Metrika (1994) 41:163–181

Metrika

Setting Test Limits Under Prescribed Consumer Loss

W. ALBERS, W. C. M. KALLENBERG AND G. D. OTTEN

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

Abstract: In inspection of manufactured parts inaccuracies of the measurement process compel to use test limits which are more strict than the given specification limits. To avoid violation of the prescribed bound γ on the consumer loss, in practice overly conservative informal approaches are used leading to a loss of yield. In this paper approximations are derived for which γ is realized with high precision. The expressions involved are sufficiently simple and explicit to allow extension to the more realistic case where the parameters of the underlying distributions need to be estimated. Various criteria are presented to measure the closeness of the resulting stochastic consumer losses to γ. Some attention is devoted to the sampling schemes involved.

Key words: Specification limits, yield, bivariate normal distribution.

1 Introduction

Nowadays a lot of effort is devoted to statistical process control. In this area the emphasis is on reducing variability, in order to increase steadily the fraction of conforming products. However, in many situations the more passive approach of inspection remains indicated. This occurs when rework costs are very low or when relations between process characteristics and final product characteristics are not all that clear, as is often the case for example in semiconductor industry. The ongoing interest for inspection methods is also evident from the literature, see e.g. Easterling et al. (1991).

In the inspection approach, each of a usually very large number of manufactured parts is measured and compared to a given specification limit. Measuring not being infallible, it is customary to accept parts only if they conform to a test limit which is slightly more strict than the specification limit. In this way it is hoped that the consumer loss (*CL*), which is the probability that a product is both nonconforming and accepted, stays below a given bound γ. Typically γ is quite small, e.g. 10–100 ppm (parts per million), especially when multiple measurements are made and a part has to pass each test involved. Clearly, test limits which are too liberal will lead to violation of the bound γ on the *CL*. On the other hand, choices which are too conservative will unnecessarily increase the

0026–1335/94/3–4/163–181 $2.50

producer loss (PL), which is the probability of products being both conforming and rejected. The resulting loss of yield can be quite costly, even if only tenths of percents are involved.

Consequently, it is of considerable importance to set test limits accurately. Here we introduce a way of doing so, based on a simple but adequate approximation to the bivariate normal distribution function. Besides depending on γ and the specification limit, the resulting test limits also depend on the mean and variance of the measured characteristic and those of the measurement error. It turns out that our proposals greatly improve on the informal approaches used in practice. In fact, our second order approximation is for practical purposes indistinguishable from the exact test limit, obtained through numerical evaluation. Of course, the approximation is more easily evaluated than the exact solution, but this may be deemed of secondary importance in this age of widely available computing facilities. Undisputable, however, remains the advantage of analytical tractability: the approximations enable one to study the properties of the procedures involved (cf. Cox & Wermuth (1991), p. 263).

This advantage becomes even more important once we move on to the more realistic case from a practical point of view where the parameters of the underlying distributions are no longer known, and thus need to be estimated. To be able to first appraise the considerable effect of plugging in estimated rather than true values, and then to subsequently correct for this effect, we really need the explicit structure provided through the approximations.

The paper is organized as follows. In section 2 the model is introduced and the various approximations are derived, assuming all parameter values known. A special case of the model has been considered in Albers, Kallenberg and Otten (1992), but no formal results or proofs are given there. In fact, that paper focusses on the presentation of the test limit formulas together with some intuitive explanation. Moreover, it provides guidelines for implementation, a numerical study of the accuracy achieved and a worked-out real example from semiconductor industry. Hence we shall be very brief about these aspects here and concentrate on the mathematical derivations involved. After section 2, we use estimators for the unknown parameters, which leads to stochastic test limits and thus to stochastic CL's. Hence the choice of limit will depend on the criterion used. In section 3 we use $E(\widehat{CL} - \gamma)$, while in section 4 we require that $P(\widehat{CL} > \gamma) = \alpha$ for some given, typically small, α. Finally, in section 5 some attention is devoted to the relaton between the sample sizes chosen for the estimators and the number of parts to be judged on the basis of the test limits obtained in this way.

2 Test Limits for Known Parameters

In the standard inspection situation we observe

$$\tilde{X} = X + U \ , \tag{2.1}$$

where the true measurement X and the measurement error U are independent and normally distributed: $X \sim N(\mu_X, \sigma_X^2)$ and $U \sim N(0, \sigma_U^2)$. Without essential loss of generality we choose a specification interval of the form $(-\infty, s]$, where always $s > \mu_X$. Hence

$$\pi = P(X > s) \tag{2.2}$$

is the probability that a part is nonconforming. To fix ideas, in many applications π will be in the range from 0.001 to 0.15. For a given bound γ on the CL, the test limit t should maximize the yield $P(\tilde{X} < t)$ under the restriction that $CL = P(X > s, \tilde{X} < t) \leq \gamma$. Hence the optimal choice $t^e = t^e(\mu_X, \sigma_X, \sigma_U, s, \gamma)$ can be found exactly by solving

$$P(X > s, \tilde{X} < t^e) = \gamma \ . \tag{2.3}$$

The solution of (2.3) can be obtained numerically, using that $(X, \tilde{X}) \sim N(\mu_X, \mu_X, \sigma_X^2, \sigma_X^2 + \sigma_U^2, \rho)$, with $\rho = (1 + \sigma_U^2/\sigma_X^2)^{-1/2}$. But, as argued in the introduction, we shall prefer approximations to t^e. For a given proposal t, the problem is typically rescaled by introducing

$$a = (s - t)/\sigma_U, \ a^e = (s - t^e)/\sigma_U \ . \tag{2.4}$$

Then $t = s - a\sigma_U$ and the question now is whether a and a^e are sufficiently close. In applications we will typically be interested in nonnegative a. Hence we shall tacitly assume $a > 0$ in the sequel, just like we assumed $s > \mu_X$. Many of the results we are going to present remain true for, or can be adapted to negative a, but it does not seem worthwhile to bother about this. Popular in practice is to use a fixed value for a, like 2 or 3. This clearly gives no control at all over the CL. In passing note that this type of choice makes clear that implicitly it is supposed that in applications of practical interest

$$\sigma_U \ll \sigma_X \ . \tag{2.5}$$

Otherwise, test limits like $s - 3\sigma_U$ would not make much sense and would in the two-sided case lead to the interval $(-s + 3\sigma_U, s - 3\sigma_U)$ being empty.

A slightly better proposal is to use $t^{(0)} = s - a^{(0)}\sigma_U$, where

$$a^{(0)} = \Phi^{-1}(1 - \gamma/\pi) \ , \tag{2.6}$$

in which Φ^{-1} is the inverse of the standard normal distribution function Φ and π is given in (2.2). The consumer loss $CL^{(0)}$ corresponding to using $t^{(0)}$ satisfies

$CL^{(0)} = \pi P(U < s - a^{(0)}\sigma_U - X | X > s) \leq \pi P(U < -a^{(0)}\sigma_U) = \gamma$. Hence $a^{(0)} \geq a^e$ from (2.4), i.e. $t^{(0)}$ is a conservative test limit. However, it is still quite inaccurate, as it rests on a rather crude approximation of the bivariate normal distribution.

Simple but accurate approximations to this distribution can be obtained from e.g. Cox and Wermuth (1991) and Albers and Kallenberg (1991). The latter paper specializes to large ρ. As X and $\tilde{X}$ have $\rho = (1 + \sigma_U^2/\sigma_X^2)^{-1/2}$, this is in view of (2.5) very much the case here. Hence we shall use some results from Albers and Kallenberg (1991). Let

$$\sigma = \frac{\sigma_U}{\sigma_X}, \; Y = -\frac{U}{\sigma_U}, \; \overline{X} = \frac{X - \mu_X}{\sigma_X}, \; \bar{s} = \frac{s - \mu_X}{\sigma_X} \; , \tag{2.7}$$

then $\sigma \ll 1$, $\overline{X}$ and Y are independent and standard normal and $\bar{s} > 0$, while

$$CL = P(X > s, X - \sigma_U Y < t) = P(Y > a, \bar{s} < \overline{X} < \bar{s} + \sigma(Y - a)) \; . \tag{2.8}$$

For $a^{(0)}$ in (2.6) an upper bound $P(Y > a, \bar{s} < \overline{X})$ is used, which in view of (2.8) is crude: as σ from (2.7) is small, only a small subset of the region $\{y > a, \bar{x} > \bar{s}\}$ is needed. We can write, letting $\phi = \Phi'$,

$$CL = \int_a^\infty \{\Phi(\bar{s} + \sigma(y - a)) - \Phi(\bar{s})\}\phi(y)dy \tag{2.9}$$

and expand this integral in powers of σ. In doing so, we encounter the functions

$$g_k(a) = \int_a^\infty (y - a)^k \phi(y)dy \; , \tag{2.10}$$

for $k \geq 0$. It is easily verified that, for $k \geq 1$,

$$g_{k+1}(a) = kg_{k-1}(a) - ag_k(a), \; g_k'(a) = -kg_{k-1}(a) \; , \tag{2.11}$$

$$g_1(a) = \phi(a) - a(1 - \Phi(a)), \; g_2(a) = (a^2 + 1)(1 - \Phi(a)) - a\phi(a) \; . \tag{2.12}$$

As the g_k decrease, we can define

$$a^{(1)} = g_1^{-1}\left(\frac{\gamma}{\sigma\phi(\bar{s})}\right) \; . \tag{2.13}$$

Our first proposal will now be to replace $a^{(0)}$ from (2.6) by $a^{(1)}$. Hence the test limit becomes $t^{(1)} = s - a^{(1)}\sigma_U$. The following lemma gives informaton on the performance of this limit.

Lemma 2.1: For $a^{(1)}$ *from* (2.13) *we have that*

$$0 \leq \frac{\gamma - P(X > s, \tilde{X} < s - a^{(1)}\sigma_U)}{\gamma} \leq \frac{3}{4}\sigma \max(\bar{s}, 1) \ . \tag{2.14}$$

Proof: Observe that

$$\Phi(\bar{s} + \sigma(y - a)) - \Phi(\bar{s}) - \sigma\phi(\bar{s})(y - a) = \tfrac{1}{2}\sigma^2\phi'(\bar{s} + \theta_y\sigma(y - a))(y - a)^2 \ , \tag{2.15}$$

with $0 \leq \theta_y \leq 1$. As for $z \geq \bar{s} \geq 0$ we have that $0 \leq -\phi'(z) \leq -\phi'(\max(\bar{s}, 1))$, it follows from (2.9), (2.15) and (2.10) that

$$\tfrac{1}{2}\sigma^2\phi'(\max(\bar{s}, 1))g_2(a) \leq P(X > s, \tilde{X} < t) - \sigma\phi(\bar{s})g_1(a) \leq 0 \ , \tag{2.16}$$

which in its turn, together with (2.13), leads to

$$\frac{1}{2}\sigma^2\phi'(\max(\bar{s}, 1))g_2\left(g_1^{-1}\left(\frac{\gamma}{\sigma\phi(\bar{s})}\right)\right) \leq P(X > s, \tilde{X} < s - a^{(1)}\sigma_U) - \gamma \leq 0 \ . \tag{2.17}$$

Now according to (2.11), $g_3(a) = 2g_1(a) - ag_2(a)$. As all g_k are nonnegative, it follows that $0 \leq g_3 \leq 2g_1$. Moreover the fact that $z^2 \leq (z + z^3)/2$ for $z \geq 0$, implies in view of (2.10) that $g_2 \leq (g_1 + g_3)/2 \leq 3g_1/2$. Consequently, the left-hand side of (2.17) is bounded from below by $3\sigma\gamma\phi'(\max(\bar{s}, 1))/(4\phi(\bar{s}))$, which is at least $-3\gamma\sigma \max(\bar{s}, 1)/4$. □

It follows from (2.14) that $t^{(1)}$, like $t^{(0)}$, is conservative and that $a^{(1)}$, like $a^{(0)}$, is an upper bound for a^e from (2.4). But it is also evident from (2.14) that the relative error in γ caused by using $a^{(1)}$, is under control. In fact, for values of σ in the range of a few percents, the upper bound in (2.14) is probably already sufficiently small for practical purposes. However, for larger values of σ and/or $\bar{s}$ (e.g. $0.10 \leq \sigma \leq 0.20$, $2 \leq \bar{s} \leq 3$), improvement is certainly called for.

The way to achieve this improvement already suggests itself in the proof of lemma 2.1: from (2.16) we deduce that the first order approximation $\sigma\phi(\bar{s})g_1(a)$

to $P(X > s, \tilde{X} < t)$ can be sharpened by adding the second order term $\frac{1}{2}\sigma^2\phi'(\bar{s})g_2(a)$. In analogy to $a^{(1)}$ from (2.13), we then want to introduce $\tilde{a}$ such that

$$g_1(\tilde{a}) - \frac{\sigma\bar{s}}{2}g_2(\tilde{a}) = \frac{\gamma}{\sigma\phi(\bar{s})} \ . \tag{2.18}$$

Note however, that $\tilde{a}$ is not necessarily defined uniquely by (2.18): the g_k are monotone, but this does not have to hold for the function on the left-hand side of this equation. The following lemma gives a simple necessary and sufficient condition. Introduce the function

$$k = \phi/(1 - \Phi) \ , \tag{2.19}$$

then we formulate:

Lemma 2.2: As $k(x) - x$ *decreases and* $\lim_{x\to-\infty} k(x) - x = \infty$, $\lim_{x\to\infty} k(x) - x = 0$, *there is a unique* x_0 *such that*

$$k(x_0) - x_0 = 1/(\sigma\bar{s}) \ , \tag{2.20}$$

for given σ *and* $\bar{s}$. *The function* $g_1(x) - \frac{\sigma\bar{s}}{2}g_2(x)$ *decreases for* $x > x_0$.

Proof: Combining (2.11), (2.12) and (2.19), we see that $k(x) - x = -g_1(x)/g_1'(x)$. Hence the derivative of $k(x) - x$ equals $h/(g_1')^2$, where $h = g_1 g_1^{(2)} - (g_1')^2$. Now $h' = g_1 g_1^{(3)} - g_1' g_1^{(2)}$, and thus $h'(x) = x\phi(x)g_1(x) - \phi(x)g_1'(x) = \phi(x)g_2(x) > 0$. As moreover g_1, g_1', $g_1^{(2)}$, and hence also h itself, tend to 0 as $x \to \infty$, it follows that $h < 0$. Consequently, $k(x) - x$ decreases.

The derivative of $g_1 - \frac{\sigma\bar{s}}{2}g_2$ equals $g_1' + \sigma\bar{s}g_1$, which is negative iff $k(x) - x < 1/(\sigma\bar{s})$. This will hold for $x > x_0$, with x_0 as in (2.20), because $k(x) - x$ is decreasing. □

The lemma shows that for typical values of γ, σ and $\bar{s}$ no monotonicity problems can occur. In fact, to obtain $x_0 = 0$ as solution of (2.20), it is required to have $\sigma\bar{s}$ as large as 1.25, which bound will hardly ever be violated. Hence we shall implicity assume in what follows that the parameters are chosen such that $\tilde{a}$ is well above $x_0 = x_0(\sigma, \bar{s})$ from (2.20). In analogy to lemma 2.1 we have

Lemma 2.3: For $\bar{s} \geq 1$ *we have for* $\tilde{a}$ *from* (2.18) *that*

$$0 \leq \frac{P(X > s, \tilde{X} < s - \tilde{a}\sigma_U) - \gamma}{\gamma} \leq \frac{1}{3}\sigma^2 \frac{\max(\bar{s}^2 - 1, 2)}{(1 - \frac{3}{4}\sigma\bar{s})} \ . \tag{2.21}$$

Proof: For $z \geq \bar{s} \geq 1$ we have that $0 \leq \phi''(z) \leq \phi''(\max(\bar{s}, \sqrt{3}))$, which leads to (cf. (2.17))

$$0 \leq P(X > s, \tilde{X} < s - \tilde{a}\sigma_U) - \gamma \leq \tfrac{1}{6}\sigma^3\phi''(\max(\bar{s}, \sqrt{3}))g_3(\tilde{a}) \ . \tag{2.22}$$

Now $g_2 \leq \frac{3}{2}g_1$, and therefore $g_1\left(1 - \frac{3}{4}\sigma\bar{s}\right) \leq g_1 - \frac{\sigma\bar{s}}{2}g_2$. As moreover $g_3 \leq 2g_1$, it follows that $g_3 \leq 2\left(g_1 - \frac{\sigma\bar{s}}{2}g_2\right)\Big/\left(1 - \frac{3}{4}\sigma\bar{s}\right)$ and thus $g_3(\tilde{a}) \leq 2\gamma/\{\sigma\phi(\bar{s})(1 - \frac{3}{4}\sigma\bar{s})\}$. In combination with (2.22) this leads to (2.21). □

It follows from (2.21) that $\tilde{a}$ is a lower bound for a^e. Hence we now have

$$\tilde{a} \leq a^e \leq a^{(1)} \ . \tag{2.23}$$

Moreover, $\tilde{a}$ is a second order approximation, and as such more precise than $a^{(1)}$ which can be seen from comparing the bounds in (2.14) and (2.21). In principle it is possible to continue in this fashion, by e.g. giving a third order upper bound. There are several reasons, however, not to pursue this. In the first place, the results become more complicated. Moreover, the lower bound on $\bar{s}$ will increase. Finally, the first and second order bounds seem adequate for practical purposes: for small σ and $\bar{s}$, the first order result may suffice, whereas for the remaining cases the second order result will be sufficiently precise. It seems more rewarding from a practical point of view to look for a more convenient version of the second order approximation.

We introduce therefore

$$a^{(2)} = a^{(1)} - \frac{\sigma\bar{s}}{2}\{(a^{(1)})^2 + 1 - a^{(1)}k(a^{(1)})\} \ , \tag{2.24}$$

with k as in (2.19). Then we have

Lemma 2.4: The approximations $a^{(2)}$ *and* $\tilde{a}$ *agree to second order, while moreover*

$$\tilde{a} \leq a^{(2)} \leq a^{(1)} \ . \tag{2.25}$$

Proof: From (2.13) and (2.18) it follows that

$$g_1(\tilde{a}) - \frac{\sigma\bar{s}}{2} g_2(\tilde{a}) = g_1(a^{(1)}) \ .$$

Letting $\delta = a^{(1)} - \tilde{a}$, we note that $\delta \geq 0$ and moreover $\delta = -\frac{\sigma\bar{s}}{2} g_2(\tilde{a}) / g_1'(\tilde{a} + \theta\delta)$. As $(-g_1'(\tilde{a} + \theta\delta))^{-1} = (1 - \Phi(\tilde{a} + \theta\delta))^{-1} \geq (-g_1'(\tilde{a}))^{-1}$, this leads to

$$\delta \geq -\frac{\sigma\bar{s}}{2} \frac{g_2(\tilde{a})}{g_1'(\tilde{a})} = \frac{\sigma\bar{s}}{2} \{\tilde{a}^2 + 1 - \tilde{a}k(\tilde{a})\} \ , \tag{2.26}$$

where we use (2.11) and (2.12) for the second step in (2.26). Comparison of (2.24) and (2.26) shows that in order to prove (2.25) it suffices to make clear that $-g_2/g_1'$ is decreasing and $\lim_{x \to \infty} (-g_2(x)/g_1(x)) = 0$.

To show this, observe that $(-g_2/g_1')' = (-g_2'g_1' + g_2 g_1^{(2)})/(g_1')^2$, which is negative iff $g_2' - g_2 g_1^{(2)}/g_1' < 0$. This in its turn is < 0 iff $2g_1 - g_2 k > 0$ (use (2.11), (2.12) and (2.19)), which is equivalent to $\{2 + xk(x)\} g_1(x) > -k(x) g_1'(x) = \phi(x)$. This translates into $(1 + xk(x))\phi(x) > (2 + xk(x))x(1 - \Phi(x))$, which is equivalent to

$$k(x) > x + \frac{x}{1 + xk(x)} \ . \tag{2.27}$$

For $x > 0$, we obtain from Kotz and Johnson (1985), p. 505, the inequality $k(x) > \ell(x) = \{3x + \sqrt{x^2 + 8}\}/4$. Hence it suffices to prove (2.27) with k replaced ℓ on both sides. Straightforward evaluation shows the latter inequality to be equivalent to $x^6 + 12x^4 + 36x^2 + 32 > x^6 + 12x^4 + 36x^2$, which is indeed true (and quite sharp for larger x!).

The remaining assertion of the lemma, concerning the agreement to second order between $a^{(2)}$ and $\tilde{a}$, is evident from the proof given. □

Combining (2.23) and (2.25), we obtain that $\tilde{a} \leq a^{(2)}$, $a^e \leq a^{(1)}$. As moreover $a^{(2)}$ and $\tilde{a}$ agree to second order, this suggests that $a^{(2)}$ will be the better approximation of the two. This impression could be checked by looking at the terms of next order. Since this is simple but quite tedious, we shall not bother to do so and restrict attention to numerical evidence. As mentioned in the introduction, such evidence can be found in Albers, Kallenberg and Otten (1992), so here we will just summarize the results. For values of practical interest ($0.01 < \sigma < 0.30$, $0.0025 < \pi < 0.15$ and γ between 1 and 100 ppm), $a^{(1)}$, $\tilde{a}$ and $a^{(2)}$ are compared to a^e, and the corresponding CL's to γ. The foremost conclusion is that $a^{(2)}$

performs strikingly well: $(CL^{(2)} - \gamma)/\gamma$ is completely negligible for practical purposes and $a^{(2)}$ can be assumed identical to a^e. Moreover, it is indeed found that $a^{(1)}$ is less accurate than $\tilde{a}$, which (almost always) is less accurate than $a^{(2)}$. In agreement with the bounds in (2.14) and (2.21), the errors grow as σ and $\bar{s}$ increase.

Next we briefly consider the yield $P(\tilde{X} < t)$ and the producer loss (PL) $P(X < s, \tilde{X} > t)$. The PL can be evaluated in the same way as the CL. Note, however,that it is easier to use the relation

$$PL = (1 - \pi) - P(X < s, \tilde{X} < t) = CL + (1 - \pi) - \text{yield} \ , \tag{2.28}$$

and to note subsequently that the yield is the simple univariate probability

$$\text{yield} = \Phi((s - a\sigma_U - \mu_X)/(\sigma_X^2 + \sigma_U^2)^{1/2}) = \Phi((\bar{s} - a\sigma)/(1 + \sigma^2)^{1/2}) \ . \tag{2.29}$$

For $a = a^e$ from (2.4), the correct PL and yield (i.e. the ones corresponding to $CL = \gamma$) are obtained. It is easily verified that for $a = a^{(2)}$ the errors in both PL and yield are negligible. Also note that it is straightforward to obtain from (2.29) the gain in yield achieved by using $a^{(2)}$ rather than the overly conservative $a^{(0)}$ from (2.6). For the (σ, π, γ)-values above, $a^{(0)}$ varies between 1.8 and 4.3, the difference $a^{(0)} - a^{(2)}$ between 0.4 and 2.0 and the gain in yield between 0.02% and 5.97%. (see Albers, Kallenberg and Otten (1992) for more details)

3 Second Order Unbiased Test Limits

Inspection of (2.24), (2.13) and (2.7) once more makes explicit that for the actual evaluation of the advocated approximate test limit $t^{(2)} = s - a^{(2)}\sigma_U$, we need to know μ_X, σ_X, σ_U, s and γ. To assume s and γ known seems fair, but for μ_X, σ_X and σ_U this is much less realistic. Quite often we shall have to rely for these parameters on corresponding estimators $\hat{\mu}_X$, $\hat{\sigma}_X$ and $\hat{\sigma}_U$. In practice it is customary to plug in the resulting estimates and to proceed as in section 2. But this approach is only sound for very large sample sizes, as will become apparent in this and the following section.

To begin with, observe that from now on we are dealing with stochastic test limits $t = t(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)$, leading in their turn to stochastic CL's. Which t we prefer will depend on the criterion we choose to judge the closeness of such a $\widehat{CL}$ to γ. In this section we will take the point of view of unbiased estimation: we look for $t^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)$ such that for a new item to be tested we have

$$\widehat{ECL}^{(3)} = EP(X > s, \tilde{X} < t^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)) = \gamma \tag{3.1}$$

with sufficient precision. Arguing as in (2.7)–(2.9), we obtain that (3.1) can be translated into

$$E \int_0^\infty \{\Phi(\bar{s} + \sigma y) - \Phi(\bar{s})\}\phi(y + \hat{\sigma}_U a^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)/\sigma_U)dy = \gamma \ , \tag{3.2}$$

where in analogy to (2.4) we have $t^{(3)} = s - a^{(3)}\hat{\sigma}_U$, and thus

$$a^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U) = \{s - t^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\}/\hat{\sigma}_U \ . \tag{3.3}$$

To determine the function $a^{(3)}$, we can use an expansion similar to that of lemma 2.3, leading from (3.2) to

$$E\left\{g_1\left(\frac{\hat{\sigma}_U}{\sigma_U}a^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\right) - \frac{\sigma\bar{s}}{2}g_2\left(\frac{\hat{\sigma}_U}{\sigma_U}a^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\right) + \cdots\right\} = \frac{\gamma}{\sigma\phi(\bar{s})} \ , \tag{3.4}$$

again to sufficient precision. Comparison of (3.4) to (2.18) suggests that the use of consistent estimators for the three parameters involved will lead to an $a^{(3)}$ which is closely related to $\tilde{a}$, and hence to $a^{(2)}$ from (2.24). The difference will consist of some appropriate terms to correct for the effect of having to estimate instead of being able to use the known values.

However, for arbitrary $(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)$ the explicit expression for this difference will be unduly complicated. Therefore we first present two choices for these estimators which are likely to be used in practice. After this, we can take advantage of their properties to derive a relatively simple correction $c = a^{(3)} - a^{(2)}$. Typically, estimation of σ_U will entail a series of repeated measurements. Suppose we consider n parts and apply p replications:

$$X_{ij} = X_i + U_{ij}, i = 1, \ldots, n, j = 1, \ldots, p \ . \tag{3.5}$$

Denoting averaging over an index by a "$\cdot$", we have (cf. e.g. Scheffé (1959) p. 228) the independent statistics

$$X_{\cdot\cdot}, S_w^2 = \sum_i \sum_j (X_{ij} - X_{i\cdot})^2/(n(p-1)), S_b^2 = \sum_i (X_{i\cdot} - X_{\cdot\cdot})^2/(n-1) \ . \tag{3.6}$$

As moreover $X_{\cdot\cdot}$ is $N(\mu_X, \sigma_X^2/n + \sigma_U^2/(np))$, $n(p-1)S_w^2/\sigma_U^2$ is $\chi^2(n(p-1))$ and $(n-1)S_b^2/(\sigma_X^2 + \sigma_U^2/p)$ is $\chi^2(n-1)$, this leads to the $UMVU$ estimators (cf.

Lehmann (1983), p 198)

$$\hat{\mu}_X = X_{..}, \hat{\sigma}_X^2 = S_b^2 - S_w^2/p, \hat{\sigma}_U^2 = S_w^2 \ . \tag{3.7}$$

It is straightforward to obtain (co)variances. We e.g. have $\text{var}(\hat{\sigma}_X^2) = 2(\sigma_X^2 + \sigma_U^2/p)^2/(n-1) + 2\sigma_U^4/(np^2(p-1))$ and $\text{cov}(\hat{\sigma}_X^2, \hat{\sigma}_U^2) = -2\sigma_U^4/(np(p-1))$. From (2.5) we know that $\sigma = \sigma_U/\sigma_X \ll 1$, which allows simplification to (equalities holding up to $\mathcal{O}(\sigma^2)$)

$$n \text{ var}\left(\frac{\hat{\mu}_X}{\sigma_X}\right) = \frac{(n-1)}{2} \text{ var}\left(\frac{\hat{\sigma}_X^2}{\sigma_X^2}\right) = \frac{n(p-1)}{2} \text{ var}\left(\frac{\hat{\sigma}_U^2}{\sigma_U^2}\right) = 1 + \mathcal{O}(\sigma^2),$$

$$\text{cov}\left(\frac{\hat{\sigma}_X^2}{\sigma_X^2}, \frac{\hat{\sigma}_U^2}{\sigma_U^2}\right) = \mathcal{O}(\sigma^2) \tag{3.8}$$

Given a budget of np observations to base the estimators on, the question remains how to choose p. Inspection of (3.8) suggests strongly to let $p = 2$. Indeed, for $\theta = \sigma_X^2/\sigma_U^2 \geq 1$, which is clearly the case here, Scheffé (1959, p. 238) shows that this is the optimal choice. Hence we shall always let p equal 2 in what follows. Summarizing our first choice, we observe that the estimators defined through (3.6) and (3.7) are unbiased, and to the order considered, independent with variances n^{-1}, $2n^{-1}$ and $2n^{-1}$ for $\hat{\mu}_X/\sigma_X$, $\hat{\sigma}_X^2/\sigma_X^2$ and $\hat{\sigma}_U^2/\sigma_U^2$, respectively.

Also of practical interest is the following design. Quite often the test data described above form part of a typically much larger set of production data. Hence we have a sample $\tilde{X}_1, \ldots, \tilde{X}_m$ from $\tilde{X}$ in (2.1), with $m > n$ (and usually $m \gg n$). A subset of size n is measured twice, giving rise to $\tilde{X}_{j_i} = X_{i1}$ and a replication X_{i2} for $i = 1, \ldots, n$. The X_{ij} are again of the form (3.5), with $p = 2$. For σ_U^2 the situation is the same as before: just as in (3.7) it is estimated by (using that $p = 2$)

$$\hat{\sigma}_U^2 = \frac{1}{2n} \sum_{i=1}^{n} (X_{i2} - X_{i1})^2 \ . \tag{3.9}$$

But the estimators for μ_X and σ_X^2 can be improved, using the non-replicated $m - n$ production data. In fact, the mean $\tilde{X}_{\cdot}^{(m-n)}$ of these latter data is $N(\mu_X, (\sigma_X^2 + \sigma_U^2)/(m-n))$, and $\tilde{X}_{\cdot}^{(m-n)}$ can be combined optimally with $X_{..}$ from (3.6), which is $N(\mu_X, (\sigma_X^2 + \sigma_U^2/2)/n)$. It is immediately clear, however, that to $\mathcal{O}(\sigma^2)$ it is also optimal to use simply $\hat{\mu}_X = \tilde{X}_{\cdot}$: the gain in variance reduction by using the replications X_{i2}, is utterly negligible, especially when besides $\sigma_U \ll \sigma_X$ we also have $m \gg n$. For $\hat{\sigma}_X^2$, completely analogous arguments hold. Hence in this second situation we suggest

$$\hat{\mu}_X = \tilde{X}_\cdot, \hat{\sigma}_X^2 = \frac{1}{m-1}\sum_{i=1}^{m}(\tilde{X}_i - \tilde{X}_\cdot)^2 - \hat{\sigma}_U^2 , \tag{3.10}$$

with $\hat{\sigma}_U^2$ as in (3.9). Again the estimators are unbiased, and to the order considered independent, but now with variances m^{-1}, $2m^{-1}$ and $2n^{-1}$ for $\hat{\mu}_X/\sigma_X$, $\hat{\sigma}_X^2/\sigma_X^2$ and $\hat{\sigma}_U^2/\sigma_U^2$, respectively.

For either of the two choices above, choose the function $a^{(3)} = a^{(3)}(\mu_X, \sigma_X, \sigma_U)$ such that $a^{(3)} - a^{(2)}$ equals

$$\frac{k(a^{(1)})\{2a^{(1)}k(a^{(1)}) + 1 - (a^{(1)})^2\}}{4n} + \frac{(\bar{s}^4 + 4\bar{s}^2 + 1)(k(a^{(1)}) - a^{(1)})}{4m} , \tag{3.11}$$

with $a^{(2)}$, $a^{(1)}$ and k as in (2.24), (2.13) and (2.19), respectively. Clearly (3.11) is adapted to the choice of $\hat{\mu}_X$ and $\hat{\sigma}_X^2$ according to (3.10); if (3.7) is appropriate, we should read $m = n$ in (3.11). In the special case where μ_X and σ_X are known and only σ_U needs to be estimated, the last term of (3.11) can be omitted. We have:

Theorem 3.1: Let $\bar{s}$ be bounded and suppose that $n \to \infty$ and $\sigma \to 0, \gamma \to 0$ such that $a^{(1)}$ is bounded. For $t^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U) = s - a^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\hat{\sigma}_U$, with $\hat{\mu}_X, \hat{\sigma}_X$ and $\hat{\sigma}_U$ as in (3.7)–(3.10) *and the function $a^{(3)}$ as in* (3.11), *we have, for some $C > 0$,*

$$EP(X > s, \tilde{X} < t^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)) = \gamma\{1 + \mathcal{O}(n^{-2} + \sigma^2)\} + \mathcal{O}(e^{-Cn}) . \tag{3.12}$$

Proof: Once the theorem has been proved for the choice from (3.7), the case where $m > n$ follows automatically. Hence we will assume $m = n$. Moreover, let

$$V = (\hat{\mu}_X - \mu_X)/\sigma_X, W = \hat{\sigma}_X^2/\sigma_X^2 - 1, Z = \hat{\sigma}_U^2/\sigma_U^2 - 1 . \tag{3.13}$$

For every $\tilde{c} > 0$, we have $P(A) = \mathcal{O}(e^{-Cn})$ for some $C > 0$, where $A = \{\max(|V|, |W|, |Z|) > \tilde{c}\}$. Since probabilities are bounded, $E\{|P(X > s, \tilde{X} < t^{(3)}) - \gamma| I_A\} = \mathcal{O}(P(A)) = \mathcal{O}(e^{-Cn})$, where I_A denotes the indicator function of the set A. Hence the inclusion of the last term in (3.12) allows us to restrict attention to A^c in what follows, i.e. to assume that each estimator is close to the parameter it corresponds to, which in its turn makes it possible to apply expansions.

Let $c^{(3)}(\mu_X, \sigma_X, \sigma_U) = a^{(3)}(\mu_X, \sigma_X, \sigma_U) - a^{(1)}(\mu_X, \sigma_X, \sigma_U)$. From (3.11) and (2.24) it follows that $c^{(3)} = \mathcal{O}(n^{-1} + \sigma)$. Together with (3.2)–(3.4) this gives that the expected consumer loss $\widehat{ECL}^{(3)}$ equals

$$E\left\{\sigma\phi(\bar{s})\left[g_1\left(\frac{\hat{\sigma}_U}{\sigma_U}a^{(1)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\right)\right.\right.$$

$$\left.+\frac{\hat{\sigma}_U}{\sigma_U}c^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)g_1'\left(\frac{\hat{\sigma}_U}{\sigma_U}a^{(1)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\right)\right]$$

$$\left.-\frac{\sigma^2\bar{s}}{2}\phi(\bar{s})g_2\left(\frac{\hat{\sigma}_U}{\sigma_U}a^{(1)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\right)\right\}I_{A^c} + \mathcal{O}(\gamma(n^{-2} + \sigma^2) + e^{-Cn}) \ . \tag{3.14}$$

As $a^{(1)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U) = g_1^{-1}(\hat{\sigma}_X\gamma/(\hat{\sigma}_U\phi((s - \hat{\mu}_X)/\hat{\sigma}_X)))$, we observe that (3.14) requires expansion of $m(X, Y) = g_1(Xg_1^{-1}(bY/X))$ around $b = m(1, 1)$, where

$$b = \frac{\sigma_X\gamma}{\sigma_U\phi(\bar{s})}, \ X = \frac{\hat{\sigma}_U}{\sigma_U}, \ Y = \frac{\hat{\sigma}_X}{\sigma_X}\frac{\phi(\bar{s})}{\phi((s - \hat{\mu}_X)/\hat{\sigma}_X)} \ . \tag{3.15}$$

To evaluate the expectation of the resulting expansion, we need $E(X - 1)^k(Y - 1)^\ell$, for integers k and ℓ with $0 \le k, \ell \le 4$. In view of (3.13) we have $X - 1 = (1 + Z)^{1/2} - 1 = Z/2 - Z^2/8 + \cdots$. From (3.7) it is known that $EZ = 0$, $EZ^2 = 2n^{-1}$. Moreover, it is also easily verified that EZ^3 and EZ^4 are both $\mathcal{O}(n^{-2})$. Hence $E(X - 1) = -(4n)^{-1}$ and $E(X - 1)^2 = (2n)^{-1}$, apart from the contribution to the remainder in (3.12), while $E(X - 1)^3$ and $E(X - 1)^4$ only contribute to the remainder. For $(Y - 1)$ we observe that (3.13) and (3.15) give

$$Y = (1 + W)^{1/2}\phi(\bar{s})/\phi((\bar{s} - V)(1 + W)^{-1/2}) \ , \tag{3.16}$$

which can be expanded in terms of V and W. Proceeding as above, we obtain that to the desired order $E(Y - 1) = (4n)^{-1}(\bar{s}^4 + 4\bar{s}^2 + 1)$. The remaining (mixed) moments can either be relegated to the remainder, or will not occur in the expansion (this is the case with $E(Y - 1)^2$). In fact, the expansion of $(m(X, Y) - b)/g_1'(a^{(1)})$ up to and including second order terms equals

$$(X - 1)k(a^{(1)}) - (Y - 1)(k(a^{(1)}) - a^{(1)}) - \tfrac{1}{2}(X - 1)^2k(a^{(1)})a^{(1)}$$

$$\times (2k(a^{(1)}) - a^{(1)}) + (X - 1)(Y - 1)k(a^{(1)})a^{(1)}(k(a^{(1)}) - a^{(1)}) \ . \tag{3.17}$$

Using (3.17) and the evaluated moments, we now obtain from (3.14) that $\widehat{ECL}^{(3)} - \gamma$ equals

$$\sigma\phi(\bar{s})g_1'(a^{(1)})\left\{-\frac{k(a^{(1)})[2a^{(1)}k(a^{(1)})+1-(a^{(1)})^2]}{4n}\right.$$

$$\left.-\frac{(\bar{s}^4+4\bar{s}^2+1)(k(a^{(1)})-a^{(1)})}{4n}+c^{(3)}\right\}-\frac{\sigma^2\bar{s}}{2}\phi(\bar{s})g_2(a^{(1)})$$

$$+\mathcal{O}(\gamma(n^{-2}+\sigma^2)+e^{-Cn}) \tag{3.18}$$

But from (3.18) it is immediate that the choice for $a^{(3)}$ in (3.11) will lead to a $c^{(3)} = a^{(3)} - a^{(1)}$ which will cancel all terms on the right hand side of (3.18), apart from the remainder. □

To apply the theorem, one begins by evaluating $a^{(1)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U) = g_1^{-1}(\hat{\sigma}_X\gamma/(\hat{\sigma}_U\phi((s-\hat{\mu}_X)/\hat{\sigma}_X)))$ (cf. (2.13)), then uses this to obtain $a^{(2)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)$ through (2.24), and subsequently $a^{(3)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_X)$ through (3.11). Since $k(x) \geq x$, both correction terms in (3.11) are nonnegative. This agrees with our intuition, as they stand for the complication of having to estimate the parameters, rather than knowing them. Moreover, note that the behavior of the correction terms in $a^{(3)} - a^{(2)}$ differs from that of $a^{(2)} - a^{(1)}$. The coefficient $(a^{(1)})^2 + 1 - a^{(1)}k(a^{(1)})$ of $\sigma\bar{s}/2$ is quite small (e.g. it decreases from 1 for $a^{(1)} = 0$ to 0.26 for $a^{(1)} = 2$). But in (3.11) the coefficient of $(4n)^{-1}$ increases considerably as $a^{(1)}$ grows (e.g. from 0.8 for $a^{(1)} = 0$ to 15.4 for $a^{(1)} = 2$), while the factor $(\bar{s}^4 + 4\bar{s}^2 + 1)$ in the coefficient of $(4m)^{-1}$ also clearly increases steeply in $\bar{s}$.

This suggests that considerable sample sizes will be required before the correction $a^{(3)} - a^{(2)}$ can be neglected. In fact, in Albers, Kallenberg and Otten (1992), a simulation study has been carried out for the special case where μ_X and σ_X are known and only σ_U is unknown, using sample sizes $n = 40$ and $n = 80$. It turns out that using $a^{(3)}$ indeed leads to an average $\widehat{CL}^{(3)}$ which is quite close to γ, and as such completely satisfactory in practice, while ignoring the correction and using $a^{(2)}$ from (2.24) leads to average $\widehat{CL}^{(2)}$'s which are much too large (up to 2γ for $n = 40$). Hence for such intermediate sample sizes the first correction term in (3.11) indeed shows its value. Extension of the simulation study to the general case shows that the picture is quite similar for the second correction term. For $n = 40$ or 80 it is found that, in the case where m and n are about equal, inclusion of this second correction in general drastically improves the result. On the other hand, in the case where $m \gg n$ (e.g. $m = 2500$), the improvement is indeed negligible. Another aspect which manifests itself clearly in the simulations, is the steepness of $(\bar{s}^4 + 4\bar{s}^2 + 1)$. In the case where $m \approx n$, the improvement is large for $\pi = 0.01$, while for $\pi = 0.15$ it has lost its importance.

4 Test Limits for Which γ is Violated with Small Probability

It is gratifying to note in the previous section that our second order asymptotics works well in the sense that $\widehat{CL}^{(3)}$ on the average is close to the prescribed γ. However, simulations also show that $\widehat{CL}^{(3)}$ varies widely around γ (e.g. leading to upper and lower 5% quantiles like 30 and 215 ppm, respectively, when $\gamma = 100$ ppm, $n = 40$ and only σ_U is estimated). If new estmates are obtained regularly, for example for each new batch of products, the long run average of $\widehat{CL}^{(3)}$ will indeed tend to γ, notwithstanding the considerable variation between batches. For certain applications, for example one and the same consumer receiving all batches, this will be quite satisfactory. In general, however, second order unbiasedness will be too weak.

It is not difficult to identify the cause of the variation signalled above. A glance at the proof of theorem 3.1 reveals that in the expansion of $\widehat{CL}^{(3)}$ factors like $\hat{\sigma}_U/\sigma_U - 1$ (cf. (3.15) and (3.17)) occur. But while $E(\hat{\sigma}_U/\sigma_U - 1)$ is of order n^{-1}, we have that $\hat{\sigma}_U/\sigma_U - 1$ itself is of order $n^{-1/2}$ in probability. As we noted in section 3 that terms of order n^{-1} are not at all negligible in correcting the expectation for intermediate n, it stands to reason that terms of order $n^{-1/2}$ will be quite important. In fact, let

$$\ell(x) = k(x)/(k(x) - x) \ , \tag{4.1}$$

then we can show that

Lemma 4.1: Under the conditions of theorem 3.1 *we have for* $\widehat{CL}^{(j)} = P(X > s, \tilde{X} < t^{(j)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)), j = 2, 3,$

$$E\left(\frac{\widehat{CL}^{(j)} - \gamma}{\gamma}\right)^2 = \frac{\ell^2(a^{(1)})}{2n} + \frac{\bar{s}^4 + 1}{2m} + \mathcal{O}(n^{-2} + \sigma^4 + \gamma^{-2}e^{-Cn}) \ . \tag{4.2}$$

Proof: Using (3.1)–(3.4) in combination with (3.14) and (3.17) we obtain that (on the set A^c from the proof of theorem 3.1)

$$\widehat{CL}^{(3)} - \gamma = \sigma\phi(\bar{s})g_1'(a^{(1)})\{(X - 1)k(a^{(1)}) - (Y - 1)(k(a^{(1)}) - a^{(1)}) + a^{(3)} - a^{(2)}\} + \mathcal{O}(\gamma\{(X - 1)^2 + (Y - 1)^2 + n^{-2} + \sigma^2\}) \ . \tag{4.3}$$

Now $\sigma\phi(\bar{s}) = \gamma/b = \gamma/g_1(a^{(1)})$, and therefore $\sigma\phi(\bar{s})g_1'(a^{(1)}) = -\gamma/(k(a^{(1)}) - a^{(1)})$, which in its turn leads to, using (4.1),

$$(\widehat{CL}^{(3)} - \gamma)/\gamma = -(X-1)\ell(a^{(1)}) + (Y-1) - (a^{(3)} - a^{(2)})\ell(a^{(1)})/k(a^{(1)}) + \mathcal{O}(n^{-2} + \sigma^2 + (X-1)^2 + (Y-1)^2) \ . \tag{4.4}$$

Through (3.16) it can be shown that $E(Y-1)^2 = (2m)^{-1}(\bar{s}^4 + 1)$ to the desired order. We already know that $E(X-1)^2 = (2n)^{-1}$ and that all other moments involved are $\mathcal{O}(n^{-2})$. Hence the result in (4.2) follows for $j = 3$. As $a^{(3)} - a^{(2)}$ only plays a role in the remainder, it is immediate that for $j = 2$, with $c^{(2)} = a^{(2)} - a^{(1)}$, the result is true as well. □

A first conclusion from this lemma is that the correction step from $a^{(2)}$ in (2.24) to $a^{(3)}$ in (3.11), which in the previous section was seen to be vital in controlling the unbiasedness, is unimportant in the context of the mean squared (relative) error. More interesting, however, it is to note that (4.2) and (4.4) are quite illuminating, as these formulas provide a simple explanation for the wide variability in $\widehat{CL}^{(3)}$-values mentioned at the beginning of this section. In going from $\hat{\sigma}_U$ to $\widehat{CL}^{(3)}$, the relative error is inflated by a factor $\ell(a^{(1)})$, which is rapidly increasing in $a^{(1)}$ (e.g. $a^{(1)} = 1$ gives 3, while for $a^{(1)} = 2$ it equals 6.4). Next, the way from $\hat{\mu}_X$ and $\hat{\sigma}_X$ to $\widehat{CL}^{(3)}$ goes through $(Y-1)$. Now $(Y-1)$ lacks an inflating factor, but from (3.16) we see that here the effect is hidden inside, leading to a factor $\bar{s}^4 + 1$ in (4.2). Together, these effects explain why a heuristic argument, according to which the fact that for moderately large n and m the estimates $\hat{\mu}_X$, $\hat{\sigma}_X$ and $\hat{\sigma}_U$ will be reasonably close to μ_X, σ_X and σ_U, respectively, should imply that $\widehat{CL}^{(3)}$ is reasonably close to γ as well, turns out to be quite misleading.

Fortunatey, using lemma 4.1 it is now rather straightforward to present a stronger criterion when the unbiasedness property is too weak. Just as in section 3, we will replace the test limit $t^{(2)}$ for the case of known parameters by a corrected test limit. This $t^{(4)} = t^{(4)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)$ will be smaller than $t^{(3)}$ from (3.1), but still be such that the correction $c^{(4)} = a^{(4)} - a^{(2)}$, with $a^{(4)} = (s - t^{(4)})/\hat{\sigma}_U$ (cf. (3.3)), tends to zero as $n \to \infty$. Hence the corresponding $\widehat{CL}^{(4)}$ will tend to γ in probability. The idea now is to select $c^{(4)}$ such that, for some small positive α

$$P(\widehat{CL}^{(4)} > \gamma) = \alpha \ , \tag{4.5}$$

with sufficient precision. Define the function $a^{(4)} = a^{(4)}(\mu_X, \sigma_X, \sigma_U)$ by

$$a^{(4)} = a^{(2)} + u_\alpha \left\{ \frac{k^2(a^{(1)})}{2n} + \frac{(k(a^{(1)}) - a^{(1)})^2(\bar{s}^4 + 1)}{2m} \right\}^{1/2} \ , \tag{4.6}$$

where $u_\alpha = \Phi^{-1}(1-\alpha)$ is the upper α-quantile of the standard normal distribution and $a^{(2)}$, $a^{(1)}$, k and ℓ are given in (2.24), (2.13), (2.19) and (4.1), respectively. Then we obtain the following result

Lemma 4.2: Under the conditions of Theorem 3.1 *we have for* $t^{(4)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U) = s - a^{(4)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)\hat{\sigma}_U$, *with* $\hat{\mu}_X$, $\hat{\sigma}_X$ *and* $\hat{\sigma}_U$ *as in* (3.7)–(3.10) *and the function* $a^{(4)}$ *as in* (4.6), *that*

$$P(P(X > s, \tilde{X} < t^{(4)}(\hat{\mu}_X, \hat{\sigma}_X, \hat{\sigma}_U)) > \gamma) = \alpha + o(1) \ . \tag{4.7}$$

Proof: As $c^{(4)}$ is small, we can apply (4.4) to $\widehat{CL}^{(4)}$ instead of $\widehat{CL}^{(3)}$, leading to

$$\frac{\widehat{CL}^{(4)} - \gamma}{\gamma} \text{ is } AN\left(-c^{(4)}\frac{\ell(a^{(1)})}{k(a^{(1)})}, \frac{\ell^2(a^{(1)})}{2n} + \frac{(\bar{s}^4 + 1)}{2m}\right) . \tag{4.8}$$

Hence $\widehat{CL}^{(4)}$ will satisfy (4.7) if $c^{(4)}$ is chosen such that $1 - \Phi(\{c^{(4)}\ell(a^{(1)})/k(a^{(1)})\}/\{\ell^2(a^{(1)})/(2n) + (\bar{s}^4 + 1)/(2m)\}^{1/2}) = \alpha$. This precisely produces the choice $c^{(4)} = a^{(4)} - a^{(2)}$ given by (4.6). □

Note that in view of (4.8) we have to first order

$$E\widehat{CL}^{(4)} = \gamma(1 - \delta) \ , \tag{4.9}$$

with $\delta = u_\alpha\{\ell^2(a^{(1)})/(2n) + (\bar{s}^4 + 1)/(2m)\}^{1/2}$. Hence to ensure that γ is violated with probability at most α, the average consumer loss has to be lowered to $\gamma(1 - \delta)$. In passing note that for $\alpha = \frac{1}{2}$ we have $\delta = 0$, and $\widehat{CL}^{(4)}$ agrees with $\widehat{CL}^{(3)}$ to first order. A sensible criterion now requires that for given α_0 and δ_0 condition (4.5) holds for some $\alpha \leq \alpha_0$ and condition (4.9) for some $\delta \leq \delta_0$. In this way violation of γ can be kept an exception without lowering the average consumer loss too much. It immediately follows that the sample sizes n and m should satisfy

$$\ell^2(a^{(1)})/n + (\bar{s}^4 + 1)/m \leq 2\delta_0^2/u_\alpha^2 \ . \tag{4.10}$$

To give an impression of the sample sizes involved, let us specialize once again to the case where $m \gg n$. Then the second term on the left-hand side of (4.10) can be neglected and it follows that $n \geq n_0 = u_\alpha^2\ell^2(a^{(1)})/(2\delta_0^2)$. For $\alpha = 0.05$ or 0.10 and $\delta_0 = 0.1$ or 0.2, the factor $u_\alpha^2/(2\delta_0^2)$ varies from 20 to 140. Together with values of $\ell(a^{(1)})$ of about 6, it is clear that n_0 is very large. Hence the choice of δ_0 is a matter of balance: if δ_0 is small, n_0 is large, which is expensive, but if δ_0 is large, $E\widehat{CL}^{(4)}$ will be small, which will be reflected in a large loss of yield, which is also expensive. In Albers, Kallenberg and Otten (1992) a small numerical study is carried out for $n = 500$ and $n = 2000$, which shows that for such sample sizes the first order asymptotics works well.

5 Sampling Schemes

In the previous two sections we have seen that moderate sample sizes m and n allow second order unbiasedness, while large sample sizes are needed if not only the mean, but also the variation has to be controlled. However, there is still another point of view to take into account when selecting sample sizes for the estimators and this we will briefly consider in this section. Up to now, in selecting a criterion for a test limit, we have considered one single new item to be tested on the basis of our estimated test limit (cf. e.g. (3.1)). In practice we deal of course with a whole series N, with typically $N \gg n$, of new items to be judged through a single test limit, based on n (and m) earlier observations. Intuitively it is clear that n should increase in N: the longer one is going to use a certain limit, the better it should be.

To make this intuitive feeling more precise, we argue as follows. As always, we use a test limit for which $E\widehat{CL}_n = \gamma$ holds (at least) to first order. Let G_N be the average realized consumer loss over the N new items, then to first order $EG_N = \gamma$ and

$$\begin{aligned} \operatorname{var} G_N &= (1 - N^{-1}) \operatorname{var}(\widehat{CL}_n) + N^{-1}\gamma(1 - \gamma) , \\ &= \operatorname{var}(\widehat{CL}_n) + N^{-1}\gamma(1 - \gamma) , \end{aligned} \tag{5.1}$$

once more to first order, using that $\operatorname{var}(\widehat{CL}_n) = \mathcal{O}(\gamma^2)$. This suggests that $\operatorname{var}(\widehat{CL}_n)$ should be of the order of magnitude $N^{-1}\gamma(1 - \gamma)$. Taking both terms in (5.1) equal and using lemma (4.1), we obtian (cf. (4.2))

$$\ell^2(a^{(1)})/n + (\bar{s}^4 + 1)/m = 2(1 - \gamma)/(\gamma N) \approx 2/(\gamma N) . \tag{5.2}$$

Specializing as before to the case where $m \gg n$ for the sake of illustration, (5.2) reduces to $n \approx n_0 = \frac{1}{2}\ell^2(a^{(1)})\gamma N$. Having again values of $\ell(a^{(1)})$ of about 6, letting e.g. $\gamma \in (10^{-5}, 10^{-4})$ and $N \sim 10^6$ (which is implied by speaking about parts per million), we see again large sample sizes n arising. Combination with the result of section 4 leads to

$$n \geq \tfrac{1}{2}\ell^2(a^{(1)}) \max(u_\alpha^2/\delta_0^2, \gamma N) . \tag{5.3}$$

We conclude this section with the following consideration, inspired by the fact that large sample sizes n (and m) are often required. Without loss of generality assume that n, m and N are multiples of a given positive integer k. Suppose that the whole process described before is now divided into k subgroups each leading to a $G_{N/k}^{(j)}$, $j = 1, \ldots, k$. From (5.1) it follows that $\operatorname{var}(G_{N/k}^{(j)}) = \operatorname{var} \widehat{CL}_{n/k} +$

$kN^{-1}\gamma(1-\gamma)$. In view of (4.2) it is clear that var $\widehat{CL}_{n/k} = k$ var $\widehat{CL}_n$ to first order. Hence $\tilde{G}_N = k^{-1}\sum_{j=1}^{k} G_{N/k}^{(j)}$ has

$$\text{var } \tilde{G}_N = \text{var } \widehat{CL}_n + N^{-1}\gamma(1-\gamma) \ . \tag{5.4}$$

Comparison of (5.4) and (5.1) shows that var $\tilde{G}_N \approx$ var G_N. Hence there is no loss in working with a number of smaller samples, each leading to its own estimates, as compared to using one single, vary large sample. Moreover, from the perspective of robustness, it is even quite attractive to work with a number of separate steps, as this will provide better protection against deviations from the assumption that the production process is stationary.

References

Albers W, Kallenberg WCM (1991) A simple approximation to the bivariate normal distribution with large correlation coefficient. Technical Report 1011, Unversity of Twente. To appear in Journal of Multivariate Analysis

Albers W, Kallenberg WCM, Otten GD (1992) Accurate test limits with estimated parameters. Technical Report 1024, University of Twente. To appear in Technometrics

Cox DR, Wermuth N (1991) A simple approximation for bivariate and trivariate normal integrals. International Statistical Review 59:263–269

Easterling RC, Johnson ME, Bement TR, Nachtsheim CJ (1991) Statistical tolerancing based on consumer's risk considerations. Journal of Quality Technology 23:1–11

Kotz S, Johnson NL (1985) Encyclopedia of statistical sciences, Volume 5, Wiley, New York

Received 12.01.93