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 general 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 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

*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.

expected to behave as if it were based on known values. However, by now it is known that this unfortunately is too optimistic, see e.g. Woodall and Montgomery (1999) (p. 379), Ghosh et al. (1981), Quesenberry (1993), Chen (1997), Chakraborti (2000), Albers and Kallenberg (2004a) and Albers et al. (2004).

This situation can be repaired by applying suitable corrections to the estimated limits, as has been shown in Albers and Kallenberg (2004a, 2004b) (to be denoted for short as AK (2004a, 2004b) in the sequel). Due to the estimation the prescribed false alarm rate p (usually extremely small, like $p = 0.001$) is replaced by a stochastic version P_n, where n is the number of Phase I observations used and typically $Eg(P_n)$ 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 $P(RL \leq k)$ when RL has a geometric distribution with parameter p). Note that in case of estimating the parameters, conditional on the Phase I observations, RL is still geometric with parameter P_n and thus these expressions continue to make sense. For each of these three choices, suitable bias corrections are derived in AK (2004a).

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_n around p, rather than just look at its average behavior (or that of $g(P_n)$). We require that only in a fraction α (like $\alpha = 0.1$ or $\alpha = 0.2$) of the applications, one is faced with a value of P_n which is really too large in the sense that it exceeds not only p, but even $p(1 + \varepsilon)$ for some small $\varepsilon \geq 0$ (like $\varepsilon = 0.1$). Using once more the functions g leads to $P(g(P_n) > g(p)(1 + \varepsilon)) = \alpha$ for increasing g and $P(g(P_n) < g(p)(1-\varepsilon)) = \alpha$ for decreasing g. (Clearly, for $\varepsilon = 0$, g plays no role, as e.g. $P(P_n > p) = P(1/P_n < 1/p)$.) 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.) Fortunately, the impact on the out-of-control behavior will be moderate.

Hence, 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, see e.g. Chan et al. (1988), Pappanastos and Adams (1996) and Albers et al. (2004). A logical next step is to extend the normal family to a parametric model with an extra shape parameter to model distributions with heavier or lighter tails. Such families are studied in Albers, Kallenberg and Nurdiati (2004) (henceforth denoted by AKN (2004)) and there a specific choice, based on the so-called normal power family, is demonstrated to work well. The focus in AKN (2004) 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. The paper is organized as follows. In section 2 the chart for the normal power family is introduced and in section 3 corrections are derived with respect to exceedance probabilities for general parametric families. Such corrections are indeed larger than before on two counts: the model is more general 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. Next, section 5 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. The final section summarizes the conclusions.

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 one-sided case with a lower limit is completely analogous and the two-sided case is simply defined as the combination of the lower and upper limit, each with prescribed false alarm rate $p/2$. 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 more general model should be selected. Note that the standardized upper limit $(\widehat{UL} - \hat{\mu})/\hat{\sigma}$ in (2.1) simply equals u_p. For heavier tails, this value should increase, while for lighter tails the opposite should happen. This readily suggests to generalize u_p into a power like $u_p^{1+\gamma}$. Obviously $\gamma > 0$ (< 0) corresponds to heavy (light) tails, with the normal case as an interior point for $\gamma = 0$. The so obtained normal power family has intuitive appeal, is flexible, contains the normal, and moreover clearly covers a broad range of distributions. Consequently, as well as in AKN (2004), it will be the choice we shall adopt in the present paper.

After the explanation above, it merely remains to make matters precise, among others by introducing a normalization constant in order to ensure that σ still represents the standard deviation. 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 the difference between the normal chart and the normal power chart is that for the latter u_p in (2.1) needs to be replaced by $c(\hat{\gamma})u_p^{1+\hat{\gamma}}$, which indeed corresponds to the basic idea suggested above.

It is not our aim to fit the distribution globally, but only at the (far) right tail. This is done by concentrating on the ordinary tail (where we have observations), assuming that the behavior of the ordinary tail is informative for the far tail. Thus the restriction to symmetry occuring in the normal power family can be taken for granted. (When dealing with the two-sided case the fitted member of the normal power family for the lower control limit will be different from that for the upper control limit!) Consequently, instead of using all observations, we prefer to concentrate on the upper tail and to 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 $0 < q < r < 1/2$ (it turns out that $q = 0.05$ and $r = 0.25$ are good choices), define

$$\hat{\gamma}^* = \frac{X_{[n+1-qn]:n} - \overline{X}}{X_{[n+1-rn]:n} - \overline{X}}, \tag{2.3}$$

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

$$\widehat{UL} = \hat{\mu} + \hat{\sigma} c(\hat{\gamma}) u_p^{1+\hat{\gamma}}, \tag{2.4}$$

where $\hat{\gamma}$ is given through (2.3), with $q = 0.05$ and $r = 0.25$.

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 are now going to replace these $\widehat{UL}$ by

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

in which $c_e(\gamma)$, or c_e for short, well be an appropriate correction term. Obviously, by letting $c_e = 0$, we will always be able to reproduce the uncorrected charts, as $\widehat{UL}_0 = \widehat{UL}$.

Next we will make explicit what imposing the exceedance criterion on $\widehat{UL}_c$ from (3.1) means. Let $UL = \mu + \sigma\overline{K}_\gamma^{-1}(p)$ be the limit in probability of $\hat{\mu} + \hat{\sigma}\overline{K}_{\hat{\gamma}}^{-1}(p)$ under F. As P_n equals $\overline{F}(\widehat{UL}_c)$, the question now is how likely it is that $g(P_n)$, for example for 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. Note that we are only dealing with the so-called stochastic error here. The model error, $g(\overline{F}(UL)) - g(p)$, 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.

After specifying the criterion, we can now derive c_e as a function of n, p, α, ε and γ. First we introduce some notation. Let F_0 be the df of $(X_{n+1} - \mu)/\sigma$ and let E_0 and P_0 denote expectation and probability under F_0 (i.e. for the case where we simply have $\mu = 0$ and $\sigma = 1$). We assume the natural condition that $(\hat{\mu} - \mu)/\sigma$, $\hat{\sigma}/\sigma$ and $\hat{\gamma}$ have the same distribution under F as $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\gamma}$ under F_0, respectively. This condition obviously holds for $\hat{\mu} = \overline{X}$, $\hat{\sigma} = S$ and $\hat{\gamma}$ as in (2.4). Let

$$b_0 = \overline{F}_0^{-1}(g^{-1}(\{g(\overline{F}_0(\overline{K}_\gamma^{-1}(p)))(1+\varepsilon)\})). \tag{3.4}$$

As ε is supposed to be small, this implies that $b_0 - \overline{K}_\gamma^{-1}(p) \approx -\varepsilon Q$, with

$$Q = g(\overline{F}_0(\overline{K}_\gamma^{-1}(p)))/\{g'(\overline{F}_0(\overline{K}_\gamma^{-1}(p)))f_0(\overline{K}_\gamma^{-1}(p))\}, \tag{3.5}$$

where f_0 is the corresponding density $f_0 = F_0'$. Moreover, let

$$V = \hat{\mu} + \hat{\sigma}\overline{K}_{\hat{\gamma}}^{-1}(p) - \overline{K}_\gamma^{-1}(p), \tag{3.6}$$

then we have:

Lemma 3.1. *Suppose V from (3.6) satisfies $V/(E_0V^2)^{\frac{1}{2}} \overset{P_0}{\to} N(0,1)$, then $lim_{n\to\infty}$ $P(W_c > \varepsilon) = \alpha$ holds for*

$$c_e = (E_0V^2)^{\frac{1}{2}}u_\alpha + b_0 - \overline{K}_\gamma^{-1}(p). \tag{3.7}$$

Proof. Note that (3.3) translates into $P(\widehat{UL}_c < b) \le \alpha$, where b is b_0 from (3.4) with F rather than F_0 and UL instead of $\overline{K}_\gamma^{-1}(p)$. Consequently, $b = \mu + \sigma b_0$. From (3.1) it follows that $P(\widehat{UL}_c < b) = P((\hat{\mu} - \mu)/\sigma + (\hat{\sigma}/\sigma)\overline{K}_{\hat{\gamma}}^{-1}(p) + c_e < b_0) = P_0(V + c_e < b_0 - \overline{K}_\gamma^{-1}(p))$, with V as in (3.6). The desired result now is immediate in view of the conditions on V.$\square$

Typically E_0V and E_0V^2 are of order n^{-1} and hence the 'standardization' of V reduces to $V/(E_0V^2)^{\frac{1}{2}}$. Next we study the behavior of c_e from (3.7) in relation to the underlying parameters. Of main interest in this respect are n and ε.

1) role of n

Implicitly n is present in (3.7) through $(E_0V^2)^{\frac{1}{2}}$, which is typically of order $n^{-\frac{1}{2}}$. Note that this will indeed be the case for the normal power model, with $\hat{\mu} = \overline{X}$, $\hat{\sigma} = S$ and $\hat{\gamma}$ as in (2.4). Also observe that this order $n^{-\frac{1}{2}}$ is considerably larger than the order n^{-1} from the bias case (see AKN (2004)). Nevertheless, as n increases, this part of c_e will eventually become negligible. The speed of the convergence will decrease as the number of parameters to be estimated increases. For the normal chart from (2.1), (3.6) boils down to $V = \hat{\mu} + (\hat{\sigma} - 1)u_p$ and thus $E_0V^2 \approx (u_p^2 + 2)/(2n)$. For the normal power family, in addition to μ and σ we need to estimate γ, and E_0V^2 will be larger.

2) role of ε

First of all, for the boundary case where $\varepsilon = 0$, we have from (3.4) that $b_0 = \overline{K}_\gamma^{-1}(p)$ and the second part of (3.7) vanishes, as should be the case. As ε is small, this part can moreover be approximated by $-\varepsilon Q$, with Q given bij (3.5). To simplify this a bit further, let $\tilde{h}(p) = g(p)/\{pg'(p)\}$. For $g(p) = p$ we get $\tilde{h}(p) = 1$, for $g(p) = 1/p$ we find $\tilde{h}(p) = -1$, while for $g(p) = 1-(1-p)^k$ we obtain $\tilde{h}(p) = (1-p)\{(1-p)^{-k} - 1\}/(kp) \approx (e^{kp} - 1)/(kp)$.

Usually $\delta = kp$ will be small and in the last case $\tilde{h}(p) \approx 1 + \delta/2$. As a result we obtain as a further approximation

$$c_e = (E_0V^2)^{1/2}u_\alpha - \varepsilon\lambda\overline{F}_0(\overline{K}_\gamma^{-1}(p))/f_0(\overline{K}_\gamma^{-1}(p)), \tag{3.8}$$

where $\lambda = 1$ for the first two choices of g (in the second case $\tilde{h}(p) = -1$, 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.

The coefficient of ε in (3.8) obviously is negative, which agrees with the fact that increasing ε makes the criterion (3.3) less strict and thus requires a smaller correction. To provide some illustration, we consider the following example.

Example 3.1.
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 (3.8). If we are moreover in the normal case, this approximately boils down to u_p^{-1} and consequently (3.8) reduces to the approximation (cf. AK (2004b), (9)),

$$c_e = \left\{\frac{u_p^2 + 2}{2n}\right\}^{1/2} u_\alpha - \frac{\varepsilon\lambda}{u_p}. \tag{3.9}$$

According to (3.9), correction becomes superfluous once $n \approx \frac{1}{2}u_\alpha^2 u_p^2(u_p^2+2)/(\varepsilon\lambda)^2$. Choosing for example $p = 0.001$, $\varepsilon = 0.1$, $\alpha = 0.25$ and $\lambda = 1$, this produces $n \approx 2500$. Hence for (substantially) smaller n, the correction c_e will be positive and non-negligible. For further numerical illustruation consult AK (2004b). □

3) role of g
The effect of the choice of g, i.e., whether we concentrate on the signal probability itself, or rather on the ARL, or even on the distribution of RL, is represented concisely by the factor λ in (3.8). Note in particular that the fact that $\lambda = 1$ for both $g(p) = p$ and $g(p) = 1/p$ is of course no coincidence. In fact, in the latter case (3.3) gives $P(W_c < -\varepsilon) = P(\overline{F}(UL)/P_n < 1-\varepsilon) = P(P_n > \overline{F}(UL)/(1-\varepsilon))$, which is nothing but $P(W_c > \tilde{\varepsilon})$ in the case $g(p) = p$, with $\tilde{\varepsilon} = \varepsilon/(1-\varepsilon) \approx \varepsilon$ for small ε. Hence any result for the signal probability itself immediately translates into one about the ARL.

As concerns the remaining parameters of c_e in (3.7) or (3.8), we shall be more concise. If α decreases, (3.3) becomes more strict, which is reflected by the fact that the positive term in (3.7) increases linearly in u_α. Moreover, in accordance with the fact that very small p cause very large relative errors, E_0V^2 decreases in p. For the normal power family we will see this in the next section, while for the normal family this is already clear from (3.9), where the positive (negative) term increases (decreases) in u_p.

The next step towards application is the estimation of the unknown parts in (3.7) and (3.8). The fact that c_e is small to begin with, causes these changes to have negligible effects and approximate equality in (3.3) will continue to hold for the estimated versions $\hat{c}_e$ of c_e. First assume that we are in fact within the parametric family, and thus $F_0 = K_\gamma$. Hence we can simply use $\overline{K}_{\hat{\gamma}}$ for $\overline{F}_0$, $k_{\hat{\gamma}}$ for f_0 and also evaluate E_0V^2 under $K_{\hat{\gamma}}$. Denoting the latter by $\widehat{E_0V^2}$, we arrive from (3.7) at

$$\hat{c}_e = c_e(\hat{\gamma}) = (\widehat{E_0V^2})^{1/2}u_\alpha + \overline{K}_{\hat{\gamma}}^{-1}(g^{-1}(\{g(p)(1+\varepsilon)\})) - \overline{K}_{\hat{\gamma}}^{-1}(p), \tag{3.10}$$

which reduces through (3.8) to

$$\hat{c}_e = c_e(\hat{\gamma}) = (\widehat{E_0 V^2})^{1/2} u_\alpha - \frac{\varepsilon \lambda p}{k_{\hat{\gamma}}(\overline{K}_{\hat{\gamma}}^{-1}(p))}. \tag{3.11}$$

Hence with $\hat{c}_e$ from (3.10) or (3.11), the estimated upper limit

$$\widehat{UL}_c = \hat{\mu} + \hat{\sigma}\{\overline{K}_{\hat{\gamma}}^{-1}(p) + \hat{c}_e\} \tag{3.12}$$

from (3.1) will now produce approximate equality in (3.3) under the model $K_\gamma((x-\mu)/\sigma)$. This will also be true in the vicinity of this model: as long as F_0 is close to K_γ, estimation of F_0, f_0 and E_0V^2 in $\hat{c}_e$ through $K_{\hat{\gamma}}$ will make sense. Actual application in case of the special normal model already requires (see (3.9)) 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 $\widehat{E_0 V^2}$ 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 (2004) 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 (3.10) and (3.11) work out in practice for our prototype example, the normal power family. The first step is:

Lemma 4.1. *For the normal power family the results (3.10)-(3.12) specialize to: use*

$$\widehat{UL}_c = \hat{\mu} + \hat{\sigma}\{c(\hat{\gamma}) u_p^{1+\hat{\gamma}} + \hat{c}_e\}, \tag{4.1}$$

with

$$\hat{c}_e = (\widehat{E_0 V^2})^{1/2} u_\alpha + c(\hat{\gamma})\{u_{\tilde{p}}^{1+\hat{\gamma}} - u_p^{1+\hat{\gamma}}\}, \tag{4.2}$$

where $\tilde{p} = g^{-1}(\{g(p)(1+\varepsilon)\})$, or with the further simplification (writing ϕ for the standard normal density)

$$\hat{c}_e = (\widehat{E_0 V^2})^{1/2} u_\alpha - \varepsilon \lambda p (1+\hat{\gamma}) c(\hat{\gamma}) u_p^{\hat{\gamma}} / \phi(u_p). \tag{4.3}$$

Proof. In the normal power family $K_\gamma^{-1}(t) = c(\gamma)|\Phi^{-1}(t)|^{1+\gamma} sign(\Phi^{-1}(t))$ (cf. (2.2)). Together with (3.10) this readily gives (4.2). As moreover $\{1/k_\gamma(\overline{K}_\gamma^{-1}(1-t))\} = 1/k_\gamma(K_\gamma^{-1}(t)) = \{\overline{K}_\gamma^{-1}(t)\}' = (1+\gamma)c(\gamma)\Phi^{-1}(t)^\gamma/\phi(\Phi^{-1}(t))$ for $t > \frac{1}{2}$, using (3.11) rather than (3.10) leads to (4.3). □

The estimator to be used is again $\hat{\gamma}$, see (2.4). Indeed the main obstacle is $\widehat{E_0 V^2}$, which has to be expressed in terms of $\hat{\gamma}$ as well. After laborious computations the following result is obtained.

Lemma 4.2. *Using*

$$\hat{\gamma} = 1.1218 \log \left(\frac{X_{[0.95n+1]:n} - \overline{X}}{X_{[0.75n+1]:n} - \overline{X}} \right) - 1 \tag{4.4}$$

we arrive at $\{\widehat{E_0 V^2}\}^{1/2} \approx n^{-1/2} A(\hat{\gamma}, u_p)$, *where*

$$A(\hat{\gamma}, u_p) = -4.00 - 12.54\hat{\gamma} - 10.02\hat{\gamma}^2 + 2.91u_p + 6.47\hat{\gamma}u_p + 4.42\hat{\gamma}^2 u_p. \tag{4.5}$$

The proof of Lemma 4.2 is given in the Appendix.

Note that $\widehat{UL}_c$ from (4.1), with $\hat{c}_e$ from (4.2) (or the further aproximation from (4.3)) is now made completely explicit through (4.4) and (4.5) and can be applied in a straightforward manner. This leads to the following control limits for $g(p) = p$ and $g(p) = 1/p$, respectively,

$g(p)$	aim	upper control limit
p	$P\left(P_n > p\left(1+\varepsilon\right)\right) \le \alpha$	$\overline{X} + S\left\{c(\widehat{\gamma})u_{p(1+\varepsilon)}^{1+\widehat{\gamma}} + \frac{A(\widehat{\gamma}, u_p)u_\alpha}{\sqrt{n}}\right\}$
$1/p$	$P\left(\frac{1}{P_n} < \frac{1}{p}\left(1-\varepsilon\right)\right) \le \alpha$	$\overline{X} + S\left\{c(\widehat{\gamma})u_{p/(1-\varepsilon)}^{1+\widehat{\gamma}} + \frac{A(\widehat{\gamma}, u_p)u_\alpha}{\sqrt{n}}\right\}$

with $\widehat{\gamma}$ and $A(\widehat{\gamma}, u_p)$ given by (4.4) and (4.5), respectively. The lower control limits are obtained by replacing $+$ between $\overline{X}$ and S by $-$, while the two-sided control limits are simply given by the upper- and lower control limits with p replaced by $p/2$.

To exemplify these control limits we consider a real life example concerning the production of electric shavers by Philips. In an electrochemical process razor heads are formed. The measurements concern the thickness of the razor heads. A histogram of the sample is given in Figure 1.

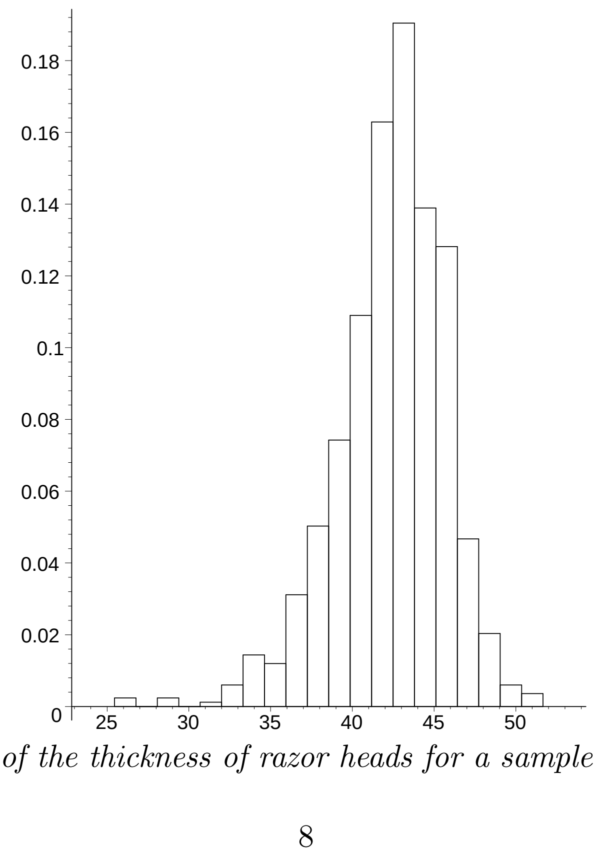

Figure 1. Histogram of the thickness of razor heads for a sample of 835 measurements.

The sample consists of 835 measurements. Application to the example gives for $p = 0.001$, $\varepsilon = 0.1, \alpha = 0.1$ the following upper control limits. We have $\overline{X} = 42.366, S = 3.311$, $n = 835, X_{[0.95n+1]:n} = 47.03, X_{[0.75n+1]:n} = 44.54, \widehat{\gamma} = -0.144$, while $u_p = 3.090, u_{p(1+\varepsilon)} = 3.062, u_{p/(1-\varepsilon)} = 3.059$ and $u_\alpha = 1.282$. We get $\widehat{UL}_c = 52.001$ when $g(p) = p$ and $\widehat{UL}_c = 51.994$ when $g(p) = 1/p$. The uncorrected traditional upper control limit (see (2.1)) equals 52.597. Indeed, it is seen in Figure 1 that we have a thin upper tail giving $\widehat{\gamma} < 0$ and hence the upper control limits are somewhat smaller than the uncorrected one. In the lower tail the situation is reversed. The corrected lower control limits are 28.306 when $g(p) = p$ and 28.324 when $g(p) = 1/p$. Comparison with the uncorrected lower control limit, 32.135, shows that the heavy lower tail (see Figure 1) causes a substantial correction.

We shall next perform a simulation study along the lines of AKN (2004). We shall concentrate on g$(p) = p$; from the derivations given (cf. remark 3) following Lemma 3.1) 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 (2004), where the values 100, 250 and 500 are considered, with the emphasis 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 (2004). 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 $\gamma = -0.5$, –0.25, 0.25, 0.75 and 1 ($\gamma = 0$ 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 $\lambda = -0.1$, 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 $P(P_n > p)$, which is seen to stabilize around $\frac{1}{2}$. Hence indeed, p turns out to be close to the 50%-quantile of the distribution of P_n. (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 $n \to \infty$, $P(P_n > p(1+\varepsilon)) \to 0$ 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	50	40	24	24	23	22	22	22	22	21
$K_{-0.5}$	50	47	50	46	50	45	51	43	21	21	19	19	20	20	20	20
$K_{-0.25}$	51	48	50	46	50	43	50	41	23	23	22	22	21	21	21	21
$K_{0.25}$	52	49	51	46	51	43	51	41	25	25	23	23	23	23	22	22
$K_{0.5}$	54	50	52	46	51	44	50	39	25	25	23	23	22	22	21	21
$K_{0.75}$	53	49	53	47	51	43	50	39	25	25	23	23	22	22	21	20
K_1	55	51	53	48	52	44	51	40	27	27	24	24	22	22	22	22
T	53	47	51	43	51	39	50	34	28	26	26	23	26	21	26	19
RM	53	47	51	43	51	39	50	34	26	25	24	21	24	21	23	18
DM	52	47	51	43	50	39	51	37	25	24	25	22	24	21	23	19
$TU(-0.1)$	53	47	51	43	51	38	51	34	29	27	28	24	25	21	26	19
$TU(0)$	52	47	50	43	51	40	51	37	27	26	25	23	24	21	24	20
$TU(0.14)$	51	48	51	47	51	44	51	42	24	24	22	22	22	22	23	23
O	52	49	51	46	51	45	51	43	23	24	22	23	21	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. As we have seen, during the in-control stage, the uncorrected chart tends to stop considerably earlier than anticipated in an unacceptably large fraction of its applications. The corrections serve to bring this fraction down to acceptable proportions ('controlling the in-control behavior') and as such typically delay the moment of stopping somewhat. Clearly, this tendency will be noticeable during out-of-control as well. But there it actually is unwelcome, as it means that detection will take a bit longer as well. Hence a balance will have to be found between controlling performance during in-control and loss of detection power during 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_n, which in view of (3.12) 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 (2004a)). 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

$$E(\Delta, \hat{c}_e) = E_\Delta g(\overline{F}_0(\widehat{UL}_c - \Delta)), RC = \begin{cases} 1 - E(\Delta, \hat{c}_e)/E(\Delta, 0) & \text{for increasing } g \\ E(\Delta, \hat{c}_e)/E(\Delta, 0) - 1 & \text{for decreasing } g. \end{cases} \quad (5.1)$$

Clearly, $E(\Delta, \hat{c}_e)$ and $E(\Delta, 0)$ stand for $E_\Delta g(P_n)$ with and without correction, respectively. Moreover, RC expresses the relative cost incurred by having to use the correction $\hat{c}_e$. 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 $\hat{c}_e$ has pushed up the ARL 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 $p = 0.001$, see Table 5.1: the simulated values of $E(\Delta, 0)$ 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 $\hat{c}_e$, respectively.

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, as expected, that 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 that 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.

Table 5.1 *Simulated values of* $E(\Delta, 0) = E_\Delta P_n$ *(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	0.50	1	1	0.50	0	0	0.50	0	0
$K_{-0.25}$	1	0.065	42	38	0.061	31	28	0.059	23	19	0.059	17	12
	2	0.35	19	17	0.35	13	12	0.35	9	8	0.35	7	5
$K_{0.25}$	2	0.061	47	43	0.055	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	0.094	39	33	0.088	29	23	0.085	21	15
	4	0.39	47	43	0.36	39	33	0.34	30	24	0.32	22	16
$K_{0.75}$	3	0.054	55	51	0.044	43	38	0.040	33	26	0.037	25	17
	4	0.20	57	52	0.16	45	39	0.14	34	28	0.13	25	18
K_1	3	0.032	57	53	0.024	46	40	0.021	35	28	0.020	26	19
	4	0.11	60	56	0.076	48	42	0.062	37	30	0.057	28	20
T	2	0.091	40	37	0.081	31	27	0.077	23	19	0.074	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	40	36	0.11	30	26	0.10	22	18	0.10	16	11
	3	0.41	29	26	0.41	21	18	0.41	15	12	0.41	10	7
$TU(-0.1)$	2	0.071	41	37	0.061	31	27	0.057	24	19	0.055	17	12
	3	0.29	39	36	0.27	30	26	0.26	23	18	0.26	17	12
$TU(0)$	2	0.098	41	37	0.088	31	27	0.084	23	18	0.083	17	12
	3	0.37	32	29	0.35	24	21	0.35	18	14	0.35	13	9
$TU(0.14)$	2	0.15	37	34	0.14	28	24	0.14	20	17	0.14	14	11
	3	0.46	23	22	0.46	17	15	0.46	12	10	0.46	8	6
O	2	0.092	44	40	0.086	33	29	0.084	24	19	0.083	18	12
	3	0.32	32	29	0.31	24	20	0.30	17	14	0.30	12	9

In view of the discussion above, we shall devote the remainder of this section to obtaining 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), with the following result.

Lemma 5.1. *RC from (5.1) approximately equals* $\widetilde{RC}$*, where*

$$\widetilde{RC} = \frac{4(1+u_{\tilde{p}})}{5\lambda(1+\gamma)c(\gamma)u_{\tilde{p}}^{\gamma}} A(\gamma, u_p) n^{-1/2} u_\alpha - \varepsilon \frac{(1+u_{\tilde{p}})u_p^{\gamma}}{(1+u_p)u_{\tilde{p}}^{\gamma}}, \tag{5.2}$$

in which $\tilde{p} = \overline{K}_\gamma(\overline{K}_\gamma^{-1}(p) - \Delta)$ *and* A *as given by (4.5).*

The proof of Lemma 5.1 is given in the Appendix.

By way of illustration, consider the following example.

Example 5.1
Select the values $p = 0.001$, $\lambda = 1$ and $\alpha = 0.2$ which are used throughout Table 5.1 and consider the case $\gamma = 0$, corresponding to $F_0 = \Phi$. Then (5.2) reduces to $\widetilde{RC} = 3.36(4.09 - \Delta)n^{-1/2} - \varepsilon(1 - \Delta/4.09)$. 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 $\gamma = 0$ considered in the example above, another interesting comparison can be made to the situation where normality is assumed to begin with, and thus estimation of γ is not needed. We have

Example 5.2
Suppose normality has been assumed and a normal chart is used. Then, in the situation of Example 5.1, we deal, according to (3.9) 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 (5.2) 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 (5.2) 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 (5.2) 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 $4(1 + u_{\tilde{p}})/\{5(1 + \gamma)c(\gamma)u_{\tilde{p}}^{\gamma}\}$ and $A(\gamma, u_p)$. Let γ be fixed, then we observe that it also decreases in p, as $A(\gamma, u_p)$ increases linearly in u_p (cf. (4.5)). 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. By way of illustration we mention an upper bound for certain values of γ and $\tilde{p}$ that might by considered reasonable, in the sense that in many applications the given ranges will apply.

Example 5.3
It can be checked numerically that for $p = 0.001$ and $\alpha = 0.2$, the coefficient of $n^{-1/2}$ in (5.2) will not exceed 9.43 for γ and $\tilde{p}$ such that $-0.25 \leq \gamma \leq 0.3$ and $0.1 \leq \tilde{p} \leq 0.3$. □

A final remark is that the first term in (5.2) 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

$$n \geq \left\{ \frac{4(1 + u_{\tilde{p}})A(\gamma, u_p)u_\alpha}{5(1 + \gamma)c(\gamma)u_{\tilde{p}}^{\gamma}\beta} \right\}^2. \tag{5.3}$$

Example 5.4
Continuing Example 5.3, it follows that for this case (5.5) boils down to $n \geq \{9.43/\beta\}^2$, 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 in Example 5.4 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.

6 Conclusions

Standard control charts simply use $\widehat{UL} = \hat{\mu} + \hat{\sigma} u_p$ (see (2.1)). If normality fails, this causes large errors in the sense that during applications of the chart quite often the resulting stochastic signal probability and ARL differ unpleasantly much from the intended values. To bring the corresponding exceedance probabilities within prescribed bounds, it is demonstrated that a more general model, based on the so-called normal power family, offers good results. Here $\widehat{UL} = \hat{\mu} + \hat{\sigma} c(\hat{\gamma}) u_p^{1+\hat{\gamma}}$ is used (see (2.5)), where $\hat{\gamma}$ (see (2.4)) is an estimator which selects the appropriate type of distribution. In this way the model error is reduced substantially, with on the other hand a larger stochastic error. As a remedy for this latter increase, corrections c_e (see (3.7) or (3.8)) are derived. By adding such a c_e to $c(\hat{\gamma}) u_p^{1+\hat{\gamma}}$ (see Lemma 4.1) in $\widehat{UL}$, the exceedance probabilities mentioned above are indeed adequately controlled (see Table 4.1). The premium to be paid for this protection during in-control unavoidably is some delay in stopping during out-of-control. An expression for the relative cost RC (see (5.1)) caused by applying c_e is obtained. It turns out that RC indeed becomes small as n increases, (see Table 5.1), but not really fast ($n \geq 1000$ would be nice). This is in agreement with the fact that c_e offers protection against both nonnormality and deviations in single applications of the chart. Hence, if possible, it is definitely advised to use c_e, but some care is needed. To this end, an approximation $\widetilde{RC}$ for RC is given (see Lemma 5.1), which leads to a lower bound on the required n in terms of the parameters involved (see (5.5)).

Acknowledgements The authors cordially thank dr. ir. J. Praagman (CQM Eindhoven) for providing us with the electric shaver data used in this paper.

References

Albers, W. and Kallenberg, W.C.M. (2004a). Estimation in Shewhart control charts: effects and corrections. *Metrika* **59**, 207-234.

Albers, W. and Kallenberg, W.C.M. (2004b). Are estimated control charts in control? *Statistics* **38**, 67-79.

Albers, W., Kallenberg, W.C.M. and Nurdiati, S. (2004). Parametric control charts. *J. Statist. Plann. Inference* **124**, 159-184.

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. *Statist. 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.

Appendix

Proof of Lemma 4.2. Note that $\hat{\gamma}$ from (4.4) is nothing but the explicit $\hat{\gamma} = h^{-1}(\hat{\gamma}^*)$, with $\hat{\gamma}^*$ as in (2.3), using again $q = 0.05$ and $r = 0.25$. In particular, $\log(u_{0.05}/u_{0.25}) = 1/1.1218$. The computation of $(E_0V^2)^{\frac{1}{2}}$ is a very technical and rather complicated matter. Fortunately, this task has been performed already in the Appendix of AKN (2004) for E_0V and E_0V^2, in order to obtain an explicit form of the bias correction. Hence for briefness sake we refer to that paper for technical details and merely present here an outline of the steps involved. The basic idea is actually quite straightforward. V from (3.6) is a function $\tilde{H}(\hat{\mu}, \hat{\sigma}, \hat{\gamma}^*)$, to which a second order expansion around $\tilde{H}(0, 1, \gamma^*)$ is applied. (Note that working under E_0 means that $\mu = 0$ and $\sigma = 1$.) After this the first and second order (mixed) moments of $\hat{\mu}$, $(\hat{\sigma} - 1)$ and $(\hat{\gamma}^* - \gamma^*)$ need to be calculated, all of which are of order n^{-1} (cf. $(A.1) - (A.5)$ from AKN (2004)). For those involving only $\hat{\mu}$ and/or $\hat{\sigma}$, this is immediate; as soon as $(\hat{\gamma}^* - \gamma^*)$ is involved, the computation is more tedious. Plugging the moment results into the expansion gives as a first version of the desired result an approximation up to $o(n^{-1})$ (cf. $(A.9)$ in AKN (2004)). However, as the expansion itself contains complicated coefficients such as the derivative of $\overline{K}^{-1}_{h^{-1}(\gamma^*)}(p)$ (cf. (2.3) and (3.6)), the result is still quite complicated. Therefore the asymptotic results are complimented with further numerical approximations for such coefficients leading to more amenable results. Using these two groups of results on $(E_0V^2)^{\frac{1}{2}}$ produces (4.5). □

Proof of Lemma 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 $\pm\, c_e f_0(\overline{K}_\gamma^{-1}(p) - \Delta)(g'/g)(\overline{F}_0(\overline{K}_\gamma^{-1}(p) - \Delta))$ with the + sign for increasing g and the $-$ sign for decreasing g. Using similar steps for the three choices of g involved as those leading to (3.8), it follows that RC approximately equals

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

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 (A.1), 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 $0<\tilde{p}\leq\frac{1}{2}$,

$$u_{\tilde{p}}^{-\gamma}k(u_{\tilde{p}})/\{(1+\gamma)c(\gamma)\}, \tag{A.2}$$

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 for example $0<x\leq 3.09=u_{0.001}$, the function k can be approximated quite adequately by $4(1+x)/5$. Combination of (4.3) and (4.5), together with (A.1) and (A.2) then leads for $\tilde{p}<\frac{1}{2}$ to the result in (5.2). $\square$