Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Journal of Statistical Planning and
Inference 124 (2004) 159–184

journal of
statistical planning
and inference

www.elsevier.com/locate/jspi

Parametric control charts☆

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

Faculty of Mathematical Sciences, University of Twente, P.O. Box 217, 7500 AE Enschede, Netherlands

Received 6 June 2002; accepted 19 March 2003

Abstract

Standard control charts are based on the assumption that the observations are normally distributed. In practice, normality often fails and consequently the false alarm rate is seriously in error. Application of a nonparametric approach is only possible with many Phase I observations. Since nowadays such very large sample sizes are usually not available, there is need for an intermediate approach by considering a larger parametric model containing the normal family as a submodel. In this paper control limits are presented in such a larger parametric model, the so called normal power family. Correction terms are derived, taking into account that the parameters are estimated. Simulation results show that the control limits are accurate, not only in the considered parametric family, but also for common distributions outside the parametric family, thus covering a broad class of distributions.

MSC: 62 F 12; 62 P 30; 65 C 05

Keywords: Statistical process control; Phase II control limits; Second order unbiasedness; Normal power family; Model error; Stochastic error

1. Introduction

Let X be a monitoring random variable in a production process which provides an out-of-control signal when X is larger than a certain control limit. Let p be the false alarm rate, that is the probability of concluding that the process is out-of-control when the process is in control, and let the in-control situation be modeled by assuming

☆ This research was supported by the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs.

* Corresponding author. Tel.: +31-53-4893374; fax: +31-53-4893069.
E-mail address: w.c.m.kallenberg@math.utwente.nl (W.C.M. Kallenberg).

0378-3758/$ - see front matter
doi:10.1016/S0378-3758(03)00200-3

that X has a normal distribution with known expectation μ and variance σ^2. Then the out-of-control signal is given if $X > \mu + \sigma u_p$, with $\Phi(u_p) = 1 - p$, where Φ is the standard normal distribution function.

In practice, it is hard to believe that we know μ and σ. Therefore, very often it is assumed that X belongs to the class of normal distributions with unknown parameters μ and σ. Based on observations from the past, so called Phase I observations, the parameters μ and σ have to be estimated. A recent reference in which the problems concerning estimation in control charts are explicitly mentioned, is Woodall and Montgomery (1999) (see p. 379); also see Ghosh et al. (1981), Quesenberry (1993), Chen (1997), Chakraborti (2000) and Neduraman and Pignatiello (2001). In Albers and Kallenberg (2000, 2001) it is shown that simply plugging in the estimators of μ and σ leads to inaccurate results, unless very large sample sizes are used. Since nowadays short production runs are more and more in demand, such large sample sizes are usually not available. Fortunately, simple but efficient correction terms can be derived, leading to control charts performing according to the required criterion; see Albers and Kallenberg (2000, 2001).

Therefore, as long as normality describes the behavior of X rather well, the corrected control charts can be applied in practice. But unfortunately, the probability of incorrectly producing an out-of-control signal may be seriously in error when the distributional form of the observations differs from normality, see e.g. Chan et al. (1988), Pappanastos and Adams (1996, Table 7 on p. 222). Basically, the problem is that the normal approximation may be fair for the central part of the distribution, but produces large relative errors in the tails. And, in view of the small values of p typically used, the tails are what we are dealing with. Therefore, a larger model is needed, providing more flexibility to describe the behavior in the far tail.

For that purpose we consider the so called normal power family, containing the normal distributions as submodel. One may object that this larger parametric family again does not cover all distributions and that a nonparametric approach should be taken. As typically the 0.999-quantile should be estimated, it is clear that with, say, 100 observations in Phase I such a quantile cannot be estimated nonparametrically. Hence, some assumptions should be made about the relation between the behavior of X in the far tail (where we should estimate an appropriate quantile) and the behavior of X “somewhat more to the middle”, where we have observations and can do the estimation. A similar type of argument is applied e.g. in de Haan and Sinha (1999), Hall and Weissman (1997) and Dekkers and de Haan (1989). The normal power family is on the one hand sufficiently rich, but on the other hand not that large that estimation is impossible or too inaccurate.

There are many other possibilities to extend the family of normal distributions, like random or deterministic mixtures. However, it turns out that several standard extensions of the family of normal distributions lead to very substantially more complicated control limits. We refer to Albers et al. (2002) for more details.

Using the normal power family gives more flexibility and hence an improvement over simply using normality and ignoring the well-known facts that in practice normality often fails and that this causes big errors in the false alarm rate. On the other hand, it is not our claim that the normal power family always is the “right” model.

But firstly, if control limits based on normality are applied, this should implicitly mean that the distribution is approximately normal, and in that case the normal power family is certainly appropriate, since normality is a submodel of the normal power family. Secondly, by the extension to the normal power family many other commonly used distributions show up, which are covered sufficiently well by members of the normal power family. Thirdly, the normal power family is not so large that accurate estimation is only possible with huge sample sizes as in the nonparametric approach.

In Section 2 an exposition of the problem is given with a discussion on model error due to misspecification and stochastic error implied by estimation. When assuming the restricted model of normality the corresponding model error turns out to be very large for many common distributions, thus showing the need for a larger parametric model.

Having reduced the model error by taking a larger parametric model, the next step is to avoid a large stochastic error due to estimation of the parameters. A correction term agreeing with the criterion at hand is required. The correction term depends on the chosen estimator as well. Since the normal power family is put forward as making the link between the (very) far tail and the behavior somewhat more to the middle, is it natural to base the estimator of the additional parameter on some order statistics in the ordinary tail. We use the 95% and 75% sample quantiles for this purpose. In Section 3 our recommended control limit is presented.

Theoretical properties of this control limit are discussed in Section 4. It is shown that there exists a control limit which (apart from a set with extremely low probability) reduces the bias due to estimation (up to $o(n^{-1})$). Unfortunately, this control limit is very complicated to calculate. Our recommended control limit is a numerical approximation of it, which on the one hand can be calculated quite straightforwardly and on the other hand is close to the complicated control limit.

A simulation study is performed to see how well the asymptotic results come true for finite sample sizes. It turns out that the recommended control limit works very well, not only when sampled from distributions belonging to the normal power family, but also for other distributions like random and deterministic mixtures. This provides a great improvement of the false alarm rate based on the normality assumption, which may fail seriously. The loss when really having normal observations is small. The simulation results are presented in Section 5.

In Section 6 a discussion is given of the results in view of the main questions stated in Section 2 including the out-of-control behavior. The paper is closed by an appendix containing some derivations of the theoretical results of Section 4 and asymptotic formulas used in Section 5.

2. The main questions

Let $X_1,\ldots,X_n,X_{n+1}$ be i.i.d. random variables (r.v.'s) with distribution function F, that is, we consider the in-control situation. The out-of-control behavior is treated in Section 6, where X_{n+1} comes from a shifted distribution function $F(x-\Delta)$. The r.v.'s $X_1,\ldots,X_n$ are the data from Phase I on which the estimators of the unknown parameters are based and X_{n+1} is the monitoring characteristic. The monitoring r.v. may be based on m observations, but here we consider the situation $m=1$ of individual measurements to avoid additional complications, thus facilitating the explanation of the basic arguments. The case $m>1$ will be treated in a next paper.

The true, but unknown distribution function F is modeled by the normal power family. Let Z be a r.v. with a standard normal distribution. For $\gamma > -1$ define $Z_\gamma = c(\gamma)|Z|^{1+\gamma}\,\mathrm{sign}(Z)$, with $c(\gamma)$ a normalizing constant given by $c(\gamma)=\{E|Z|^{2(1+\gamma)}\}^{-1/2}$. A r.v. X belongs to the normal power family if

$$X=\mu+\sigma Z_\gamma,$$

for some $\mu\in\mathbb{R}$ and $\sigma>0$. The distribution function of X is denoted by $K_\gamma((x-\mu)/\sigma)$ and hence K_γ is the distribution function of Z_γ. Hopefully F equals $K_\gamma((x-\mu)/\sigma)$ for some (μ,σ,γ), or even for some $(\mu,\sigma,0)$, but this is not necessarily true.

For any distribution function H we write $\bar{H}=1-H$ and H^{-1} and $\bar{H}^{-1}$ for the respective inverse functions. The upper p-quantile $\left(0<p<\frac{1}{2}\right)$ of Z_γ is given by

$$\bar{K}_\gamma^{-1}(p)=c(\gamma)u_p^{1+\gamma}=\pi^{1/4}2^{-(1+\gamma)/2}\Gamma\left(\gamma+\frac{3}{2}\right)^{-1/2}u_p^{1+\gamma}.$$

Taking $\gamma=0$ we get (as submodel) the family of normal distributions, to which we refer as the *restricted model*. The extra parameter γ gives more flexibility to model the upper p-quantile. Especially, heavier tails than those of the normal distribution are of interest. In terms of high upper quantiles this means larger values than the normal upper quantiles. For instance, with $p=0.001$ the upper p-quantile of the standard normal distribution equals 3.09, while $\bar{K}_{0.25}^{-1}(0.001)=3.69$ and $\bar{K}_{0.5}^{-1}(0.001)=4.30$. Smaller values than the normal upper quantile are available for negative γ as for instance $\bar{K}_{-0.25}^{-1}(0.001)=2.51$.

If $F(x)=K_\gamma((x-\mu)/\sigma)$ and the parameters μ,σ,γ are known, the control limit equals $\mu+\sigma\bar{K}_\gamma^{-1}(p)=\mu+\sigma c(\gamma)u_p^{1+\gamma}$. Often F is unknown and two problems arise: (i) F may be (slightly) outside the normal power family (or the family of normal distributions if the supposed model is the restricted model) and (ii) the parameters μ,σ,γ (or μ,σ if the supposed model is the restricted model) are unknown. This leads to two kinds of errors, the *model error* and the *stochastic error*. We speak of the *restricted model error* and the *restricted stochastic error* if the supposed model is the family of normal distributions. When the supposed model is the normal power family we simply speak of model error and stochastic error.

Note that due to estimation of the parameters the false alarm rate is no longer a number, but a r.v. which we denote by P. The parameters μ and σ will be estimated in this paper by the usual location and scale estimators $\bar{X}=n^{-1}\sum X_i$ and $S=\sqrt{S^2}$ with $S^2=(n-1)^{-1}\sum(X_i-\bar{X})^2$. We also sometimes write $\hat{\mu}$ and $\hat{\sigma}$ for $\bar{X}$ and S, respectively. As the actual distribution function F drifts away from Φ, it becomes feasible that more robust estimators are used. However, we do not go too deeply into these kinds of problems, since they are known to exist already for a long time and have no specific relation to the present setup.

Let us first consider the situation that the supposed model is the restricted model of normality. The total error $P - p$ is split up as

$$P - p = P(X_{n+1} > \hat{\mu} + \hat{\sigma} u_p) - p = \bar{F}(\hat{\mu} + \hat{\sigma} u_p) - p = \mathrm{RME} + \mathrm{RSE} \tag{2.1}$$

with the restricted model error RME given by

$$\mathrm{RME} = \bar{F}(\mu + \sigma u_p) - p = \bar{F}(\mu + \sigma u_p) - \bar{F}(\bar{F}^{-1}(p)) \tag{2.2}$$

and the restricted stochastic error RSE defined by

$$\mathrm{RSE} = \bar{F}(\hat{\mu} + \hat{\sigma} u_p) - \bar{F}(\mu + \sigma u_p). \tag{2.3}$$

Example 2.1. Suppose that the true distribution comes from the normal power family, that is $F(x) = K_\gamma((x - \mu)/\sigma)$. Then

$$\mathrm{RME} = \overline{K_\gamma}(u_p) - p = \bar{\Phi}\left(\left(\frac{u_p}{c(\gamma)}\right)^{1/(1+\gamma)}\right) - p.$$

For $p = 0.001$ and $\gamma = -\frac{1}{4}$ we get $c(-\frac{1}{4}) = 1.0783$ and hence $\mathrm{RME} = -0.00098$. For $p = 0.001$ and $\gamma = \frac{1}{2}$ we get $c\left(\frac{1}{2}\right) = \frac{1}{2}\sqrt[4]{2\pi}$ and hence $\mathrm{RME} = 0.00558$, while for $p = 0.001$ and $\gamma = 1$ we have $c(1) = 3^{-1/2}$ and thus $\mathrm{RME} = 0.00935$.

Let the true distribution be the Student distribution with six degrees of freedom. Taking $p = 0.001$ we get $\mathrm{RME} = 0.00356$. For a number of other examples of RME's we refer to Albers et al. (2002).

Remark 2.1. From the calculations of RME it is seen that RME can be negative, implying a lower false alarm rate. From the point of view of the false alarm rate this looks nice. However, this will certainly have harmful consequences for the out-of-control behavior. For instance, when $\gamma = -\frac{1}{4}$ in the normal power family, $\mathrm{RME} = -0.00098$. To illustrate what this might imply for the out-of-control behavior, we consider the more simple situation of a control chart for a standard normal distribution with $p = 0.001$ and $p = 0.001 - 0.00098 = 0.00002$. Then the expected run length to detect a shift 2 equals 7.3 for $p = 0.001$ and no fewer than 57.0 for $p = 0.00002$. Hence, both positive and negative model errors should be reduced.

Conclusion. From Example 2.1 it is seen that the restricted model error can be quite large. Often RME is several factors larger than the prescribed p. Therefore, if normality fails, there is a need for a larger model, thus reducing the model error.

Next we consider as supposed model the normal power family. Then the (uncorrected) control limit has the following form

$$\hat{\mu} + \hat{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p),$$

with $\hat{\gamma}$ a suitable estimator of γ. Hence, the total error can be split up as

$$P - p = P(X_{n+1} > \hat{\mu} + \hat{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p)) = \bar{F}(\hat{\mu} + \hat{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p)) - p = \mathrm{ME} + \mathrm{SE} \tag{2.4}$$

with the model error ME given by

$$\mathrm{ME} = \bar{F}(\mu + \sigma\bar{K}_{\gamma}^{-1}(p)) - p = \bar{F}(\mu + \sigma\bar{K}_{\gamma}^{-1}(p)) - \bar{F}(\bar{F}^{-1}(p)) \tag{2.5}$$

for some suitably chosen point γ and the stochastic error SE defined by

$$\mathrm{SE} = \bar{F}(\hat{\mu} + \hat{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p)) - \bar{F}(\mu + \sigma\bar{K}_{\gamma}^{-1}(p)). \tag{2.6}$$

The "suitably chosen point γ" is determined by the estimator of γ. We will always take a consistent estimator of γ and hence if the true distribution belongs to the normal power family, the "suitably chosen point γ" is simply the parameter value γ of the true distribution. If the true distribution is outside the normal power family, the "suitably chosen point γ" is determined by the functional associated with the estimator of γ. (In fact the same holds for the μ and σ appearing in (2.2), (2.3), (2.5) and (2.6). They are given by $\mu = \int x\,\mathrm{d}F(x)$ and $\sigma = \sqrt{\int (x-\mu)^2\,\mathrm{d}F(x)}$, the functionals related to $\bar{X}$ and S.)

Our recommended estimator of γ is based on the 95% and 75% sample quantiles. Writing $X_{m:n}$ for the mth order statistic of $X_1,\dots,X_n$ and writing $u_t = \bar{\Phi}^{-1}(t), 0 < t < \frac{1}{2}$ it is defined by

$$\frac{\log((X_{[0.95n+1]:n} - \bar{X})/(X_{[0.75n+1]:n} - \bar{X}))}{\log(u_{0.05}/u_{0.25})} - 1, \quad \left(\frac{1}{\log(u_{0.05}/u_{0.25})} \approx 1.1218\right),$$

where $[x]$ denotes the entier of x. Let F be the true distribution function, corresponding to the r.v. X and let F_0 be the distribution function of $(X-\mu)/\sigma$. Then the suitably chosen point γ equals

$$\begin{aligned} &\frac{\log((\bar{F}^{-1}(0.05) - \int x\,\mathrm{d}F(x))/(\bar{F}^{-1}(0.25) - \int x\,\mathrm{d}F(x)))}{\log(u_{0.05}/u_{0.25})} - 1 \\ &= \frac{\log(\overline{F_0}^{-1}(0.05)/\overline{F_0}^{-1}(0.25))}{\log(u_{0.05}/u_{0.25})} - 1 \end{aligned} \tag{2.7}$$

and indeed this gives γ if X belongs to the normal power family with parameter γ. Obviously, the model error for distributions in the normal power family is reduced to 0. But also a reduction for distributions outside the normal power family is achieved. For instance, when the true distribution is the Student distribution with 6 degrees of freedom, the model error is reduced from 0.00356 to 0.00208. For more examples we refer to the tables in Section 5.

The larger the parametric family, the smaller in general the model error, but the more difficult the estimation problem (more parameters involved) and hence the larger the stochastic error. We want to control both the model error and the stochastic error. To establish this, we consider three cases:

1. We are in the restricted model, that is the observations are normally distributed.
2. The observations are from the normal power family.
3. We are outside the normal power family.

The main questions treated in the paper are the following.

1. When the restricted model is true, that is normality holds, both RME and ME are equal to 0. Hence, we only have to control RSE and SE. As a rule, SE will be larger than RSE. How large is this difference? Can RSE and SE be reduced by (simple) corrections? How large are the corrected RSE and SE, when normality holds?
2. When the observations are from the normal power family, ME equals 0, but as a rule RME does not. How bad can RME be? How does this balance with the larger SE? Can RSE and SE be reduced by (simple) corrections? How large are the corrected RSE and SE, when the observations are from the normal power family?
3. When we are outside the normal power family, the ME's can in principle be very large if we are far away from the normal power family. In the present paper we concentrate on the parametric approach and hence we consider only situations outside the normal power family where the ME's are not too big. In a forthcoming paper the nonparametric approach will be considered and then also departures farther away from the normal power family are considered. But, as already observed in the introduction, the nonparametric approach only makes sense if some flexibility is allowed w.r.t. n and/or p: for e.g. $n = 100$ and $p = 0.001$ as given quantities, nothing can be done.
   How large are the uncorrected and corrected SE and the uncorrected and corrected RSE, when we are outside the normal power family and ME is not too big? How is the total error when applying the corrected control limit with as supposed model the normal power family and how does this compare with the total error when applying the corrected control limit with as supposed model normality?
4. What is the impact on the out-of-control behavior of the chart? Does the process stop at a reasonable time when it has gone out-of-control?

Remark 2.2. The notions of model error and stochastic error can be defined in a quite general set-up, invoking also more than one additional parameter. Such a theory is treated in Albers et al. (2002), where also an extensive discussion on other possible models than the normal power family is presented. The conditions for an appropriate general model are however rather comprehensive. Therefore, several classical ways of extending the normal model turn out to cause (technical) difficulties.

Remark 2.3. For estimating the parameter γ we may use (standard) estimators based on all observations; see Albers et al. (2002) for an estimator of that type and the corresponding (corrected) control limit. However, we only believe in our model mildly. Especially, we hope that the behavior in the ordinary tail has some relevance for the information in the far tail, where we need to go. Hence, we want to rely on the largest k order statistics only. The same type of reasoning is applied in extreme-value theory, see e.g. Dekkers and de Haan (1989) and Hall and Weissman (1997).

As a second argument, note that a restriction as symmetry, for instance occurring in the normal power family, can be taken for granted, since we are only interested in fitting the tail of the distribution. Especially when dealing with the largest k order statistics only, symmetry plays no role and hence is no restriction at all. Indeed, the recommended estimator is based on the largest order statistics.

Remark 2.4. The estimator based on only one quantile may lead to problems, since the corresponding functional $\bar{K}_\gamma^{-1}(q) = c(\gamma)(\bar{\Phi}^{-1}(q))^{1+\gamma}$ is not monotone in γ for e.g. $q = 0.1$.

3. Corrected control limit

Due to estimation the (observed) false alarm rate is a random variable P, given by (2.1) or (2.4). We want P to be close to p. To compare the stochastic P and the deterministic p, several criteria can be used. The most obvious one is to consider EP and to compare it with p. Another possibility is to investigate the average run length, leading to a comparison of $E(1/P)$ and $1/p$. Consideration of the probability that the run length is at most some specified value k, leads to comparison of $E\{1-(1-P)^k\}$ with $1-(1-p)^k$.

More generally, we consider a function $g(p)$, estimate it by $g(P)$ and compare $Eg(P)$ with $g(p)$. In particular, we focus on the previously mentioned functions

$$g(p) = p, \quad g(p) = \frac{1}{p}, \quad g(p) = 1-(1-p)^k. \tag{3.1}$$

Typical values of interest for k are small fractions of the average run length, that is $k = [\delta/p]$ with small δ.

Unfortunately, even in the case of normal r.v.'s the classical empirical rules for choosing the number of observations to estimate μ and σ are inadequate, see Quesenberry (1993), Chen (1997), Roes (1995) and Albers and Kallenberg (2000). On the other hand, when normality holds, simple corrections can be developed, leading to control limits that meet for common sample sizes the required conditions posed by the criteria involved. Here we derive similar corrections in the present set-up in order to get $Eg(P)$ close to $g(p)$.

Elimination of the systematic error is of course not the only criterion. Due to the variation in the possible outcomes of $g(P)$ one may require that only in relatively few cases the prescribed $g(p)$ is exceeded. Corrections (of a stronger type than the one met here) can be proposed to satisfy such an exceedance probability criterion, see Albers and Kallenberg (2001). However, in this paper we restrict attention to the bias as criterion.

Theory and simulations clearly show (see Sections 4 and 5) that for sample sizes like $n = 100$ large stochastic errors occur. Starting with the control limit $\mu + \sigma\bar{K}_\gamma^{-1}(p)$, which has for known parameters μ, σ and γ probability p of incorrectly concluding that the process is out-of-control, and subsequently simply plugging in the estimators, we arrive for $Eg(P)$ at a value not equal to $g(p)$. Therefore, we change the starting value p to q, say, in such a way that, when estimating the parameters, we end up with $Eg(P) = g(p)$ (or at least close to it). In other words, we do not use $\bar{K}_\gamma^{-1}(p)$, but $\bar{K}_\gamma^{-1}(q)$ for an appropriate value of q. Instead of $\bar{K}_\gamma^{-1}(q)$ we write $\bar{K}_\gamma^{-1}(p) + c$ with c being the correction term. One might argue that the corrected control limit could be obtained by simply adding a correction term $\tilde{c}$. However, the smaller σ, the smaller the correction term and hence we give it the form $c\sigma$, estimate σ and arrive at

$$\hat{\mu} + \hat{\sigma}(\bar{K}_{\hat{\gamma}}^{-1}(p) + c).$$

Furthermore, note that the correction term depends on n, p, the function g and γ. The last parameter is estimated again and one may ask whether this estimation step in the correction term gives no renewed bias term. However, the effect here is much smaller as the correction term itself is of order n^{-1}.

The idea behind the derivation of the correction term is as follows. With a given correction term c, first we give an expression for the bias (as a function of c). Then we calculate the correction term c such that it eliminates the bias. In order to do these calculations we have to make a lot of approximations, which are of two types. There are numerical approximations to complicated functions like $\bar{K}_{\gamma}^{-1}(p)$ in order to do the analysis. Secondly, we apply asymptotics with respect to n, ignoring o(n^{-1})-terms. The derivation is still complicated, but fortunately leads to a proposal that can be applied straightforwardly. For a detailed discussion of the derivation we refer to Albers et al. (2002), but see also Section 4 and the Appendix.

We recommend the following control limit:

$$\hat{\mu} + \hat{\sigma}\left\{\bar{K}_{\hat{\gamma}}^{-1}(p) - C1(\hat{\gamma})C2(\hat{\gamma}) - \frac{C3(\hat{\gamma})}{n} + \lambda\,\frac{C4(\hat{\gamma})}{n}\right\}, \tag{3.2}$$

where $\hat{\mu} = n^{-1}\sum X_i$ and $\hat{\sigma} = S = \sqrt{S^2}$ with $S^2 = (n-1)^{-1}\sum(X_i - \bar{X})^2$, $\bar{K}_{\gamma}^{-1}(p) = \pi^{1/4}2^{-(1+\gamma)/2}\Gamma(\gamma+\frac{3}{2})^{-1/2}u_p^{1+\gamma}$ refers to the normal power family, $\lambda = 1, -1$, or $1-\delta$ according to $g(p) = p$, $g(p) = \frac{1}{p}$, or $g(p) = 1-(1-p)^k$, $k = [\delta/p]$, respectively, and where moreover

$$\hat{\gamma} = 1.1218\log\left(\frac{X_{[0.95n+1]:n} - \bar{X}}{X_{[0.75n+1]:n} - \bar{X}}\right) - 1, \tag{3.3}$$

$$C1(\gamma) = -1.23 - 0.63\gamma + 0.73\gamma^2 + 0.74u_p - 0.08\gamma u_p - 0.14\gamma^2 u_p,$$

$$C2(\gamma) = \left(\frac{\Phi^{-1}\left(\frac{[0.95n+1]}{n+1}\right)}{\Phi^{-1}\left(\frac{[0.75n+1]}{n+1}\right)}\right)^{1+\gamma} - 2.4387^{1+\gamma},$$

$$C3(\gamma) = -10.86 - 27.77\gamma - 22.36\gamma^2 + 4.72u_p + 9.98\gamma u_p + 7.29\gamma^2 u_p,$$

$$C4(\gamma) = -87.23 - 147.89\gamma - 104.29\gamma^2 + 40.25u_p + 63.69\gamma u_p + 44.47\gamma^2 u_p.$$

Remark 3.1. Incidentally, $(X_{[0.95n+1]:n} - \bar{X})/(X_{[0.75n+1]:n} - \bar{X})$ may be negative. This happens with extremely small probability and therefore it does not matter much how this is repaired. E.g. we may take the absolute value leading to $\hat{\gamma} = 1.1218\log(|(X_{[0.95n+1]:n} - \bar{X})/(X_{[0.75n+1]:n} - \bar{X})|) - 1$.

4. Theoretical results

It is shown in Albers and Kallenberg (2000, 2001) that, even when normality holds, without correction $Eg(P)$ may differ substantially from $g(p)$, unless very large sample sizes are used. But under normality simple corrections can be derived which produce for common sample sizes results that are sufficiently accurate with respect to the criteria involved. A similar correction can be derived in the present set-up in order to get $Eg(P)$ close to $g(p)$. It turns out that this correction term is rather complicated. Therefore, we have developed the simplification already given in (3.2) and (3.3). It will be shown both theoretically (in this section) and by simulations (in Section 5) that for this recommended control limit $Eg(P)$ is close to $g(p)$.

Writing in general $c_u = c_u(\hat{\gamma})$ for a correction term, leading to the control limit $\hat{\mu} + \hat{\sigma}\{\bar{K}_{\hat{\gamma}}^{-1}(p) + c_u(\hat{\gamma})\}$, the false alarm rate is now given by

$$P(X_{n+1} > \hat{\mu} + \hat{\sigma}\{\bar{K}_{\hat{\gamma}}^{-1}(p) + c_u\}) = \bar{F}(\hat{\mu} + \hat{\sigma}\{\bar{K}_{\hat{\gamma}}^{-1}(p) + c_u\}) = \bar{F}_0\left(\bar{K}_{\gamma}^{-1}(p) + V + \frac{\hat{\sigma}}{\sigma}c_u\right),$$

where F_0 is the distribution function of $\sigma^{-1}(X_{n+1} - \mu)$ and where

$$V = \frac{\hat{\mu} + \hat{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p) - \{\mu + \sigma\bar{K}_{\gamma}^{-1}(p)\}}{\sigma} = \frac{\hat{\mu} - \mu}{\sigma} + \frac{\hat{\sigma}}{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p) - \bar{K}_{\gamma}^{-1}(p).$$

It is seen from the definitions of our estimators $\hat{\mu}, \hat{\sigma}$ and $\hat{\gamma}$ that without loss of generality we may assume that $\mu = 0$ and $\sigma = 1$ and that $X_1, \ldots, X_n, X_{n+1}$ have distribution function F_0. Therefore, the false alarm rate reads as $\bar{F}_0(\bar{K}_{\gamma}^{-1}(p) + V + \hat{\sigma}c_u)$ with

$$V = \hat{\mu} + \hat{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p) - \bar{K}_{\gamma}^{-1}(p). \tag{4.1}$$

For the estimators $\hat{\mu}, \hat{\sigma}$ and $\hat{\gamma}$ we restrict attention to neighborhoods of $\mu(=0)$, $\sigma(=1)$ and γ. The error involved by this is presented in the following theorem. For mathematical convenience the formal definition of the estimator $\hat{\gamma}$ in the theorems is the "exact" representation with $1/\log(u_{0.05}/u_{0.25})$ instead of its numerical value up to four decimals 1.1218. Similarly, we use the "exact" form of $C2$, with $u_{0.05}/u_{0.25}$ instead of its numerical value up to four decimals 2.4387. The difference due to this replacement is extremely small (e.g. $\bar{K}_{\gamma}(\bar{K}_{\tilde{\gamma}}^{-1}(p)) - \bar{K}_{\gamma}(\bar{K}_{\gamma}^{-1}(p))$ with $\tilde{\gamma} = 1.1218\log(u_{0.05}/u_{0.25})^{1+\gamma} - 1$ is of order 10^{-7}) and is therefore ignored. Let

$$\hat{\gamma}^* = \frac{X_{[0.95n+1]:n} - \bar{X}}{X_{[0.75n+1]:n} - \bar{X}}.$$

When the observations are from the normal power family, $\hat{\gamma}^*$ converges to

$$\gamma^* = \frac{\bar{K}_{\gamma}^{-1}(0.05)}{\bar{K}_{\gamma}^{-1}(0.25)} = \left(\frac{u_{0.05}}{u_{0.25}}\right)^{1+\gamma} = h(\gamma), \quad \text{say.}$$

Its inverse is given by

$$h^{-1}(x) = \frac{\log(x)}{\log(u_{0.05}/u_{0.25})} - 1$$

and $\hat{\gamma} = h^{-1}(\hat{\gamma}^*)$.

Theorem 1. *Let $X_1,\dots,X_n$ be i.i.d. r.v.'s with a normal power distribution with parameter γ. Then for each $\varepsilon > 0$*

$$\limsup_{n\to\infty} n^{-\min(1,2/(1+\gamma))} \log P(|\bar{X}| > \varepsilon) < 0, \tag{4.2}$$

$$\limsup_{n\to\infty} n^{-\min(1,1/(1+\gamma))} \log P(|S^2 - 1| > \varepsilon) < 0 \tag{4.3}$$

and

$$\limsup_{n\to\infty} n^{-\min(1,2/(1+\gamma))} \log P(|\hat{\gamma}^* - \gamma^*| > \varepsilon) < 0. \tag{4.4}$$

The proof of Theorem 1 is based on large deviation theory. Note that for $\gamma > 1$, the moment generating function of X_i, having a normal power distribution with parameter γ, does not exist and therefore the results of Theorem 1 and the proof are not quite standard. Because of lack of space we do not give the proof here. Details can be obtained from the authors.

Define

$$A_n(\varepsilon) = \{|\hat{\mu}| \leqslant \varepsilon, |\hat{\sigma}^2 - 1| \leqslant \varepsilon, |\hat{\gamma}^* - \gamma^*| \leqslant \varepsilon\}$$

and $A_n^c(\varepsilon)$ its complement, then Theorem 1 implies $P(A_n^c(\varepsilon)) \leqslant \exp\{-\eta n^{\min(1,1/(1+\gamma))}\}$ for some $\eta > 0$ and hence for each $\varepsilon > 0$ we have $P(A_n^c(\varepsilon)) = o(n^{-1})$ as $n \to \infty$.

Let

$$c_{u1}(\hat{\gamma}) = \left\{ -B1_n(\hat{\gamma}) + \frac{1}{2} B2_n(\hat{\gamma}) \left[\frac{g''(p)}{g'(p)} k_{\hat{\gamma}}(\bar{K}_{\hat{\gamma}}^{-1}(p)) - \frac{k'_{\hat{\gamma}}}{k_{\hat{\gamma}}} (\bar{K}_{\hat{\gamma}}^{-1}(p)) \right] \right\} 1_{A_n(\varepsilon)},$$

where $1_{A_n(\varepsilon)}$ is the indicator function of the set $A_n(\varepsilon)$, k_γ is the density of Z_γ and where $B1_n(\gamma)$ and $B2_n(\gamma)$ are given in (A.7) and (A.9), respectively. Let $X_1,\dots,X_n,X_{n+1}$ be i.i.d. r.v.'s with a normal power distribution with parameter γ. It is shown in Appendix A that $B1_n(\gamma) = O(n^{-1})$, $B2_n(\gamma) = O(n^{-1})$, that for sufficiently small $\varepsilon > 0$

$$EV 1_{A_n(\varepsilon)} = B1_n(\gamma) + o(n^{-1}), \quad EV^2 1_{A_n(\varepsilon)} = B2_n(\gamma) + o(n^{-1})$$

and that

$$Ec_{u1}(\hat{\gamma}) = -B1_n(\gamma) + \frac{1}{2} B2_n(\gamma) \left[\frac{g''(p)}{g'(p)} k_\gamma(\bar{K}_\gamma^{-1}(p)) - \frac{k'_\gamma}{k_\gamma} (\bar{K}_\gamma^{-1}(p)) \right] + o(n^{-1}).$$

Let

$$c_{u2}(\hat{\gamma}) = \left\{ -C1(\hat{\gamma})C2(\hat{\gamma}) - \frac{C3(\hat{\gamma})}{n} + \lambda \frac{C4(\hat{\gamma})}{n} \right\} 1_{A_n(\varepsilon)}, \quad \text{cf. (3.2) and (3.3)},$$

then

$$Ec_{u2}(\hat{\gamma}) = -C1(\gamma)C2(\gamma) - \frac{C3(\gamma)}{n} + \lambda\,\frac{C4(\gamma)}{n} + \mathrm{o}(n^{-1}).$$

The corresponding false alarm rates are denoted by $P^{(1)}$ and $P^{(2)}$, respectively, that is

$$P^{(j)} = \bar{K}_\gamma(\bar{K}_\gamma^{-1}(p) + V + \hat{\sigma}c_{uj}(\hat{\gamma})), \quad j = 1,2.$$

The correction term $c_{u2}(\hat{\gamma})$ is on the set $A_n(\varepsilon)$ the same as our recommended one. So, apart from a set with extremely small probability, the false alarm rate of the recommended control limit equals $P^{(2)}$. Define

$$W_g(p,\gamma) = -B1_n(\gamma) + \frac{1}{2}\,B2_n(\gamma)\left[\frac{g''(p)}{g'(p)}\,k_\gamma(\bar{K}_\gamma^{-1}(p)) - \frac{k'_\gamma}{k_\gamma}\,(\bar{K}_\gamma^{-1}(p))\right]$$
$$-\left\{-C1(\gamma)C2(\gamma) - \frac{C3(\gamma)}{n} + \lambda\,\frac{C4(\gamma)}{n}\right\},$$

then $Ec_{u1}(\hat{\gamma}) - Ec_{u2}(\hat{\gamma}) = W_g(p,\gamma) + \mathrm{o}(n^{-1})$.

Theorem 2. *Let $X_1,\dots,X_n,X_{n+1}$ be i.i.d. r.v.'s with a normal power distribution with parameter γ. Let g be three times continuously differentiable at p. Then for sufficiently small $\varepsilon > 0$ we have*

$$E\{g(P^{(1)})1_{A_n(\varepsilon)}\} = g(p) + \mathrm{o}(n^{-1}) \quad \text{as } n \to \infty, \tag{4.5}$$

and

$$|E\{g(P^{(1)})1_{A_n(\varepsilon)}\} - E\{g(P^{(2)})1_{A_n(\varepsilon)}\}|$$
$$\leqslant |W_g(p,\gamma)g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p))|n^{-1} + \mathrm{o}(n^{-1}) \tag{4.6}$$

as $n \to \infty$. For $g(p) = p$ we get $|W_g(p,\gamma)g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p))| \leqslant 14.4p$ for $-0.5 \leqslant \gamma \leqslant 1, 0.001 \leqslant p \leqslant 0.01$ and hence a maximal relative error equal to $14.4n^{-1} + \mathrm{o}(n^{-1})$ as $n \to \infty$.

If g is bounded, then $E\{g(P^{(j)})1_{A_n^c(\varepsilon)}\} = \mathrm{o}(n^{-1})$ as $n \to \infty$ for $j = 1,2$ and hence $1_{A_n(\varepsilon)}$ can be deleted in (4.5) and (4.6).

Proof. Let $X_1,\dots,X_n,X_{n+1}$ be i.i.d. r.v.'s with a normal power distribution with parameter γ. Let g be three times continuously differentiable at p and let $\varepsilon > 0$ be sufficiently small. It is easily seen that both $Ec_{u1}(\hat{\gamma})$ and $Ec_{u2}(\hat{\gamma})$ are $\mathrm{O}(n^{-1})$ as $n \to \infty$. By Taylor's expansion, noting that $P(A_n^c(\varepsilon)) = \mathrm{o}(n^{-1})$, we get for $j = 1,2$

$$E\{g(P^{(j)})1_{A_n(\varepsilon)}\}$$
$$= E\{g(\bar{K}_\gamma(\bar{K}_\gamma^{-1}(p) + V1_{A_n(\varepsilon)} + \hat{\sigma}c_{uj}(\hat{\gamma})))1_{A_n(\varepsilon)}\}$$
$$= g(p) - E\{V1_{A_n(\varepsilon)} + \hat{\sigma}c_{uj}(\hat{\gamma})\}g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p))$$
$$+ \frac{1}{2}\,E\{V1_{A_n(\varepsilon)} + \hat{\sigma}c_{uj}(\hat{\gamma})\}^2\{g''(p)k_\gamma^2(\bar{K}_\gamma^{-1}(p)) - g'(p)k'_\gamma(\bar{K}_\gamma^{-1}(p))\}$$

$$
\begin{aligned}
&+\mathrm{o}(n^{-1})\\
&= g(p) - E\{V1_{A_n(\varepsilon)} + c_{uj}(\hat{\gamma})\}g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p))\\
&\quad + \frac{1}{2}EV^2 1_{A_n(\varepsilon)}\{g''(p)k_\gamma^2(\bar{K}_\gamma^{-1}(p)) - g'(p)k'_\gamma(\bar{K}_\gamma^{-1}(p))\} + \mathrm{o}(n^{-1})\\
&= g(p) - \{B1_n(\gamma) + Ec_{u1}(\hat{\gamma})\}g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p))\\
&\quad + \frac{1}{2}B2_n(\gamma)\{g''(p)k_\gamma^2(\bar{K}_\gamma^{-1}(p)) - g'(p)k'_\gamma(\bar{K}_\gamma^{-1}(p))\}\\
&\quad + \{Ec_{u1}(\hat{\gamma}) - Ec_{uj}(\hat{\gamma})\}g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p)) + \mathrm{o}(n^{-1})\\
&= g(p) + \{Ec_{u1}(\hat{\gamma}) - Ec_{uj}(\hat{\gamma})\}g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p)) + \mathrm{o}(n^{-1})
\end{aligned}
$$

as $n \to \infty$. This completes the proof of (4.5) and (4.6). Numerical calculations show that $|W_g(p,\gamma)g'(p)k_\gamma(\bar{K}_\gamma^{-1}(p))| \leqslant 14.4p$ for $-0.5 \leqslant \gamma \leqslant 1, 0.001 \leqslant p \leqslant 0.01$ when $g(p) = p$. If g is bounded, then $E\{g(P^{(j)})1_{A_n^c(\varepsilon)}\} = \mathrm{O}(P(A_n^c(\varepsilon))) = \mathrm{o}(n^{-1})$. □

5. Simulation and other numerical results

A simulation study is performed to see to what kind of improvement the several steps lead: firstly, extending the restricted model of normality to the larger model of the normal power family and secondly, the application of the correction terms in the larger model. For comparison we also consider the restricted model with correction terms. In the simulation study we want to cover the restricted model of normality, the normal power family and distributions outside the normal power family (but not too far away from it).

As criterion we take EP and compare this to p. In terms of Sections 3 and 4 this means that we take $g(p) = p$ in the simulation study. Similar results as presented here hold for the other functions of Section 3, cf. (3.1). In the simulations we always choose $p = 0.001$ and for n we take 100, 250 and 500. Our main attention is on $n = 100$, since nowadays large samples are usually not available. The columns with $n = 250, 500$ give an impression of the rate of convergence. The number of repetitions in the simulation study equals 100,000.

In the simulation study we take the following distributions (with $\mu = 0$ and $\sigma = 1$):

Φ: standard normal distribution function;

K_γ: normal power distribution function with $\gamma = -0.5, -0.25, 0 (= \Phi), 0.25, 0.5, 0.75, 1$;

T: standardized Student distribution function with 6 degrees of freedom;

RM: random mixture: $\frac{1}{2}\Phi + \frac{1}{2}T$;

DM: deterministic mixture, given by: $DM^{-1} = c^*\{\Phi^{-1} + T^{-1}\}$ with c^* a constant to ensure unit variance;

TU: Tukey's λ-family, based on a r.v. $c(\lambda)\{U^\lambda - (1-U)^\lambda\}$ with U a uniform r.v. on $(0,1)$ and $c(\lambda)$ a constant to ensure unit variance, with $\lambda = -0.1, \lambda = 0$ (standardized logistic distribution), $\lambda = 0.14$ (very close to the standard normal (outside the tails!));

O: Orthonormal family with $k=3$ and $(\gamma_1,\gamma_2,\gamma_3)=(-0.1,-0.1,0.1)$; the orthonormal family is defined as follows: for $k=1,2,\ldots$ and $j=1,\ldots,k$, let γ_j be a coefficient, π_j be the jth Legendre polynomial on $(0,1)$ and consider the density $f(y,\gamma_1,\ldots,\gamma_k)$ proportional to $\exp\{\sum_{j=1}^k \gamma_j\pi_j(y)\}$; if Y is a r.v. with density f, then take $\Phi^{-1}(Y)$ and standardize that r.v. to have zero mean and variance one.

We also calculate numerically the (restricted) model errors and two approximations, called App1 and App2. The first approximation of EP, App1, is obtained by taking into account convergence up to order o(1). Since the correction terms are of order $\mathrm{O}(n^{-1})$, this means that both for the uncorrected and the corrected control limit App1 equals $\bar{F}_0(u_p)$ with F_0 the distribution function of $\sigma^{-1}(X_{n+1}-\mu)$ when the supposed model is normality and $\bar{F}_0(\bar{K}_\gamma^{-1}(p))$ with γ the suitably chosen point given by (2.7) when the supposed model is the normal power family. Hence the first approximation is related to the (restricted) model error by App1 $= p+(R)ME$, see (2.2) and (2.5).

The second approximation is obtained by expanding EP up to terms of order $\mathrm{O}(n^{-1})$. Note that there are no $\mathrm{O}(n^{-1/2})$-terms. In general the derivation of these formulas is as follows. We get, with V given by (4.1),

$$EP = E\{\bar{F}_0(\bar{K}_\gamma^{-1}(p)+V+\hat{\sigma}c_u)\}.$$

In the uncorrected case we have $c_u=0$. Furthermore, when dealing with the restricted model of normality we take $\bar{K}_\gamma^{-1}(p)=u_p$ and $V=\bar{X}+Su_p-u_p$. Typically we have that EV and EV^2 are of order $\mathrm{O}(n^{-1})$, and $E|V|^k=\mathrm{o}(n^{-1})$ for $k\geqslant 2$. The correction term c_u for correcting the bias typically is of order $\mathrm{O}(n^{-1})$. Let F_0 have density f_0. Taylor expansion gives, writing $\approx$ for ignoring terms of order V^k for $k\geqslant 3$ and terms of order $c_u(\hat{\sigma}-1)$, c_uV and $(c_u)^k$ for $k\geqslant 2$,

$$EP \approx \bar{F}_0(\bar{K}_\gamma^{-1}(p)) - f_0(\bar{K}_\gamma^{-1}(p))E(V+c_u) - \frac{1}{2}EV^2 f_0'(\bar{K}_\gamma^{-1}(p)). \tag{5.1}$$

App2 is obtained from the right-hand side of (5.1) by taking suitable approximations of $E(V+c_u)$ and EV^2. The formulas are given below in the corresponding subsections.

The control limit depends on the supposed model, on the estimators of the parameters and on whether we make a correction or not. We consider the following cases. For convenience the unit in the tables equals 0.001, thus, for instance, 1.25 means 0.00125.

(i) *Supposed model*: *normality*, *no correction.*

Assuming that we are in the restricted model of normality and applying no correction for using estimators, the control limit simply equals

$$\bar{X}+Su_p.$$

For the first approximation of the expected (observed) false alarm rate we get App1 $=\bar{F}_0(u_p)=p+\mathrm{RME}$ with F_0 the distribution function of $\sigma^{-1}(X_{n+1}-\mu)$. The second approximation of EP equals, cf. also (A.1),

$$\mathrm{App2} = \bar{F}_0(u_p) + \frac{(\mu_4-1)f_0(u_p)u_p}{8n} - \frac{f_0'(u_p)\left\{1+\mu_3u_p+\frac{1}{4}(\mu_4-1)u_p^2\right\}}{2n},$$

where μ_k is the kth moment under F_0.

Table 1
Simulated expected (observed) false alarm rate without correction for the estimation of the mean and variance and corresponding first and second approximation assuming normality as model, with in the column RME the restricted model error. The unit in the table is 0.001

F_0	Simulation			App1	RME	App2		
	$n=100$	$n=250$	$n=500$			$n=100$	$n=250$	$n=500$
Φ	1.36	1.14	1.07	1.00	0.00	1.33	1.13	1.07
$K_{-0.5}$	0.00	0.00	0.00	0.00	−1.00	0.00	0.00	0.00
$K_{-0.25}$	0.06	0.03	0.03	0.02	−0.98	0.05	0.03	0.03
$K_{0.25}$	4.44	3.96	3.81	3.66	2.66	4.39	3.95	3.81
$K_{0.5}$	7.70	7.01	6.79	6.58	5.58	7.67	7.02	6.80
$K_{0.75}$	10.31	9.45	9.15	8.86	7.86	10.30	9.44	9.15
K_1	12.13	11.01	10.71	10.35	9.35	12.17	11.08	10.71
T	5.30	4.87	4.72	4.56	3.56	5.37	4.89	4.73
RM	3.31	2.99	2.89	2.78	1.78	3.34	3.00	2.89
DM	3.45	3.13	3.03	2.92	1.92	3.43	3.13	3.02
TU(−0.1)	6.61	6.07	5.89	5.71	4.71	6.66	6.09	5.90
TU(0)	4.29	3.91	3.79	3.67	2.67	4.26	3.90	3.79
TU(0.14)	1.21	0.98	0.91	0.85	−0.15	1.19	0.98	0.91
O	2.81	2.39	2.26	2.13	1.13	2.76	2.38	2.26

The simulation results, the restricted model error and the first and second approximations are presented in Table 1.

It is seen from Table 1 that the false alarm rate may be completely wrong if we are not in the restricted model, that is if normality does not hold. This confirms previous results of e.g. Chan et al. (1988) and Pappanastos and Adams (1996). They only give the total error. By splitting up the total error in the restricted model error and the restricted stochastic error more insight is obtained about the roles of misspecification on the one hand and estimation of the location and scale parameter on the other hand.

The restricted model error RME is the difference between App1 and 0.001 ($=p$). It is illuminating that, while in the middle of the distribution there is hardly any difference between TU(0.14) and the standard normal distribution, the far tails are different. The effect of this difference is clearly seen by comparing TU(0.14) and Φ in Table 1. Often it is also stated that in the middle there is not much difference between normal and logistic distributions. Comparison of TU(0) and Φ in Table 1 shows that for the problem at hand sloppy inspection of the data in the middle, leading to the conclusion that “normality is not so bad” may produce serious errors.

Apart from the misspecification the effect of the estimation is also not negligible. Due to the estimation the total error becomes substantially larger in case of positive RME’s, which may occur in practice more often than negative RME’s. In the latter situation the total error is compensated by the (positive) restricted stochastic error. This, for instance, occurs for TU(0.14). However, such compensation is in an uncontrolled way and may be far too small (see $K_{-0.25}$) or may lead to overcompensation (TU(0.14), $n=100$).

Clearly, App1 gives an impression of *EP*, the expected (observed) false alarm rate, but is not very precise. App2 gives a very good prediction of the simulation.

(ii) *Supposed model*: *normality*, *with correction.*

Assuming that we are in the restricted model of normality and applying the suitable correction term for using estimators, the control limit equals

$$\bar{X} + Su_p + \frac{Su_p}{4n}(u_p^2 + 3).$$

The correction term is of order $\mathrm{O}(n^{-1})$ and hence App1 is the same as in Table 1. In view of (5.1), cf. also (A.1), we get

$$\mathrm{App2} = \bar{F}_0(u_p) - f_0(u_p)\frac{u_p}{4n}\left\{u_p^2 + 3 - \frac{1}{2}(\mu_4 - 1)\right\}$$
$$-\frac{1}{2}f_0'(u_p)\frac{1 + \mu_3 u_p + \frac{1}{4}(\mu_4 - 1)u_p^2}{n}.$$

It is seen that indeed $\mathrm{App2} = p$ for $F_0 = \Phi$.

The simulation results, the restricted model error and the first and second approximations are presented in Table 2. Note that RME and App1 are the same as in Table 1.

It is seen from the row denoted by Φ in Table 2 that the correction for estimation of the parameters works very well, cf. also Albers and Kallenberg (2000) for more

Table 2
Simulated expected (observed) false alarm rate with correction for the estimation of the mean and variance and corresponding first and second approximation assuming normality as model, with in the column RME the restricted model error. The unit in the table is 0.001

F_0	Simulation			App1	RME	App2		
	$n=100$	$n=250$	$n=500$			$n=100$	$n=250$	$n=500$
Φ	1.01	1.00	1.00	1.00	0.00	1.00	1.00	1.00
$K_{-0.5}$	0.00	0.00	0.00	0.00	−1.00	0.00	0.00	0.00
$K_{-0.25}$	0.03	0.03	0.02	0.02	−0.98	0.03	0.03	0.02
$K_{0.25}$	3.68	3.67	3.66	3.66	2.66	3.65	3.66	3.66
$K_{0.5}$	6.72	6.63	6.60	6.58	5.58	6.71	6.63	6.61
$K_{0.75}$	9.27	9.03	8.94	8.86	7.86	9.28	9.03	8.95
K_1	11.10	10.67	10.51	10.35	9.35	11.17	10.68	10.51
T	4.64	4.61	4.59	4.56	3.56	4.73	4.63	4.60
RM	2.80	2.80	2.79	2.78	1.78	2.86	2.81	2.80
DM	2.89	2.91	2.93	2.92	1.92	2.90	2.91	2.92
TU(−0.1)	5.86	5.78	5.75	5.71	4.71	5.94	5.80	5.76
TU(0)	3.63	3.65	3.66	3.67	2.67	3.62	3.65	3.66
TU(0.14)	0.87	0.85	0.85	0.85	−0.15	0.87	0.85	0.85
O	2.18	2.15	2.14	2.13	1.13	2.16	2.14	2.14

details. Comparison of Tables 1 and 2 shows that the correction reduces the stochastic error not only for the normal distribution, but as well for the other distributions, thus bringing the total error closer to the restricted model error. Nevertheless, if normality fails still the false alarm rate may be completely wrong due to the restricted model error. As a consequence of the bias correction, App1 and App2 are much closer to each other. Again, App2 gives a very good prediction of the simulation.

(iii) *Supposed model*: *normal power family*, *no correction.*

Assume that our observations are from the normal power family and that we use the estimator $\hat{\gamma}$ based on quantiles in the ordinary tail. Then the control limit without correcting for the estimation equals

$$\bar{X} + S\bar{K}_{\hat{\gamma}}^{-1}(p).$$

The first approximation of the false alarm rate is given by $\text{App1} = \bar{F}_0(\bar{K}_{\gamma}^{-1}(p))$ with γ the suitably chosen point given by

$$\gamma = 1.1218 \log\left(\frac{\bar{F}_0^{-1}(0.05)}{\bar{F}_0^{-1}(0.25)}\right) - 1. \tag{5.2}$$

It is related to the model error by $\bar{F}_0(\bar{K}_{\gamma}^{-1}(p)) = p + \text{ME}$, cf. (2.5). Inserting $c_u = 0$ in (5.1) the second approximation of EP equals

$$\text{App2} = \bar{F}_0(\bar{K}_{\gamma}^{-1}(p)) - f_0(\bar{K}_{\gamma}^{-1}(p))\widetilde{B1}_n(\gamma) - \frac{1}{2}\widetilde{B2}_n(\gamma) f_0'(\bar{K}_{\gamma}^{-1}(p))$$

with $\widetilde{B1}_n(\gamma)$ and $\widetilde{B2}_n(\gamma)$ given by (A.10). The involved moments μ_i should be taken under F_0.

The simulation results, the model error and the first and second approximations are presented in Table 3.

Compared to Tables 1 and 2 it is seen from Table 3 that the difference between the expected (observed) false alarm rate and the required value 0.001 is seriously reduced by considering the normal power family as parametric model. Moreover, when the restricted model of normality holds, the loss is not large. The second approximation gives a very good prediction of the simulation throughout the table. This makes it rather easy to analyze and predict the behavior of the control chart under all kinds of distributions.

(iv) *Supposed model*: *normal power family*, *with correction.*

The control limit in this case equals

$$\bar{X} + S\left\{\bar{K}_{\hat{\gamma}}^{-1}(p) - C1(\hat{\gamma})C2(\hat{\gamma}) - \frac{C3(\hat{\gamma})}{n} + \frac{C4(\hat{\gamma})}{n}\right\}.$$

Table 3
Simulated expected (observed) false alarm rate using $\hat{\gamma}$ as estimator without correction for the estimation of the mean, the variance and the parameter γ and corresponding first and second order approximation assuming the normal power family as model, with in the column ME the model error. The unit in the table is 0.001

F_0	Simulation			App1	ME	App2		
	$n=100$	$n=250$	$n=500$			$n=100$	$n=250$	$n=500$
Φ	2.53	1.59	1.27	1.00	0.00	2.26	1.52	1.25
$K_{-0.5}$	3.41	1.92	1.42	1.00	0.00	3.15	1.89	1.43
$K_{-0.25}$	2.73	1.66	1.30	1.00	0.00	2.50	1.63	1.30
$K_{0.25}$	2.39	1.54	1.25	1.00	0.00	2.15	1.48	1.23
$K_{0.5}$	2.34	1.51	1.23	1.00	0.00	2.11	1.46	1.22
$K_{0.75}$	2.33	1.51	1.23	1.00	0.00	2.12	1.46	1.22
K_1	2.34	1.51	1.23	1.00	0.00	2.18	1.48	1.24
T	4.68	3.78	3.39	3.08	2.08	4.56	3.73	3.38
RM	3.63	2.76	2.43	2.16	1.16	3.44	2.71	2.42
DM	3.78	2.90	2.57	2.28	1.28	3.61	2.85	2.55
TU(−0.1)	4.92	3.97	3.58	3.25	2.25	4.79	3.93	3.56
TU(0)	3.96	3.00	2.63	2.33	1.33	3.75	2.94	2.61
TU(0.14)	2.33	1.41	1.08	0.81	−0.19	2.11	1.35	1.07
O	2.18	1.23	0.94	0.69	−0.31	1.82	1.15	0.91

The correction term is of order $\mathrm{O}(n^{-1})$ and hence App1 is the same as in Table 3. The second approximation equals, with γ the suitably chosen point given by (5.2),

$$\mathrm{App2} = \bar{F}_0(\bar{K}_\gamma^{-1}(p)) - f_0(\bar{K}_\gamma^{-1}(p))\left\{\widetilde{B1}_n(\gamma) - C1(\gamma)C2(\gamma) - \frac{C3(\gamma)}{n} + \frac{C4(\gamma)}{n}\right\}$$
$$-\frac{1}{2}\widetilde{B2}_n(\gamma)f_0'(\bar{K}_\gamma^{-1}(p))$$

with $\widetilde{B1}_n(\gamma)$ and $\widetilde{B2}_n(\gamma)$ given by (A.10). The involved moments μ_i should be taken under F_0.

The simulation results, the model error and the first and second approximations are presented in Table 4.

Table 4 shows that the correction for estimation of the parameters works very well, leading to EP (very) close to p for distributions from the normal power family. Comparison with Table 2 shows that the corrected control limit based on $\hat{\gamma}$ performs much better than the corrected standard control limit based on normality. Both the restricted model error and the restricted stochastic error are reduced to an acceptable level by the recommended control limit. App2 gives a very good prediction of the simulation, thus making it easy to analyze and predict the behavior of this control chart for all kind of distributions.

The recommended control limit gives a simple explicit control limit with small total error in the normal power family and reasonable total error for distributions outside this family.

Table 4
Simulated expected (observed) false alarm rate using $\hat{\gamma}$ as estimator and the recommended version of the control limit with correction for the estimation of the mean, the variance and the parameter γ and corresponding first and second approximation assuming the normal power family as model, with in the column ME the model error. The unit in the table is 0.001

F_0	Simulation			App1	ME	App2		
	$n=100$	$n=250$	$n=500$			$n=100$	$n=250$	$n=500$
Φ	1.14	1.05	1.02	1.00	0.00	1.06	1.02	1.01
$K_{-0.5}$	0.77	0.93	0.97	1.00	0.00	1.06	1.02	1.01
$K_{-0.25}$	1.02	1.01	1.01	1.00	0.00	1.08	1.03	1.02
$K_{0.25}$	1.16	1.05	1.02	1.00	0.00	1.02	1.01	1.00
$K_{0.5}$	1.18	1.05	1.01	1.00	0.00	1.01	1.00	1.00
$K_{0.75}$	1.20	1.05	1.01	1.00	0.00	1.03	1.01	1.00
K_1	1.21	1.06	1.01	1.00	0.00	1.08	1.03	1.02
T	2.91	3.04	3.03	3.08	2.08	2.80	2.99	3.03
RM	2.09	2.14	2.14	2.16	1.16	2.02	2.11	2.13
DM	2.13	2.21	2.24	2.28	1.28	1.99	2.18	2.22
TU(−0.1)	3.12	3.21	3.21	3.25	2.25	2.96	3.17	3.19
TU(0)	2.24	2.28	2.28	2.33	1.33	2.02	2.22	2.26
TU(0.14)	1.00	0.90	0.85	0.81	−0.19	0.98	0.88	0.84
O	0.95	0.78	0.73	0.69	−0.31	0.86	0.75	0.72

6. Discussion

We start the discussion with answering the questions posed in Section 2.

(i) *Normality holds*

How large is the difference between SE and RSE? Can RSE and SE be reduced by (simple) corrections? How large are the corrected SE and the corrected RSE?

The simulated expected RSE when applying the uncorrected control limit and its (first and second) approximations are found in Table 1 on the row containing Φ by subtracting $p = 0.001$ (that is 1 unit) from the corresponding entry. Similarly, again applying the uncorrected control limit the simulated expected SE and its approximations are found in Table 3 on the row containing Φ, again subtracting 1 unit. It is seen that without correction indeed SE is larger than RSE. However, by making the appropriate corrections the differences disappear: all the corrected versions are close to 0.

We may conclude that the larger SE in the uncorrected control limit can be accurately repaired by an appropriate correction term.

(ii) *Observations from the normal power family*

How bad can RME be? How does this balance with the larger SE? Can RSE and SE be reduced by (simple) corrections? How large are the corrected RSE and the corrected SE, when the observations are from the normal power family?

RME is discussed in Section 2 and furthermore numerically calculated in Tables 1 and 2. In Albers et al. (2002) a more extensive discussion is presented. The conclusion is clear: RME can be quite large.

In the situation where no correction takes place we consider the balance between the larger SE and the possibly large RME. Therefore, we compare the results of Table 3 with those of Table 1 for the distributions $K_{-0.5}, K_{-0.25}, K_{0.25}, K_{0.5}, K_{0.75}$ and K_1. It is seen that RME has a larger effect than SE. Moreover, SE can be corrected, while RME remains.

RSE refers to the situation where the supposed model is normality. Therefore, the corresponding correction terms are based on this assumption. The correction is not tailored to the normal power family. How well the correction still helps in these kinds of situations is seen by comparing the lines containing $K_{-0.5}, K_{-0.25}, K_{0.25}, K_{0.5}, K_{0.75}$ and K_1 in Tables 2 and 1. RSE reduces also for these distributions, but due to the large RME the total error is still (very) large.

The corrected SE is quite small, see the lines containing $K_{-0.5}, K_{-0.25}, K_{0.25}, K_{0.5}, K_{0.75}$ and K_1 in Table 4.

(iii) *Outside the normal power family*

How large are the uncorrected and corrected SE and the uncorrected and corrected RSE, when we are outside the normal power family and ME is not too big? How is the total error when applying the corrected control limit with as supposed model the normal power family and how does this compare with the total error when applying the corrected control limit with as supposed model normality?

From the lines containing T, RM, DM, $\mathrm{TU}(-0.1), \mathrm{TU}(0), \mathrm{TU}(0.14), O$ in Tables 3 and 4 it is seen that the uncorrected SE is substantially improved by application of the correction terms (even although they are derived for the parametric model). Also when supposing normality the correction is very helpful, see the lines containing T, RM, DM, $\mathrm{TU}(-0.1)$, $\mathrm{TU}(0), O$ in Tables 1 and 2.

With respect to the total error the lines containing T, RM, DM, $\mathrm{TU}(-0.1)$, $\mathrm{TU}(0), \mathrm{TU}(0.14)$, O in Tables 1–4 show the following. The total error is often very large when using the classical normal control limit, due to a high restricted model error (and a restricted stochastic error, when no correction is applied). As a first step the total error is considerably reduced by application of the normal power family with, secondly, a further great improvement by taking the correction terms into account.

(iv) *Out-of-control*

What is the impact on the out-of-control behavior of the chart? Does the process stop at a reasonable time when it has gone out-of-control?

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} come from a shifted distribution function $F(x-\Delta)$. For simplicity, and without loss of generality, we again let $\mu=0$ and $\sigma=1$ and thus work under the standardized F_0. The shift Δ

Table 5
Simulated values of $E_{\Delta}P$, the expected (observed) probability of producing an out-of-control signal under $F_0(x-\Delta)$, using $\hat{\gamma}$ as estimator and the recommended version of the control limit with correction for the estimation of the mean, the variance and the parameter γ, with in the column $\tilde{p}$, the probability of producing an out-of-control signal when the distribution is completely known

F_0	Δ	$\tilde{p}$	n			F_0	Δ	$\tilde{p}$	n		
			100	250	500				100	250	500
Φ	2	0.138	0.102	0.120	0.128	T	2	0.016	0.068	0.072	0.071
	3	0.464	0.346	0.410	0.436		3	0.088	0.254	0.297	0.311
$K_{-0.5}$	1	0.227	0.144	0.190	0.208	RM	2	0.040	0.081	0.091	0.095
	2	0.500	0.477	0.496	0.499		3	0.217	0.293	0.349	0.372
$K_{-0.25}$	1	0.058	0.040	0.049	0.053	DM	2	0.045	0.082	0.092	0.097
	2	0.355	0.261	0.314	0.335		3	0.234	0.295	0.353	0.377
$K_{0.25}$	2	0.049	0.043	0.046	0.047	TU(−0.1)	2	0.012	0.055	0.055	0.054
	3	0.210	0.168	0.190	0.199		3	0.052	0.213	0.237	0.244
$K_{0.5}$	3	0.082	0.083	0.080	0.080	TU(0)	2	0.036	0.070	0.076	0.078
	4	0.300	0.257	0.292	0.305		3	0.188	0.259	0.306	0.325
$K_{0.75}$	3	0.036	0.046	0.038	0.036	TU(0.14)	2	0.148	0.099	0.118	0.126
	4	0.119	0.151	0.141	0.129		3	0.481	0.339	0.406	0.431
K_1	3	0.018	0.029	0.021	0.019	O	2	0.097	0.062	0.071	0.076
	4	0.053	0.097	0.072	0.059		3	0.340	0.224	0.263	0.280

is such that the probability of detecting the shift when the distribution is completely known, that is

$$\tilde{p} = \bar{F}_0(\bar{F}_0^{-1}(p) - \Delta),$$

is no longer extremely small, like p.

The out-of-control behavior is seen from the results in Table 5. For the same distributions as used in Tables 1–4 we have performed simulations under $F_0(x-\Delta)$. For each F_0 we have selected two values of Δ such that reasonable values of $\tilde{p}$ result. 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.

It is seen from Table 5 that repairing the damage during in control does not destroy the out-of-control behavior. Even compared to the situation of exactly knowing the distribution, which rarely occurs in practice, the recommended control chart does not loose that much when considering a shifted distribution. For instance, when normality holds and a shift of 2 standard deviations occurs, the probability of detecting such a shift goes from 0.138 to 0.101 ($n = 100$), 0.120 ($n = 250$) or 0.128 ($n = 500$). With these probabilities correspond average run lengths going from 7.3 to 9.9 ($n = 100$), 8.3 ($n = 250$), or 7.8 ($n = 500$), which indeed is not a tremendous change. For distributions outside the parametric model there may be even a gain w.r.t. the case of completely known distribution, which is due to the model error, see Table 4.

We conclude that the classical normal control limit is unreliable with often a very large total error, mainly due to the large restricted model error. Although the correction for estimating the mean and the variance is very useful in reducing the restricted stochastic error (even when normality does not hold), the model error often dominates and causes a large total error.

The (corrected) control limits based on the normal power family work very well, also outside the normal power family. The recommended one is completely explicit and can be calculated quite straightforwardly. It has a reliable approximation (App2), which makes it possible to predict in an accurate way the behavior of this control chart under all kind of distributions. The out-of-control behavior is not dropping off dramatically; the process will stop at a reasonable time when it has gone out-of-control.

Appendix A

In this appendix we present the derivation and formulas of the correction terms based on the normal power family. Note that we take $\mu = 0$ and $\sigma = 1$. In order to calculate approximations also outside the normal power family we present all formulas under F_0. For the derivation of the correction terms and the proof of Theorem 2 we only need these formulas under $K\gamma$. In that case the error terms are $o(n^{-1})$. To express this we use the notation $\doteq$. For instance,

$$E(\hat{\sigma} - 1) \doteq -\frac{\mu_4 - 1}{8n}$$

means that $E(\hat{\sigma} - 1) = -(\mu_4 - 1)/(8n) + o(n^{-1})$, when the observations are from the normal power family.

The following formulas give approximations for the moments of the estimators $\hat{\mu}, \hat{\sigma}$ and $\hat{\gamma}^*$ assuming that $X_1, \ldots, X_n$ are i.i.d. r.v.'s with distribution function F_0, thus having expectation 0 and variance 1. Writing $\mu_i = EX_1^i$, $q = 0.05, r = 0.25, x_q = F_0^{-1}(1-q)$, $q_n = 1 - [n + 1 - qn]/(n+1), x_r = F_0^{-1}(1-r), r_n = 1 - [n + 1 - rn]/(n+1)$ we have

$$E\hat{\mu} = 0, \quad E(\hat{\sigma} - 1) \doteq -\frac{\mu_4 - 1}{8n},$$

$$E\hat{\mu}^2 = \frac{1}{n}, \quad E(\hat{\sigma} - 1)^2 \doteq \frac{\mu_4 - 1}{4n}, \quad E\hat{\mu}(\hat{\sigma} - 1) \doteq \frac{\mu_3}{2n}, \tag{A.1}$$

$$E(\hat{\gamma}^* - \gamma^*) \doteq \frac{F_0^{-1}(1-q_n)}{F_0^{-1}(1-r_n)} - \frac{F_0^{-1}(1-q)}{F_0^{-1}(1-r)} - \frac{q(1-q)}{2nx_r}\frac{f_0'}{f_0^3}(x_q) + \frac{x_q r(1-r)}{2nx_r^2}\frac{f_0'}{f_0^3}(x_r)$$

$$- \frac{1}{nx_r^2}\left\{\frac{q(1-r)}{f_0(x_q)f_0(x_r)} - \frac{\int_{x_q}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_q)} - \frac{\int_{x_r}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_r)} + 1\right\}$$

$$+ \frac{x_q}{nx_r^3}\left\{\frac{r(1-r)}{[f_0(x_r)]^2} - 2\,\frac{\int_{x_r}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_r)} + 1\right\}, \qquad \text{(A.2)}$$

$$E(\hat{\sigma}-1)(\hat{\gamma}_2^* - \gamma_2^*) \doteq \frac{1}{2nx_r}\left\{\frac{\int_{x_q}^{\infty}(y^2-1)f_0(y)\,\mathrm{d}y}{f_0(x_q)} - \mu_3\right\}$$

$$- \frac{x_q}{2nx_r^3}\left\{\frac{\int_{x_r}^{\infty}(y^2-1)f_0(y)\,\mathrm{d}y}{f_0(x_r)} - \mu_3\right\}, \qquad \text{(A.3)}$$

$$E(\hat{\gamma}^* - \gamma^*)^2$$

$$\doteq \frac{1}{nx_r^2}\left\{\frac{q(1-q)}{[f_0(x_q)]^2} - 2\,\frac{\int_{x_q}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_q)} + 1\right\}$$

$$+ \frac{x_q^2}{nx_r^4}\left\{\frac{r(1-r)}{[f_0(x_r)]^2} - 2\,\frac{\int_{x_r}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_r)} + 1\right\}$$

$$- \frac{2x_q}{nx_r^3}\left\{\frac{q(1-r)}{f_0(x_q)f_0(x_r)} - \frac{\int_{x_q}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_q)} - \frac{\int_{x_r}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_r)} + 1\right\}, \qquad \text{(A.4)}$$

$$E\hat{\mu}(\hat{\gamma}^* - \gamma^*) \doteq \frac{1}{nx_r}\left\{\frac{\int_{x_q}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_q)} - 1 - \frac{x_q}{x_r}\left[\frac{\int_{x_r}^{\infty} yf_0(y)\,\mathrm{d}y}{f_0(x_r)} - 1\right]\right\}. \qquad \text{(A.5)}$$

For the normal power family we have $\mu_i = 0$ for odd i and

$$\mu_4 = EZ_\gamma^4 = \frac{\sqrt{\pi}\,\Gamma(2\gamma + \frac{5}{2})}{\Gamma(\gamma + \frac{3}{2})^2}.$$

Furthermore, in case of the normal power family (that is $F_0 = K_\gamma$, $f_0 = k_\gamma$ and $f_0' = k_\gamma'$) we write for short $C_{21}(\gamma)$ for the right-hand side of (A.2), $C_{22}(\gamma)$ for the right-hand side of (A.3), $C_{23}(\gamma)$ for the right-hand side of (A.4) and $C_{24}(\gamma)$ for the right-hand side of (A.5).

Since $V = \hat{\mu} + \hat{\sigma}\bar{K}_{\hat{\gamma}}^{-1}(p) - \bar{K}_\gamma^{-1}(p)$ and $\hat{\gamma} = h^{-1}(\hat{\gamma}^*)$, Taylor expansion gives

$$EV1_{A_n(\varepsilon)} \doteq E\hat{\mu}1_{A_n(\varepsilon)} + E\hat{\sigma}1_{A_n(\varepsilon)}\bar{K}_\gamma^{-1}(p) - \bar{K}_\gamma^{-1}(p) + E\hat{\sigma}(\hat{\gamma}^* - \gamma^*)1_{A_n(\varepsilon)}$$

$$\times \frac{\partial \bar{K}_{h^{-1}(\gamma^*)}^{-1}(p)}{\partial \gamma^*} + \frac{1}{2}E\hat{\sigma}(\hat{\gamma}^* - \gamma^*)^2 1_{A_n(\varepsilon)}\frac{\partial^2 \bar{K}_{h^{-1}(\gamma^*)}^{-1}(p)}{\partial \gamma^{*2}}. \qquad \text{(A.6)}$$

Let

$$B1_n(\gamma) = -\frac{\mu_4 - 1}{8n}\,\bar{K}_\gamma^{-1}(p) + \{C_{21}(\gamma) + C_{22}(\gamma)\}\,\frac{\partial \bar{K}^{-1}_{h^{-1}(\gamma^*)}(p)}{\partial \gamma^*} + \frac{1}{2}\,C_{23}(\gamma)\,\frac{\partial^2 \bar{K}^{-1}_{h^{-1}(\gamma^*)}(p)}{\partial \gamma^{*2}}, \tag{A.7}$$

then combination of (A.6), (A.1)–(A.4) yields

$$EV1_{A_n(\varepsilon)} \doteq B1_n(\gamma).$$

Similarly, we have

$$\begin{aligned} EV^2 1_{A_n(\varepsilon)} \doteq{} & E\hat{\mu}^2 1_{A_n(\varepsilon)} + E(\hat{\sigma}1_{A_n(\varepsilon)} - 1)^2\{\bar{K}_\gamma^{-1}(p)\}^2 \\ & + E(\hat{\gamma}^* - \gamma^*)^2 1_{A_n(\varepsilon)} \left\{ \frac{\partial \bar{K}^{-1}_{h^{-1}(\gamma^*)}(p)}{\partial \gamma^*} \right\}^2 \\ & + 2E\hat{\mu}(\hat{\sigma} - 1)1_{A_n(\varepsilon)}\bar{K}_\gamma^{-1}(p) + 2E\hat{\mu}(\hat{\gamma}^* - \gamma^*)1_{A_n(\varepsilon)}\,\frac{\partial \bar{K}^{-1}_{h^{-1}(\gamma^*)}(p)}{\partial \gamma^*} \\ & + 2E(\hat{\sigma} - 1)(\hat{\gamma}^* - \gamma^*)1_{A_n(\varepsilon)}\bar{K}_\gamma^{-1}(p)\,\frac{\partial \bar{K}^{-1}_{h^{-1}(\gamma^*)}(p)}{\partial \gamma^*} \end{aligned} \tag{A.8}$$

and hence

$$EV^2 1_{A_n(\varepsilon)} \doteq B2_n(\gamma)$$

with

$$B2_n(\gamma) = \frac{1}{n} + \frac{\mu_4 - 1}{4n}\,\{\bar{K}_\gamma^{-1}(p)\}^2 + C_{23}(\gamma)\left\{ \frac{\partial \bar{K}^{-1}_{h^{-1}(\gamma^*)}(p)}{\partial \gamma^*} \right\}^2 + 2\{C_{24}(\gamma) + C_{22}(\gamma)\bar{K}_\gamma^{-1}(p)\}\,\frac{\partial \bar{K}^{-1}_{h^{-1}(\gamma^*)}(p)}{\partial \gamma^*}. \tag{A.9}$$

Noting that

$$\frac{\bar{K}_\gamma^{-1}(1 - q_n)}{\bar{K}_\gamma^{-1}(1 - r_n)} - \frac{\bar{K}_\gamma^{-1}(1 - q)}{\bar{K}_\gamma^{-1}(1 - r)} = \mathrm{O}(n^{-1})$$

it is immediately seen that $B1_n(\gamma) = \mathrm{O}(n^{-1})$ and $B2_n(\gamma) = \mathrm{O}(n^{-1})$. By definition of $c_{u1}(\hat{\gamma})$ and $c_{u2}(\hat{\gamma})$ it follows that

$$Ec_{u1}(\hat{\gamma}) \doteq -B1_n(\gamma) + \frac{1}{2}\,B2_n(\gamma)\left[\frac{g''(p)}{g'(p)}\,k_\gamma(\bar{K}_\gamma^{-1}(p)) - \frac{k'_\gamma}{k_\gamma}\,(\bar{K}_\gamma^{-1}(p))\right]$$

and

$$Ec_{u2}(\hat{\gamma}) \doteq -C1(\gamma)C2(\gamma) - \frac{C3(\gamma)}{n} + \lambda\,\frac{C4(\gamma)}{n}.$$

Remark A.1. To represent the restricted model of normality take $\hat{\gamma}^*$ in the control limit $\hat{\mu}+\hat{\sigma}\{\bar{K}_{\hat{\gamma}}^{-1}(p)+c_{u1}(\hat{\gamma})\}$ identically equal to $\gamma^*=u_{0.05}/u_{0.25}$ and hence $h_1^{-1}(\hat{\gamma}^*)$ identically equal to 0, according to the fact that we do not need to estimate γ. Moreover, by the same reason disregard the contribution of $\partial\bar{K}_{h^{-1}(\gamma^*)}^{-1}(p)/\partial\gamma^*$ and $\partial^2\bar{K}_{h^{-1}(\gamma^*)}^{-1}(p)/\partial\gamma^{*2}$ in $B1_n(\gamma)$ and $B2_n(\gamma)$. Noting that $\mu_4=3$ in case of normality, we replace $B1_n(\gamma)$ by $-u_p/(4n)$ and $B2_n(\gamma)$ by $n^{-1}+u_p^2/(2n)$. Since

$$\frac{g''(p)}{g'(p)}\,k_0(\bar{K}_0^{-1}(p))-\frac{k_0'}{k_0}\,(\bar{K}_0^{-1}(p))$$

is approximately equal to λu_p the control limit $\hat{\mu}+\hat{\sigma}\{\bar{K}_{\hat{\gamma}}^{-1}(p)+c_{u1}(\hat{\gamma})\}$ reduces to $\hat{\mu}+\hat{\sigma}u_p+(4n)^{-1}\hat{\sigma}u_p\{1+\lambda(u_p^2+2)\}$, which, (as it should be!) is the correction term for the control limit based on normality, cf. Albers and Kallenberg (2000). (Note that here we use $\hat{\sigma}=S$ and not its bias-corrected version, explaining here the "extra" term $1/(4n)$ compared to (3.6) in Albers and Kallenberg (2000).)

Remark A.2. As $\partial\bar{K}_{h^{-1}(\gamma^*)}^{-1}(p)/\partial\gamma^*$ and $\partial^2\bar{K}_{h^{-1}(\gamma^*)}^{-1}(p)/\partial\gamma^{*2}$ are difficult to calculate we replace them by numerical approximations A and B, respectively, cf. Albers et al. (2002) for more details. They are given by

$$A=\frac{\log\gamma^*}{\gamma^*}\left\{\frac{0.7698u_p^2+2.9854\sqrt{u_p}-4u_p}{(\log(u_{0.05}/u_{0.25}))^2}\right\}$$

$$+\frac{1}{\gamma^*}\left\{\frac{-0.5773u_p^2-4.4781\sqrt{u_p}+5u_p}{\log(u_{0.05}/u_{0.25})}\right\}$$

and

$$B=\frac{\log\gamma^*}{(\gamma^*)^2}\left\{\frac{-0.3849u_p^2-1.4927\sqrt{u_p}+2u_p}{(\log(u_{0.05}/u_{0.25}))^2}\right\}$$

$$+\frac{1}{(\gamma^*)^2}\left\{\frac{0.3849u_p^2+1.4927\sqrt{u_p}-2u_p}{(\log(u_{0.05}/u_{0.25}))^2}+\frac{0.2887u_p^2+2.2391\sqrt{u_p}-2.5u_p}{\log(u_{0.05}/u_{0.25})}\right\}.$$

Inserting these numerical approximations in $B1_n(\gamma)$ and $B2_n(\gamma)$ leads to the approximation $\widetilde{B1}_n(\gamma)$ of EV and $\widetilde{B2}_n(\gamma)$ of EV^2 given by

$$\widetilde{B1}_n(\gamma)=-\frac{\mu_4-1}{8n}\,\bar{K}_\gamma^{-1}(p)+\{C_{21}(\gamma)+C_{22}(\gamma)\}A+\frac{1}{2}\,C_{23}(\gamma)B,$$

$$\widetilde{B2}_n(\gamma)=\frac{1}{n}+\frac{\mu_4-1}{4n}\,\{\bar{K}_\gamma^{-1}(p)\}^2+C_{23}(\gamma)A^2+2\{C_{24}(\gamma)+C_{22}(\gamma)\bar{K}_\gamma^{-1}(p)\}A. \tag{A.10}$$

References

Albers, W., Kallenberg, W.C.M., 2000. Estimation in Shewhart control charts. Technical Report No. 1559. Faculty of Mathematical Sciences, University of Twente, to appear in Metrika.

Albers, W., Kallenberg, W.C.M., 2001. Are estimated control charts in control? Technical Report No. 1569. Faculty of Mathematical Sciences, University of Twente, to appear in Statistics.

Albers, W., Kallenberg, W.C.M., Nurdiati, S., 2002. Parametric control charts. Technical Report No. 1623. Faculty of Mathematical Sciences, University of Twente.

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

Chan, L.K., Hapuarachchi, K.P., Macpherson, B.D., 1988. Robustness of $\bar{X}$ and R charts. IEEE Trans. Reliability 37, 117–123.

Chen, G., 1997. The mean and standard deviation of the run length distribution of $\bar{X}$ charts when control limits are estimated. Statist. Sinica 7, 789–798.

de Haan, L., Sinha, A.K., 1999. Estimating the probability of a rare event. Ann. Statist. 27, 732–759.

Dekkers, A.L.M., de Haan, L., 1989. On the estimation of the extreme-value index and large quantile estimation. Ann. Statist. 17, 1795–1832.

Ghosh, B.K., Reynolds Jr., M.R., Hui, Y.V., 1981. Shewhart $\bar{X}$-charts with estimated process variance. Comm. Statist. Theory Methods 10, 1797–1822.

Hall, P., Weissman, I., 1997. On the estimation of extreme tail probabilities. Ann. Statist. 25, 1311–1326.

Neduraman, G., Pignatiello Jr., J.J, 2001. On estimating $\bar{X}$ control chart limits. J. Quality Technol. 33, 206–212.

Pappanastos, E.A., Adams, B.M., 1996. Alternative designs of the Hodges-Lehmann control chart. J. Quality Technol. 28, 213–223.

Quesenberry, C.P., 1993. The effect of sample size on estimated limits for $\bar{X}$ and X control charts. J. Quality Technol. 25, 237–247.

Roes, C.B., 1995. Shewhart-type charts in statistical process control. Ph.D. Thesis, University of Amsterdam.

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