structure is overcome. Proceeding here in a similar fashion we obtain that $\{\widehat{EV^2}\}^{1/2} \approx n^{-1/2} A(\hat{\gamma}, u_p)$, where

$$A(\hat{\gamma}, u_p) = -4.00 - 12.54\hat{\gamma} - 10.02\hat{\gamma}^2 + 2.91u_p + 6.47\hat{\gamma}u_p + 4.42\hat{\gamma}^2 u_p, \tag{4.2}$$

with (cf. (2.4))

$$\hat{\gamma} = 1.1218 \log\Big\{\frac{X_{[0.95n+1]:n} - \overline{X}}{X_{0.75n+1]:n} - \overline{X}}\Big\} - 1. \tag{4.3}$$

Note that the resulting control limit

$$\hat{\mu} + \hat{\sigma}\{c(\hat{\gamma})u_p^{1+\hat{\gamma}} + c_e\}, \tag{4.4}$$

with c_e as given through (4.1)-(4.3), now is completely explicit and can be applied in a straightforward manner (cf. its counterpart for the bias case as given by (7.1) and (7.2) in AKN (2002)).

To see how the proposal in (4.4) actually works out, we shall next perform a simulation study along the lines of AKN (2002). To avoid repetition, however, we shall not be as extensive as in that paper. For example, we shall not dwell again on the improvement achieved concerning ME by using a parametric rather than a normal chart. This gain has been amply demonstrated in AKN (2002) and the situation in that respect is exactly the same here. Hence we shall neither consider the normal chart at any point, nor report the ME involved. What matters is whether the present corrections work well for controlling SE under the present criterion. We shall concentrate on g(p) = p; from the derivations given it is evident that completely similar results will hold for the other two choices of g. Moreover, we will always let $p = 0.001$ and use 10,000 repetitions in the simulations.

Sample sizes n involved will range from 250 to 2000. Note that these values are considerably higher than those in AKN (2002), where the values 100, 250 and 500 are considered, with the emphasison on 100. This reflects the fact that in the present situation we are dealing with corrections of order $n^{-1/2}$ rather than of order n^{-1}, thus requiring larger values of n. As concerns the constants α and ε used in setting our criterion (cf. 3.3)), we shall use $\alpha = 0.2$, and either $\varepsilon = 0$ or $\varepsilon = 0.1$.

The underlying distributions will be the same as in AKN (2002). First of all include the normal df Φ, corresponding to the normal family. Then add K_γ as defined through (2.2), for the values γ = –0.5, –0.25, 0.25, 0.75 and 1 (γ= 0 is already covered by Φ), thus representing the normal power family. Subsequently, it is only fair to add a number of cases outside either family. To begin with, include T, the Student df with 6 degrees of freedom and standardized to unit variance. Next add the random mixture $RM = \frac{1}{2}\Phi + \frac{1}{2}T$ and the deterministic mixture DM given by $DM^{-1} = c^*(\Phi^{-1} + T^{-1})$, with c^* a constant to ensure unit variance. In addition, consider Tukey's λ-family, based on a rv $c(\lambda)\{U^\lambda - (1-U)^\lambda\}$, with U a uniform rv on (0,1) and $c(\lambda)$ once more a constant to achieve unit variance. Include the corresponding df's for λ = –0.1, 0 (which corresponds to the standardized logistic df) and 0.14 (which is very close to the standard normal (outside the tails!)). Finally, take the following orthonormal family: for $k = 1, 2, \ldots$ and $j = 1, \ldots, k$, let γ_j be a coefficient, π_j be the j^{th} Legendre polynomial and consider the density $f(x, \gamma_1, \ldots, \gamma_k)$ proportional to $\exp\{\Sigma\gamma_j\pi_j(x)\}$. If Y is a rv with this density f, then consider $\Phi^{-1}(Y)$ and standardize that rv to have zero mean and unit variance. Now include the corresponding df for $k = 3, \gamma_1 = \gamma_2 = -0.1$ and $\gamma_3 = 0.1$.

The simulation results are presented in Table 4.1. As can be seen from this table, the correction works quite well. To be more specific, first consider the normal power family when no correction is used. For $\varepsilon = 0$, we are then simply looking at $P(P > p)$, which is seen to stabilize around $\frac{1}{2}$. Hence indeed, as remarked in the introduction, p turns out to be close to the 50%-quantile of the distribution of P. (Note that for distributions outside the normal power family p should be replaced by $\overline{F}(UL)$). Increasing ε from 0 to 0.1 should help in this respect: as $n \to \infty$, $P(P > p(1+\varepsilon)) \to 0$ for $\varepsilon > 0$. But from Table 4.1 we see that apparently this convergence is quite slow. Even for n as large as 2000, the exceedance probabilities are still larger than 35%. Hence corrections are certainly in order if values of α well below $\frac{1}{2}$ are desired. From Table 4.1 it is evident that applying such a correction for $\alpha = 0.2$ indeed brings the exceedance probabilities down to values which are close to this desired 20%, both for $\varepsilon = 0$ and $\varepsilon = 0.1$. For $n = 250$, the fluctuations may still be considered to be a bit large, but from $n = 500$ on, the result seems quite satisfactory for practical purposes. Also observe that, although the correction terms are based on the normal power family, they work also rather well outside this family.

Table 4.1 *Simulated exceedance probabilities (in %) without* $(P(W_0 > \varepsilon))$ *and with* $(P(W_c > \varepsilon))$ *correction, using (cf. (3.3))* $\varepsilon = 0$ *or* 0.1 *and* $\alpha = 0.2$. *The first percentage in each cell corresponds to* $\varepsilon = 0$; *the second to* $\varepsilon = 0.1$.

F_0	$P(W_0 > \varepsilon)$								$P(W_c > \varepsilon)$							
	$n = 250$		$n = 500$		$n = 1000$		$n = 2000$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
Φ	52	49	51	46	51	44	51	41	24	24	23	23	22	22	22	22
$K_{-0.5}$	50	47	50	46	49	44	50	42	20	20	20	20	19	19	20	20
$K_{-0.25}$	51	48	50	45	50	44	50	41	24	24	21	21	22	22	22	22
$K_{0.25}$	53	49	50	44	51	44	51	40	25	25	23	23	22	22	22	22
$K_{0.50}$	54	50	52	46	52	44	50	39	26	26	24	23	22	22	21	21
$K_{0.75}$	54	50	52	47	52	44	51	39	26	26	24	23	22	22	21	21
K_1	55	51	52	47	52	44	52	41	27	27	24	24	23	23	22	22
T	54	47	52	43	50	39	51	34	29	27	26	23	26	21	25	19
RM	53	47	52	43	50	39	51	34	26	24	25	22	23	20	23	18
DM	53	47	51	43	51	40	50	35	26	24	24	21	24	21	24	19
$TU(-0.1)$	54	48	51	43	51	39	50	34	30	27	28	24	27	22	26	19
$TU(0)$	53	47	52	44	51	41	50	36	27	25	26	24	24	22	24	20
$TU(0.14)$	52	49	51	46	50	45	50	42	25	25	23	23	23	23	22	22
O	52	49	52	48	51	45	50	43	23	24	23	24	22	23	21	23

5 The out-of-control situation

Here we shall study the impact of the adaptations from the previous sections on the out-of-control behavior of the chart. Avoiding (apart from a small probability) stopping unexpectedly early during the in-control period is of course desirable, but this should not be achieved by typically stopping much later once the process has gone out of control. Thus let X_{n+1} now come from a shifted df $F_0(x-\Delta)$, where Δ is such that $\tilde{p} = \overline{F}_0(UL-\Delta)$ is no longer extremely small, like p. (For simplicity, and without loss of generality, we again let $\mu = 0$ and $\sigma = 1$ and thus work under the standardized df F_0.) Consequently, the relative errors caused by the replacement of this $\tilde{p}$ by its stochastic counterpart P, which in view of (3.1) now equals $\overline{F}_0(\widehat{UL}_c - \Delta)$, will be much smaller than those during the in-control situation (also cf. tables 10 and 11 from AK (2000)). Hence during out of control there seems to be no need to use exceedance probabilities again, as a more simple first order expectation approach will already suffice to exhibit the resulting behavior of the chart. To be specific, let E_Δ denote expectation under $F_0(x-\Delta)$ and introduce

$$E(\Delta, c_e) = E_\Delta g(\overline{F}_0(\widehat{UL}_c - \Delta)), RC = |E(\Delta, c_e)/E(\Delta, 0) - 1|. \tag{5.1}$$

Clearly, $E(\Delta, c_e)$ and $E(\Delta, 0)$ stand for $E_\Delta g(P)$ with and without correction, respectively. Moreover, RC expresses the relative cost incurred by having to use the correction c_e. A simple example explains its meaning: take $g(p) = 1/p$, let $p = 0.001$ and suppose that Δ is such that $\tilde{p} = 0.05$. Then $E(\Delta, 0)$ will be close to 20, and a value for RC of 20% means that using the correction c_e has pushed up the AF during out of control by 20% to about 24. Note that of course the notion "cost" is somewhat virtual and should therefore be interpreted with care. In fact, there are three approaches to be distinguished. In the first, everything is known, no estimation is needed, and both $g(p)$ during in control and $g(\tilde{p})$ during out of control are achieved effortlessly. The only drawback is that this situation rarely occurs, so one typically has to settle for one of the two remaining ones: to correct or not to correct. In the latter, $g(\tilde{p})$ indeed is achieved (apart from a usually acceptably small relative error), but, as we saw, the price in terms of exceedance probabilities is unacceptably large with respect to $g(p)$, caused by the very small values of p that are typically used. Hence the remaining candidate uses the correction, thus repairing the damage during in control (cf. Table 4.1), but at a price under out of control as expressed by RC from (5.1). To obtain an impression of what can be expected during out of control, we have repeated the simulations from section 4 for the same choices of p, n, g, α, ε and F_0 as used there. For each F_0 we have selected two values of Δ such that reasonable values of $\tilde{p}$ result (i.e. $\tilde{p}$ considerably larger than $p = 0.001$). Usually, $\Delta = 2$ and $\Delta = 3$ will do, but as γ moves away from 0, the values $\Delta = 1$ or $\Delta = 4$ can become more appropriate. In Table 5.1 the results have been collected: presented are the simulated average P under out of control when no correction is used, which are close to $\tilde{p}$. In addition, the relative costs RC are given in percentages, using $\varepsilon = 0$ and $\varepsilon = 0.1$ in the correction c_e, respectively.

Table 5.1 *Simulated values of $E(\Delta, 0) = E_{\Delta}P$ (see (5.1)), together with the relative costs due to correction RC (in %), using $\varepsilon = 0$ and $\varepsilon = 0.1$, respectively. Throughout are used: $p = 0.001$, $g(p) = p$ and $\alpha = 0.2$.*

F_0	Δ	$n = 250$			$n = 500$			$n = 1000$			$n = 2000$		
Φ	2	0.15	37	34	0.14	27	24	0.14	20	16	0.14	14	10
	3	0.47	23	21	0.46	16	14	0.46	11	9	0.46	8	6
$K_{-0.5}$	1	0.23	24	23	0.23	17	15	0.23	12	10	0.23	8	7
	2	0.50	2	2	0.54	1	1	0.50	0	0	0.50	0	0
$K_{-0.25}$	1	0.066	42	38	0.061	31	28	0.060	23	19	0.059	16	12
	2	0.35	19	17	0.35	13	12	0.35	9	7	0.35	7	5
$K_{0.25}$	2	0.061	47	43	0.054	36	31	0.052	27	22	0.050	20	14
	3	0.24	40	37	0.22	30	26	0.22	22	18	0.21	16	11
$K_{0.50}$	3	0.11	50	45	0.09	39	34	0.09	29	23	0.08	21	15
	4	0.39	47	43	0.36	38	33	0.34	30	24	0.32	22	16
$K_{0.75}$	3	0.054	55	50	0.044	43	38	0.040	33	26	0.038	25	17
	4	0.21	57	52	0.16	45	39	0.14	34	27	0.13	25	18
K_1	3	0.031	57	53	0.024	46	40	0.021	35	28	0.020	26	19
	4	0.11	60	56	0.08	48	42	0.06	37	30	0.06	28	20
T	2	0.092	40	37	0.081	31	27	0.077	23	19	0.075	17	12
	3	0.36	34	31	0.34	26	22	0.34	19	15	0.34	13	9
RM	2	0.12	40	36	0.11	30	26	0.10	22	18	0.10	16	12
	3	0.41	29	26	0.40	21	18	0.40	15	12	0.40	11	7
DM	2	0.12	39	36	0.11	30	26	0.11	22	18	0.10	16	12
	3	0.42	28	26	0.41	21	18	0.41	15	12	0.41	10	7
$TU(-0.1)$	2	0.072	41	37	0.062	31	27	0.057	24	19	0.055	17	12
	3	0.30	39	35	0.27	30	26	0.27	23	18	0.26	17	12
$TU(0)$	2	0.098	41	37	0.089	31	27	0.085	23	18	0.082	17	12
	3	0.37	33	30	0.36	24	21	0.35	18	14	0.35	13	9
$TU(0.14)$	2	0.15	37	34	0.14	27	24	0.14	20	17	0.14	14	11
	3	0.46	23	21	0.46	17	14	0.46	12	10	0.46	8	6
O	2	0.071	49	45	0.065	37	32	0.062	27	22	0.061	20	14
	3	0.28	36	32	0.27	26	22	0.27	19	15	0.26	14	10

From Table 5.1 it is evident that the RC values are decreasing nicely as n becomes larger (cf. the much slower decrease of the exceedance probabilities for the case without correction from Table 4.1). Moreover, the percentages occurring can be considerable, and hence the use of such larger values of n can indeed be felt to be necessary. Certainly for $n = 250$, the values are quite large and only from $n = 1000$ on small percentages start to prevail. Note that becoming more liberal in (3.3) by increasing ε from 0 to 0.1, does indeed help, but not terribly much: the percentages drop, but not dramatically so. Another comment is that things go relatively well for the choices of F_0 outside the normal power family. The most unfavorable cases occur within this family, for large positive values of γ. All in all, the results in Table 5.1 tend to confirm the expectations expressed in the introduction. The present type of correction is the most ambitious one of the four types considered (correcting bias or exceedance probabilities, assuming normality or not) and the impact on the out of control behavior is no longer negligible. Applying a correction of this type may very well be a good idea (in principle even the best among the available four), but it should not be applied automatically. Some thought should be given to points such as what size of RC is still acceptable, whether the n at hand is sufficiently large (or can be made so) to realize that size, whether increasing α might be an option, etc.

In view of the discussion above, it may be useful to have a better insight into the behavior of e.g. RC as a function of the rather large number of underlying parameters, such as p, n, α, ε, γ and Δ. To this end we shall consider some further approximations to the quantities from (5.1). Note that $E(\Delta, c_e)$ can be approximated by $g(\overline{F}_0(\overline{K}_\gamma^{-1}(p)+c_e-\Delta))$ and consequently RC to first order equals $c_e f_0(\overline{K}_\gamma^{-1}(p)-\Delta)(g'/g)(\overline{F}_0(\overline{K}_\gamma^{-1}(p)-\Delta))$. Using similar steps for the three choices of g involved as those leading to (3.6), it follows that RC approximately equals

$$c_e f_0(\overline{K}_\gamma^{-1}(p)-\Delta)/\{\lambda\overline{F}_0(\overline{K}_\gamma^{-1}(p)-\Delta)\}, \tag{5.2}$$

where once more $\lambda = 1$ for either $g(p) = p$ or $g(p) = 1/p$, while $\lambda = 1 + (kp)/2$ for $g(p) = 1-(1-p)^k$. To study the behavior of the expression in (5.2), we may specialize to the family $\{K_\gamma\}$. Note that F_0 for distributions outside the family $\{K_\gamma\}$ (under consideration in this paper) is well approximated by a suitably chosen member K_γ in this family. Straightforward calculation shows that $k_\gamma(\overline{K}_\gamma^{-1}(p)-\Delta)/\overline{K}_\gamma(\overline{K}_\gamma^{-1}(p)-\Delta)$ equals, for $0 < \tilde{p} \le \frac{1}{2}$,

$$u_{\tilde{p}}^{-\gamma} k(u_{\tilde{p}})/\{(1+\gamma)c(\gamma)\}, \tag{5.3}$$

where again $\tilde{p} = \overline{K}_\gamma(\overline{K}_\gamma^{-1}(p)-\Delta)$ and $k(x) = \phi(x)/\overline{\Phi}(x)$. It is well-known that $k(x) \approx x$ for x large. But note that for smaller x, like e.g. $0 < x \le 3.09 = u_{0.001}$, the function k can be approximated quite adequately by $4(1+x)/5$ (see also (17) and (18) from AK (2001)). Combination of (3.9) and (4.2), together with (5.2) and (5.3) then leads for $\tilde{p} < \frac{1}{2}$ to the following first order estimate for RC:

$$\widehat{RC} \approx \frac{4(1+u_{\tilde{p}})}{5\lambda(1+\hat{\gamma})c(\hat{\gamma})u_{\tilde{p}}^{\hat{\gamma}}} A(\hat{\gamma}, u_p) n^{-1/2} u_\alpha - \varepsilon \frac{(1+u_{\tilde{p}})u_p^{\hat{\gamma}}}{(1+u_p)u_{\tilde{p}}^{\hat{\gamma}}}. \tag{5.4}$$

As an example, select the values $p = 0.001$, $\lambda = 1$ and $\alpha = 0.2$ which are used throughout Table 5.1 and consider the case $\hat{\gamma} = 0$, corresponding to $F_0 = \Phi$. Then (5.4) reduces to $\widehat{RC} = 3.36(4.09-\Delta)n^{-1/2} - \varepsilon(1-\Delta/4.09)$. For $\Delta = 2$ the resulting $7.03n^{-1/2} - 0.51\varepsilon$ produces for $\varepsilon = 0$ the percentages 44, 31, 22 and 16 for $n = 250, 500, 1000$ and 2000, respectively; for $\varepsilon = 0.1$, these percentages should be lowered by 5. For $\Delta = 3$ we have $3.66n^{-1/2} - 0.27\varepsilon$ and the resulting values are 23, 16, 12 and 8 for $\varepsilon = 0$, to be lowered by 2 or 3 for $\varepsilon = 0.1$. Comparison to the simulated values from Table 5.1 for $F_0 = \Phi$ shows a nice agreement, especially for $\varepsilon = 0.1$. Incidentally, for the case $\hat{\gamma} = 0$ considered here, another interesting comparison can be made to the situation where normality is assumed to begin with, and thus estimation of γ is not needed. Then, according to (3.7), we deal with $\{(u_p^2+2)/2\}^{1/2}$, rather than with $A(0,u_p)$. For $p=0.001$, the former equals 2.40, while the latter equals 4.99. Hence the fact that the additional parameter γ needs to be estimated calls in this example for a value of n which is $(4.99/2.40)^2 = 4.32$ times as high to reach the same precision. This illustrates that the impact of going from normality to a more general parametric family is indeed far from negligible.

As intended, the result in (5.4) helps to shed some light on the behavior during out of control. To begin with, note that the influence of ε is indeed seen to be quite limited, as its coefficient in (5.4) is typically smaller than 1. Hence from now on we concentrate on the case where $\varepsilon = 0$, i.e. where only the first term in (5.4) is relevant. Clearly, this term decreases in both n and α. The dependence on n and α is rather simple, since they are not mixed up with other parameters Furthermore, we can break up the remaining factor into the parts $4(1+u_{\tilde{p}})/\{5(1+\hat{\gamma})c(\hat{\gamma})u_{\tilde{p}}^{\hat{\gamma}}\}$ and $A(\hat{\gamma},u_p)$. Let $\hat{\gamma}$ be fixed, then we observe that it also decreases in p, as $A(\hat{\gamma},u_p)$ increases linearly in u_p (cf. (4.2). The situation with respect to $\tilde{p}$ is slightly more complicated: it decreases in $\tilde{p}$ $(0<\tilde{p}<\frac{1}{2})$ only as long as $u_{\tilde{p}} > \gamma/(1-\gamma)$. Finally, the behavior with respect to γ is more variable. On the one hand, $A(\gamma,u_p)$ increases quadratically in γ (for $p=0.001$, we for example have that it behaves as $(7\gamma+9)^2/16)$, but the effect of $u_{\tilde{p}}^{\gamma}$ will obviously depend on whether $u_{\tilde{p}}$ is larger than or smaller than 1. As an example, it can be checked numerically for $p=0.001$ and $\alpha=0.2$, the coefficient of $n^{-1/2}$ in (5.4) will not exceed 9.43 for γ and $\tilde{p}$ such that $-0.25 \le \gamma \le 0.3$ and $0.1 \le \tilde{p} \le 0.3$.

A final remark is that the first term in (5.4) can be used in a straightforward manner to derive a lower bound for the n required. Let β be a small positive number, like 0.2 or 0.3, and suppose we want RC to be at most β, then it readily follows that we should let

$$n \ge \left\{\frac{4(1+u_{\tilde{p}})A(\hat{\gamma},u_p)u_\alpha}{5(1+\hat{\gamma})c(\hat{\gamma})u_{\tilde{p}}^{\hat{\gamma}}\beta}\right\}^2. \tag{5.5}$$

Continuing the numerical example just mentioned, it follows that for this case (5.5) boils down to $n \ge \{9.43/\beta\}^2$, which equals 2222 and 988 for $\beta=0.2$ and $\beta=0.3$, respectively. (Incidentally, if these computations are based, without further approximation steps, on $RC \approx g(\overline{F}_0(\overline{K}_\gamma^{-1}(p)+c_e-\Delta))/g(\overline{F}_0(\overline{K}_\gamma^{-1}(p)-\Delta))-1$, the corresponding results are 1701 and 678, respectively.) Note that the values of n obtained are in line with the impression already created by Table 5.1: the correction c_e works quite well, but considerable sample sizes are needed to avoid effects during out of control which might be considered too strong.

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

Chan, L.K., Hapuarachchi, K.P. and Macpherson, B.D. (1988). Robustness of $\overline{X}$ and R charts. *IEEE Trans. Reliability* **37**, 117-123.

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.

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.

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.

Pappanastos, E.A. and Adams, B.M. (1996). Alternative designs of the Hodges-Lehmann control chart. *J. Qual. Technol.* **28**, 213-223.

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.

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