Tolerance Bounds and C_{pk} Confidence Bounds Under Batch Effects

Fritz Scholz
Boeing Shared Services Group [1]

Mark Vangel
National Institute of Standards and Technology [2]

published in

Advances in Stochastic Models for Reliability, Quality and Safety
edited by W. Kahle, E. von Collani, J. Franz, and U. Jensen
Birkäuser, Boston, 1998, pp. 361-379.

[1]Research and Technology, P.O. Box 3707, MS 7L-22,
Seattle WA 98124-2207, e-mail: fritz.scholz@boeing.com
[2]Statistical Engineering Division, Building 820/Room 353,
Gaithersburg, MD-20899-0001 e-mail: vangel@cam.nist.gov

Abstract
The capability index C_{pk} for a process, that produces parts with normally distributed characteristic X, is defined as $C_{pk} = \min(U - \mu, \mu - L)/(3\sigma) = (T - |\mu - \nu|)/(3\sigma)$, where U and L are upper and lower specification limits for X, μ and σ are process mean and standard deviation, and $\nu = (U + L)/2$, $T = (U - L)/2$. Using a sample $X_1, \ldots, X_n$ of independent observations from $\mathcal{N}(\mu, \sigma^2)$ Chou et al. (1990) (with clarification by Kushler and Hurley (1992)) showed how to get lower confidence bounds for C_{pk}. Here we extend this methodology to cover the situation where samples come in batches and the intra batch correlation reduces the amount of independent information. In parallel we also apply this extension to the closely related tolerance bounds or confidence bounds for quantiles. Introducing the simple trick of effective sample size these problems are linked quite successfully to existing tables for tolerance bounds or C_{pk} confidence bounds. The basic idea is to "approximate" the complicated data situation with an i.i.d. scenario with reduced overall sample size. The approximation is anchored by analysis to the two extreme situations where the within batch correlation is zero or one. For the in-between cases the effective sample size is chosen on a simple heuristic basis, namely by matching the variances of the sample mean under the batch effect model and its i.i.d. approximation. The coverage properties of the resulting method, examined by simulation, were found to be reasonably accurate near the extreme cases and mildly conservative in-between.

AMS Classification: Primary 62F25, 62N10; Secondary 62E17

Acknowledgement: We wish to thank the referee for constructive comments improving the presentation of this article.

Chapter 24

24.1 Introduction and Overview

It is assumed that we deal with data from a normal population $\mathcal{N}(\mu, \sigma^2)$ with mean μ and standard deviation σ. For i.i.d. samples it has long been known how to construct tolerance bounds or confidence bounds for normal p-quantiles $x_p + \sigma z_p$ based on the noncentral t-distribution. The earliest reference we found was Jennett and Welch (1939), but also see Johnson and Welch (1940), Owen (1968, 1985), and Odeh and Owen (1980) for extensive tables.

Closely related to such quantiles is the process capability index C_{pk}, introduced by Kane (1986), and defined as

$$C_{pk} = \min\left\{\frac{U-\mu}{3\sigma}, \frac{\mu-L}{3\sigma}\right\} = \frac{\frac{1}{2}(U-L) - |\mu - \frac{1}{2}(U+L)|}{3\sigma} ,$$

where U and L are given upper and lower product specification limits. Confidence bounds for C_{pk}, again for the i.i.d. case, were given by Chou et. al (1990) with clarification by Kushler and Hurley (1992). For a comprehensive overview of capability indices see Kotz and Johnson (1993).

Often the data of a production process arrive in batches with significant within batch correlation. A popular model for such batch data is $\{X_{ij},\ j = 1, \ldots, n_i,\ i = 1, \ldots, B\}$, where B is the number of batches and n_i is the size of the i^{th} batch. It is then assumed that $X_{ij} = \mu + b_i + e_{ij}$, where b_i is normal with mean zero and variance σ_b^2 and e_{ij} is normal with mean zero and variance σ_e^2. The effects b_i and $\{e_{ij}\}$ are assumed to be mutually independent. Hence X_{ij} is normally distributed with mean μ and variance $\sigma^2 = \sigma_b^2 + \sigma_e^2$. The correlation of two different observations within the same batch is $\rho = \sigma_b^2/(\sigma_b^2 + \sigma_e^2)$ which can range anywhere within $[0, 1]$. Under such a scenario one usually still wants to characterize aspects of the overall $\mathcal{N}(\mu, \sigma^2)$ population and not of individual batches. Hence it is desirable to extend the methodology for constructing tolerance bounds or C_{pk} confidence bounds to such batch data.

Although this sampling model reflects greater realism of the industrial data experience, it also makes it impossible to construct exact confidence bounds for

x_p and C_{pk}. For the latter we are aware of no attempts. For tolerance bounds several attempts have been made, with various degrees of numerical complexity, see Seeger and Thorsson (1972), Mee and Owen (1983), and Vangel (1995) who also treats additional regression covariates.

Our intent here is to "reduce" the problem to the i.i.d. case by the simple device of *effective sample size*. As with other methods we can only hope for achieved confidence levels that are approximate. The validity of this approximation is checked via simulations and contrasted with the treatment that ignores batch effects altogether. The appeal of this method is its conceptual simplicity and the reduction to a methodology with available tables and that already is widely spread in the industrial quality assurance practice.

We start out by giving the rationale for the *effective sample size*, which depends on the within batch correlation ρ, and show how to estimate it in straightforward fashion. This is followed by confidence bound construction for x_p, either exactly or approximately, for the two extreme cases: $(\sigma_e > 0, \sigma_b = 0)$ or $\rho = 0$ and $(\sigma_e = 0, \sigma_b > 0)$ or $\rho = 1$. The resulting bounds are further simplified so that they only differ in one parameter which can be identified with the effective sample size $N^\star$. The cases between these two extremes can then be interpolated using the effective sample size and using the existing tables from the i.i.d. case. This process is repeated, but more from a testing perspective, for C_{pk}. For this latter case we present some simulation results for validation and give a sample calculation using a composite material strength data set.

24.2 Effective Sample Size and its Estimation

The extreme case $(\sigma_e > 0,\ \sigma_b = 0)$ or $\rho = 0$ reduces the assumed batch data structure to $N = n_1 + \ldots + n_B$ i.i.d. observations, i.e., the effective sample size is $N^\star = N$. The other extreme case $(\sigma_e = 0,\ \sigma_b > 0)$ or $\rho = 1$ leaves us with effectively $N^\star = B$ i.i.d. observations $X_{11}, X_{21}, \ldots, X_{B1}$, since the remaining observations are just copies of those in this independent set and are of no use.

This suggests that we use an *effective sample size* $N^\star \in [B, N]$ for the intermediate cases $0 < \rho < 1$ in the following sense. We aim to approximate the given batch data set by a fictitious i.i.d. data set $X_1^\star, \ldots, X_{N^\star}^\star$, with $X_i^\star \sim \mathcal{N}(\mu, \sigma^2)$, that in some sense carries the same amount of information. Hence each individual observation in either sample has the same distribution but whereas $\{X_{ij}\}$ has sample size N with complex batch structure, the fictitious sample has the simple i.i.d. structure but with effective sample size $N^\star$.

The above vague notion of "carrying the same amount of information" could be made precise in several different ways. Here we choose $N^\star$ to match the variances of $\bar{X} = \sum_{i=1}^{B} \sum_{j=1}^{n_i} X_{ij}/N$ and $\bar{X}^\star = \sum_{i=1}^{N^\star} X_i^\star/N^\star$, i.e., find $N^\star$ such that

$$\operatorname{var}\left(\bar{X}\right) = \sigma_b^2 \sum_{i=1}^{B} \left(\frac{n_i}{N}\right)^2 + \sigma_e^2 \frac{1}{N} = \operatorname{var}\left(\bar{X}^\star\right) = \frac{\sigma_b^2 + \sigma_e^2}{N^\star} .$$

This leads to the following formula for $N^\star = N^\star(\rho)$

$$N^\star = \left[\frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2} \sum_{i=1}^{B} \left(\frac{n_i}{N}\right)^2 + \frac{1}{N} \frac{\sigma_e^2}{\sigma_b^2 + \sigma_e^2} \right]^{-1} = \left[\rho \, \frac{1}{f+1} + (1-\rho) \, \frac{1}{N} \right]^{-1} ,$$

where we write $1/(f+1) = \sum_{i=1}^{B} (n_i/N)^2$ for reasons to become clear later. For $\rho = 0$ this becomes $N^\star = N$ and for $\rho = 1$ we get $N^\star = f + 1$ which matches B when $n_1 = \ldots = n_B$. Thus in the latter case of equal batch sizes this effective sample size formula agrees with our previous notion. We will not bother with the fact that $N^\star$ may not be an integer. An actual fictitious sample $X_1^\star, \ldots, X_{N^\star}^\star$ is never used in our procedure and all calculations are based on the actual batch data $\{X_{ij}\}$.

In practice the within batch correlation ρ is unknown but one may find reasonable estimates from the data as follows. Compute the between batch and error sums of squares

$$SS_b = \sum_{i=1}^{B} n_i (\bar{X}_{i\cdot} - \bar{X})^2 \qquad \text{and} \qquad SS_e = \sum_{i=1}^{B} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i\cdot})^2 .$$

Take $\hat{\sigma}_e^2 = SS_e/(N-B)$ as unbiased estimate of σ_e^2 and $\hat{\kappa}^2 = SS_b/(B-1)$ as unbiased estimate of

$$\kappa^2 = \sigma_e^2 + \sigma_b^2 \, \frac{N}{B-1} \left(1 - \sum_{i=1}^{B} \left(\frac{n_i}{N}\right)^2 \right) = \sigma_e^2 + \sigma_b^2 \, \frac{N}{B-1} \, \frac{f}{f+1} .$$

Combining these two estimates we get $\hat{\sigma}_b^2 = \left(\hat{\kappa}^2 - \hat{\sigma}_e^2\right) (B-1)(f+1)/(N\ f)$ as unbiased estimate for σ_b^2. Unfortunately, this latter estimate may be negative. If that happens it is suggested to set the estimate to zero. We denote this modification again by $\hat{\sigma}_b^2$ but it will no longer be unbiased. The estimate of ρ is then computed as $\hat{\rho} = \hat{\sigma}_b^2/(\hat{\sigma}_b^2 + \hat{\sigma}_e^2)$. It is this estimate that is used in place of ρ in estimating $N^\star$ by $\hat{N}^\star = N^\star(\hat{\rho})$.

The notion of "effective sample size" is not new although it is not clear whether we have the earliest references. A recent one is Fisher and Van Belle (1993) (p. 828) when interpreting the information loss in the Kaplan-Meier estimate due to censoring. Earlier references, provided kindly by Thomas Lumley, are Kish (1965) (p. 162, p. 259) interpreting design effects with simple random sampling and Skinner, et al. (1989) who view the same issue from the perspective of misspecification.

24.3 Tolerance Bounds

Let $x_p = \mu + z_p\,\sigma$ denote the p-quantile of the sampled $\mathcal{N}(\mu, \sigma^2)$ population. Here $z_p = \Phi^{-1}(p)$ is the p-quantile of the standard normal population. It is desired to find lower confidence or lower tolerance bounds for x_p based on the batch data $\{X_{ij}\}$. We will approach this problem by first examining two extreme situations, namely $(\sigma_b = 0, \sigma_e > 0)$, i.e., no between batch variation, and $(\sigma_b > 0, \sigma_e = 0)$, i.e., no within batch variation, and then interpolate all intermediate situations using the effective sample size.

24.3.1 No Between Batch Variation

Here we assume $\sigma_b = 0$ and $\sigma_e > 0$, i.e, $\rho = 0$, and thus all observations X_{ij} are mutually independent. $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/N)$ and $SS_T = SS_b + SS_e \sim \sigma^2 \cdot \chi^2_{N-1}$ and both are independent of each other. In the following let

$$Z = \sqrt{N}\,\frac{\bar{X} - \mu}{\sigma} \quad \text{and} \quad V = \frac{S}{\sigma}\,, \quad \text{where} \quad S = \sqrt{\frac{SS_T}{N-1}}\,.$$

We consider $100\gamma\%$ lower tolerance bounds of the form $\bar{X} - k\,S$, where the factor k is determined such that

$$\gamma = P\left(\bar{X} - k\,S \le x_p\right) = P\left(\frac{Z - z_p\sqrt{N}}{V} \le k\sqrt{N}\right) = P\left(T_{N-1,-z_p\sqrt{N}} \le k\sqrt{N}\right)\,,$$

where $T_{N-1,-z_p\sqrt{N}}$ represents a noncentral Student t random variable with noncentrality parameter $-z_p\sqrt{N}$ and $N-1$ degrees of freedom. This results in the following expression for the factor k:

$$k = k_0(N) = \frac{1}{\sqrt{N}}\,t_{N-1,-z_p\sqrt{N},\gamma} = \sqrt{\frac{N-1}{N}}\,\frac{1}{\sqrt{N-1}}\,t_{N-1,-z_p\sqrt{N},\gamma}\,,$$

where $t_{N-1,-z_p\sqrt{N},\gamma}$ is the γ quantile of $T_{N-1,-z_p\sqrt{N}}$.

24.3.2 No Within Batch Variation

Here we assume $\sigma_b > 0$ and $\sigma_e = 0$, i.e, $\rho = 1$, and thus $\sigma^2 = \sigma_b^2$ and all observations within each batch are identical. Hence $SS_e = 0$, and thus $S^2 = SS_b/(N-1)$. Using Satterthwaite's method we will approximate the distribution of $SS_T = SS_b$ by a chi-square multiple with g degrees of freedom, i.e., $SS_T = SS_b \approx a \cdot \chi^2_g$, where a and g are determined to match the first two moments on either side. This leads to

$$g = \frac{(1-\sum w_i^2)^2}{\sum w_i^2 - 2\sum w_i^3 + (\sum w_i^2)^2} \quad \text{and} \quad a = \frac{N}{g}\,\sigma_b^2\,\left(1 - \sum_{i=1}^{B} w_i^2\right)\,,$$

where $w_i = n_i/N$. In the Appendix it is shown that this complicated expression for g can be approximated very well by a much simpler expression, namely by $f = \left(\sum w_i^2\right)^{-1} - 1$, and the approximation is exact when the n_i are all the same. We will use this simplification (f replacing g) from now on since it leads to a convenient similarity of the formulas for the factor k in the two cases studied. With this simplification we have $a \approx N\,\sigma_b^2/(f+1)$ and we can treat

$$V^2 = \frac{SS_T}{a\ f} = S^2\ \frac{(N-1)(f+1)}{f\ N\ \sigma_b^2}$$

as an approximate χ_f^2/f random variable. Further, $\bar{X} \sim \mathcal{N}(\mu, \tau^2)$ with $\tau^2 = \sigma_b^2 \cdot \sum_{i=1}^{B} w_i^2 \approx \sigma_b^2/(f+1)$, i.e., $Z = \sqrt{f+1}\ (\bar{X}-\mu)/\sigma_b$ has a standard normal distribution.

Note that when all samples sizes n_i are the same ($= n$), then the above complicated expressions for f and a (and their approximations) reduce to $f = B-1$ and $a = n\sigma_b^2$. In that case SS_b actually is exactly distributed like $n\sigma_b^2 \cdot \chi_{B-1}^2$ and then $SS_T = SS_b$ is independent of $\bar{X}$. When the samples sizes are not the same, then SS_T is approximately distributed like the above chi-square multiple and the strict independence property no longer holds. We will ignore this latter flaw in our derivation below. The simulations show that this is of no serious consequence.

Again we have

$$\begin{aligned}\gamma = P\left(\bar{X} - k\ S \le x_p\right) &= P\left(\frac{Z - z_p\sqrt{f+1}}{V} \le k\sqrt{\frac{f\ N}{N-1}}\right) \\ &= P\left(T_{f,-z_p\sqrt{f+1}} \le k\sqrt{\frac{f\ N}{N-1}}\right)\end{aligned}$$

leading to

$$k = k_1(N) = \sqrt{\frac{N-1}{N}}\ \frac{1}{\sqrt{f}}\ t_{f,-z_p\sqrt{f+1},\gamma}\ .$$

24.3.3 The Interpolation Step

We note that the two expressions for $k_0(N)$ and $k_1(N)$ share the common factor $\sqrt{(N-1)/N}$ and the remainder can be matched if we match $f+1$ and N. We propose to use the previously developed estimated effective sample size $\hat{N}^\star$ as a

simple interpolation between $f+1$ and N and use as k-factor in the general case

$$k^\star(N) = \sqrt{\frac{N-1}{N}}\ \frac{1}{\sqrt{\hat{N}^\star - 1}}\ t_{\hat{N}^\star - 1, -z_p\sqrt{\hat{N}^\star}, \gamma}\ .$$

24.4 Confidence Bounds for C_L, C_U and C_{pk}

For lower and upper specification limits L and U define

$$C_L = \frac{\mu - L}{3\sigma}\ , \quad C_U = \frac{U-\mu}{3\sigma} \quad \text{and} \quad C_{pk} = \min(C_L, C_U)\ .$$

These process capability indices are unknown but can be estimated respectively by

$$\widehat{C}_L = \frac{\bar{X} - L}{3S}\ , \quad \widehat{C}_U = \frac{U - \bar{X}}{3S} \quad \text{and} \quad \widehat{C}_{pk} = \min\left(\widehat{C}_L, \widehat{C}_U\right)\ .$$

Here S is again the sample standard deviation of all the data, i.e, $S^2 = (SS_b + SS_e)/(N-1)$. We want to use these estimates $\widehat{C}_L$, $\widehat{C}_U$, and $\widehat{C}_{pk}$ in order to decide whether the corresponding population parameters exceed a given threshold C_0. This can be accomplished either by constructing lower confidence bounds based on these estimates or by testing of appropriate hypotheses. Since the available tables so far favor the testing framework we will stay with that preference, but we will indicate confidence bounds at the appropriate places.

We focus on C_L (C_U is handled the same way) and then combine the results for C_{pk}. Consider the problem of testing the hypothesis $H_L(C_0) : C_L \le C_0$ against the alternative $K_L(C_0) : C_L > C_0$. We will reject $H_L(C_0)$ at level α whenever $\widehat{C}_L \ge C_\star$, where $C_\star = C_\star(\alpha, C_0)$ is determined such that the maximal chance of $\widehat{C}_L$ exceeding $C_\star$ is α when the hypothesis is true. Clearly $C_\star(\alpha, C_0)$ is an increasing function of C_0 and thus has an inverse $C_\star^{-1}(\alpha, \cdot)$. Solving $C_\star(\alpha, C_0) = \widehat{C}_L$ for $C_0 = C_\star^{-1}(\alpha, \widehat{C}_L)$ will give us a $100(1-\alpha)\%$ lower confidence bound $\widehat{C}_L(1-\alpha) = C_\star^{-1}(\alpha, \widehat{C}_L)$ for C_L. By this construction $\widehat{C}_L(1-\alpha) > C_0$ means that we should reject $H_L(C_0)$. Similarly $\widehat{C}_U(1-\alpha) = C_\star^{-1}(\alpha, \widehat{C}_U)$ is a $100(1-\alpha)\%$ lower confidence bound for C_U and $\widehat{C}_{pk}(1-\alpha) = \min(\widehat{C}_L(1-\alpha), \widehat{C}_U(1-\alpha))$ is a $100(1-\alpha)\%$ lower confidence bound for C_{pk}. The latter is easily seen by letting σ get arbitrarily small so that the two-sided problem reduces to the one-sided one, see also Kushler and Hurley (1992).

The main problem now is to find the proper critical value $C_\star$. We will do this again by examining the two extreme situations ($\sigma_b > 0, \sigma_e = 0$) and ($\sigma_b = 0, \sigma_e > 0$). All other situations will then be dealt with by a simple interpolation scheme. Finally, the resulting procedure is examined via simulations.

24.4.1 No Between Batch Variation

Here we assume again ($\sigma_b = 0,\ \sigma_e > 0$). Thus $\sigma = \sigma_e$ and all X_{ij} are mutally independent. $\bar{X}$ is normally distributed with mean μ and variance σ^2/N, SS_T is

distributed as $\sigma^2 \cdot \chi^2_{N-1}$ and both are independent of each other. Adopting the notation that P_{C_0} denotes a probability distribution under (μ, σ) with $C_L = C_0$ we find $C_\star$ by solving

$$\alpha = P_{C_0}\left(\widehat{C}_L \geq C_\star\right) = P_{C_0}\left(\frac{\bar{X} - L}{3S} \geq C_\star\right) = P\left(T_{N-1,3C_0\sqrt{N}} \geq 3C_\star\sqrt{N}\right)$$

which yields

$$C_\star = \frac{1}{3\sqrt{N}}\, t_{N-1,3C_0\sqrt{N},1-\alpha} = \sqrt{\frac{N-1}{N}}\, \frac{1}{3\sqrt{N-1}}\, t_{N-1,3C_0\sqrt{N},1-\alpha}\,.$$

24.4.2 No Within Batch Variation

Here we assume again ($\sigma_b > 0,\ \sigma_e = 0$) and use the same notation and approximations developed in the corresponding section on tolerance bounds. The α requirement on $C_\star$ leads to

$$\begin{aligned}\alpha &= P_{C_0}\left(\widehat{C}_L \geq C_\star\right) = P_{C_0}\left(\frac{\bar{X} - L}{3S} \geq C_\star\right) \\ &= P\left(\frac{Z+\delta}{V} \geq 3C_\star\sqrt{N/(N-1)}\sqrt{f}\right) \approx P\left(T_{f,\delta} \geq 3C_\star\sqrt{N/(N-1)}\sqrt{f}\right)\,,\end{aligned}$$

where $T_{f,\delta}$ is a noncentral Student t random variable with f degrees of freedom and noncentrality parameter $\delta = 3C_0\sqrt{f+1}$. This yields the following expression for $C_\star$

$$C_\star = \sqrt{\frac{N-1}{N}}\, \frac{1}{3\sqrt{f}}\, t_{f,3C_0\sqrt{f+1},1-\alpha}\,,$$

where $t_{f,\delta,1-\alpha}$ represents the $1-\alpha$ percentile of that noncentral Student t distribution.

24.4.3 The Interpolation Step

Note that the two formulas for $C_\star$, developed for the two extreme cases, share the factor $\sqrt{(N-1)/N}$ and the remainder can be matched if we match $f+1$ and N. We propose to use the previously developed effective sample size $N^\star$ as a simple interpolation between $f+1$ and N, namely

$$N^\star = \left[\hat{\rho}\, \frac{1}{f+1} + (1-\hat{\rho})\, \frac{1}{N}\right]^{-1}$$

and use as critical point in the general case

$$C_\star = \sqrt{\frac{N-1}{N}}\ \frac{1}{3\sqrt{N^\star - 1}}\ t_{N^\star - 1, 3C_0\sqrt{N^\star}, 1-\alpha}\ .$$

Table 3 of Chou et al. (1990) gives the value of

$$C_{\text{Table}\,\star}(N) = \sqrt{\frac{N-1}{N}}\ \frac{1}{3\sqrt{N-1}} t_{N-1, 3C_0\sqrt{N}, 1-\alpha} \quad \text{for } \alpha = .05,$$

for various values of $C_0 = .7, .8, \ldots, 2.0$ and $N = 10, 20, \ldots, 50, 75, 100, 125, 150, 200, 300, 350, 400$. These tabled values are correct when $\sigma_b = 0$, i.e, in the i.i.d. case, which was addressed by Chou et al. and then clarified by Kushler and Hurley (1992). The same table, covering a somewhat different grid, and additional tables for $\alpha = .20, .10, .01$ are given in Tables 24.1-24.4.

To allow for the possible batch effect we should, according to the above derivation, use instead the adjusted critical value

$$C_{\text{Adj}\,\star}(N) = \sqrt{\frac{N-1}{N}}\ \sqrt{\frac{N^\star}{N^\star - 1}}\ C_{\text{Table}\,\star}(N^\star).$$

This concludes the derivation of the critical point $C_\star$ for our hypothesis testing problem concerning C_L. The same $C_\star$ in conjunction with $\widehat{C}_U$ works for the testing the hypothesis $H_U : C_U \leq C_0$ against the alternative $K_U : C_U > C_0$.

To combine these two procedures into one for testing the corresponding hypothesis for C_{pk}, namely $H : C_{pk} \leq C_0$ versus $K : C_{pk} > C_0$, we simply reject H when $\widehat{C}_{pk} \geq C_\star$ with $C_\star = C_{\text{Adj}\,\star}(N)$ as developed previously. Upon rejection of H we can be at least $100(1-\alpha)\%$ confident that $C_{pk} > C_0$. When $C_L = C_U$ the confidence will be slightly higher than the target of $100(1-\alpha)\%$ but when C_L and C_U are quite different, the confidence will be approximately equal to $100(1-\alpha)\%$. Not knowing the actual values of C_L and C_U and wanting to use the simple estimated value $\widehat{C}_{pk}$ as a decision criterion, this procedure should serve its purpose reasonably well.

Table 24.1: $\alpha = .20$ or 80% confidence

sample size	**Critical Values $C_\star$** **$\alpha = .20$** C_0											
	1.00	1.10	1.20	1.30	1.33	1.40	1.50	1.60	1 .70	1.80	1.90	2.00
500	1.03	1.13	1.24	1.34	1.37	1.44	1.54	1.65	1.75	1.85	1.95	2.06
300	1.04	1.14	1.25	1.35	1.38	1.45	1.56	1.66	1.76	1.87	1.97	2.08
250	1.04	1.15	1.25	1.36	1.39	1.46	1.56	1.67	1.77	1.87	1.98	2.08
200	1.05	1.15	1.26	1.36	1.40	1.47	1.57	1.68	1.78	1.88	1.99	2.09
175	1.05	1.16	1.26	1.37	1.40	1.47	1.58	1.68	1.79	1.89	2.00	2.10
150	1.06	1.16	1.27	1.37	1.41	1.48	1.58	1.69	1.79	1.90	2.00	2.11
125	1.06	1.17	1.28	1.38	1.42	1.49	1.59	1.70	1.80	1.91	2.01	2.12
100	1.07	1.18	1.29	1.39	1.43	1.50	1.60	1.71	1.82	1.92	2.03	2.14
90	1.08	1.18	1.29	1.40	1.43	1.50	1.61	1.72	1.82	1.93	2.04	2.14
80	1.08	1.19	1.30	1.40	1.44	1.51	1.62	1.73	1.83	1.94	2.05	2.15
70	1.09	1.20	1.30	1.41	1.45	1.52	1.63	1.74	1.84	1.95	2.06	2.17
60	1.10	1.21	1.31	1.42	1.46	1.53	1.64	1.75	1.86	1.97	2.07	2.18
50	1.11	1.22	1.33	1.44	1.47	1.55	1.66	1.76	1.87	1.98	2.09	2.20
46	1.11	1.22	1.33	1.44	1.48	1.55	1.66	1.77	1.88	1.99	2.10	2.21
42	1.12	1.23	1.34	1.45	1.49	1.56	1.67	1.78	1.89	2.00	2.11	2.23
38	1.13	1.24	1.35	1.46	1.50	1.57	1.68	1.79	1.91	2.02	2.13	2.24
34	1.14	1.25	1.36	1.47	1.51	1.58	1.70	1.81	1.92	2.03	2.14	2.26
30	1.15	1.26	1.37	1.49	1.52	1.60	1.71	1.82	1.94	2.05	2.16	2.28
28	1.15	1.27	1.38	1.49	1.53	1.61	1.72	1.83	1.95	2.06	2.18	2.29
26	1.16	1.27	1.39	1.50	1.54	1.62	1.73	1.85	1.96	2.07	2.19	2.30
24	1.17	1.28	1.40	1.51	1.55	1.63	1.74	1.86	1.97	2.09	2.20	2.32
22	1.18	1.29	1.41	1.53	1.56	1.64	1.76	1.87	1.99	2.11	2.22	2.34
20	1.19	1.31	1.42	1.54	1.58	1.66	1.77	1.89	2.01	2.13	2.24	2.36
19	1.20	1.31	1.43	1.55	1.59	1.67	1.78	1.90	2.02	2.14	2.26	2.37
18	1.20	1.32	1.44	1.56	1.60	1.68	1.79	1.91	2.03	2.15	2.27	2.39
17	1.21	1.33	1.45	1.57	1.61	1.69	1.81	1.93	2.04	2.16	2.28	2.40
16	1.22	1.34	1.46	1.58	1.62	1.70	1.82	1.94	2.06	2.18	2.30	2.42
15	1.23	1.35	1.47	1.59	1.63	1.71	1.83	1.96	2.08	2.20	2.32	2.44
14	1.24	1.36	1.49	1.61	1.65	1.73	1.85	1.97	2.10	2.22	2.34	2.46
13	1.26	1.38	1.50	1.62	1.67	1.75	1.87	1.99	2.12	2.24	2.36	2.49
12	1.27	1.40	1.52	1.64	1.69	1.77	1.89	2.02	2.14	2.27	2.39	2.52
11	1.29	1.42	1.54	1.67	1.71	1.79	1.92	2.05	2.17	2.30	2.43	2.55
10	1.31	1.44	1.57	1.70	1.74	1.82	1.95	2.08	2.21	2.34	2.47	2.60
9	1.34	1.47	1.60	1.73	1.77	1.86	1.99	2.12	2.25	2.39	2.52	2.65
8	1.37	1.51	1.64	1.77	1.82	1.91	2.04	2.18	2.31	2.45	2.58	2.72
7	1.42	1.56	1.70	1.83	1.88	1.97	2.11	2.25	2.39	2.53	2.67	2.81
6	1.48	1.63	1.77	1.92	1.96	2.06	2.21	2.35	2.50	2.64	2.79	2.93
5	1.58	1.73	1.89	2.04	2.09	2.20	2.35	2.51	2.66	2.82	2.97	3.13
4	1.75	1.92	2.09	2.26	2.32	2.44	2.61	2.78	2.95	3.12	3.29	3.47
3	2.14	2.35	2.56	2.77	2.84	2.98	3.19	3.40	3.61	3.82	4.03	4.24
2	3.94	4.34	4.73	5.13	5.26	5.52	5.92	6.31	6.71	7.10	7.50	7.89

Table 24.2: $\alpha = .10$ or 90% confidence

sample size	**Critical Values $C_\star$** **$\alpha = .10$** C_0											
	1.00	1.10	1.20	1.30	1.33	1.40	1.50	1.60	1.70	1.80	1.90	2.00
500	1.05	1.15	1.25	1.36	1.39	1.46	1.57	1.67	1.78	1.88	1.98	2.09
300	1.06	1.17	1.27	1.38	1.41	1.48	1.59	1.69	1.80	1.90	2.01	2.11
250	1.07	1.17	1.28	1.39	1.42	1.49	1.60	1.70	1.81	1.91	2.02	2.13
200	1.08	1.18	1.29	1.40	1.43	1.50	1.61	1.72	1.82	1.93	2.04	2.14
175	1.08	1.19	1.30	1.40	1.44	1.51	1.62	1.72	1.83	1.94	2.05	2.15
150	1.09	1.20	1.30	1.41	1.45	1.52	1.63	1.74	1.84	1.95	2.06	2.17
125	1.10	1.21	1.32	1.42	1.46	1.53	1.64	1.75	1.86	1.97	2.08	2.19
100	1.11	1.22	1.33	1.44	1.48	1.55	1.66	1.77	1.88	1.99	2.10	2.21
90	1.12	1.23	1.34	1.45	1.49	1.56	1.67	1.78	1.89	2.00	2.11	2.22
80	1.13	1.24	1.35	1.46	1.50	1.57	1.68	1.79	1.90	2.02	2.13	2.24
70	1.14	1.25	1.36	1.47	1.51	1.58	1.70	1.81	1.92	2.03	2.15	2.26
60	1.15	1.26	1.38	1.49	1.53	1.60	1.72	1.83	1.94	2.06	2.17	2.28
50	1.17	1.28	1.40	1.51	1.55	1.63	1.74	1.85	1.97	2.08	2.20	2.31
46	1.18	1.29	1.41	1.52	1.56	1.64	1.75	1.87	1.98	2.10	2.21	2.33
42	1.19	1.30	1.42	1.53	1.57	1.65	1.77	1.88	2.00	2.12	2.23	2.35
38	1.20	1.31	1.43	1.55	1.59	1.67	1.78	1.90	2.02	2.14	2.25	2.37
34	1.21	1.33	1.45	1.57	1.61	1.69	1.80	1.92	2.04	2.16	2.28	2.40
30	1.23	1.35	1.47	1.59	1.63	1.71	1.83	1.95	2.07	2.19	2.31	2.43
28	1.24	1.36	1.48	1.60	1.64	1.72	1.84	1.97	2.09	2.21	2.33	2.45
26	1.25	1.37	1.49	1.62	1.66	1.74	1.86	1.98	2.11	2.23	2.35	2.47
24	1.26	1.39	1.51	1.63	1.67	1.76	1.88	2.00	2.13	2.25	2.38	2.50
22	1.28	1.40	1.53	1.65	1.69	1.78	1.90	2.03	2.15	2.28	2.40	2.53
20	1.30	1.42	1.55	1.68	1.72	1.80	1.93	2.06	2.18	2.31	2.44	2.57
19	1.31	1.44	1.56	1.69	1.73	1.82	1.95	2.07	2.20	2.33	2.46	2.59
18	1.32	1.45	1.58	1.71	1.75	1.83	1.96	2.09	2.22	2.35	2.48	2.61
17	1.33	1.46	1.59	1.72	1.77	1.85	1.98	2.11	2.24	2.37	2.50	2.63
16	1.35	1.48	1.61	1.74	1.78	1.87	2.00	2.14	2.27	2.40	2.53	2.66
15	1.37	1.50	1.63	1.76	1.81	1.89	2.03	2.16	2.29	2.43	2.56	2.69
14	1.38	1.52	1.65	1.79	1.83	1.92	2.06	2.19	2.33	2.46	2.60	2.73
13	1.41	1.54	1.68	1.81	1.86	1.95	2.09	2.22	2.36	2.50	2.64	2.77
12	1.43	1.57	1.71	1.85	1.89	1.99	2.13	2.26	2.40	2.54	2.68	2.82
11	1.46	1.60	1.74	1.89	1.93	2.03	2.17	2.31	2.45	2.60	2.74	2.88
10	1.50	1.64	1.79	1.93	1.98	2.08	2.22	2.37	2.52	2.66	2.81	2.95
9	1.55	1.69	1.84	1.99	2.04	2.14	2.29	2.44	2.59	2.74	2.89	3.04
8	1.61	1.76	1.91	2.07	2.12	2.22	2.38	2.54	2.69	2.85	3.00	3.16
7	1.69	1.85	2.01	2.17	2.23	2.34	2.50	2.66	2.83	2.99	3.15	3.32
6	1.80	1.97	2.15	2.32	2.38	2.49	2.67	2.84	3.02	3.19	3.37	3.54
5	1.98	2.17	2.36	2.55	2.62	2.74	2.94	3.13	3.32	3.51	3.71	3.90
4	2.31	2.53	2.76	2.98	3.05	3.20	3.43	3.65	3.88	4.10	4.33	4.55
3	3.13	3.43	3.73	4.04	4.14	4.35	4.65	4.96	5.26	5.57	5.88	6.18
2	7.95	8.75	9.55	10.3	10.6	11.1	11.9	12.7	13.5	14.3	15.1	15.9

Table 24.3: $\alpha = .05$ or 95% confidence

	Critical Values $C_\star$											
	$\alpha = .05$											
sample size	**C_0**											
	1.00	1.10	1.20	1.30	1.33	1.40	1.50	1.60	1.70	1.80	1.90	2.00
500	1.06	1.17	1.27	1.38	1.41	1.48	1.59	1.69	1.80	1.90	2.01	2.11
300	1.08	1.19	1.29	1.40	1.44	1.51	1.61	1.72	1.83	1.93	2.04	2.15
250	1.09	1.20	1.30	1.41	1.45	1.52	1.63	1.73	1.84	1.95	2.06	2.16
200	1.10	1.21	1.32	1.42	1.46	1.53	1.64	1.75	1.86	1.97	2.08	2.19
175	1.11	1.22	1.32	1.43	1.47	1.54	1.65	1.76	1.87	1.98	2.09	2.20
150	1.12	1.23	1.34	1.45	1.48	1.56	1.67	1.78	1.89	2.00	2.11	2.22
125	1.13	1.24	1.35	1.46	1.50	1.57	1.68	1.80	1.91	2.02	2.13	2.24
100	1.15	1.26	1.37	1.48	1.52	1.60	1.71	1.82	1.93	2.05	2.16	2.27
90	1.16	1.27	1.38	1.49	1.53	1.61	1.72	1.84	1.95	2.06	2.18	2.29
80	1.17	1.28	1.39	1.51	1.55	1.62	1.74	1.85	1.97	2.08	2.20	2.31
70	1.18	1.29	1.41	1.53	1.56	1.64	1.76	1.87	1.99	2.10	2.22	2.34
60	1.20	1.31	1.43	1.55	1.59	1.66	1.78	1.90	2.02	2.13	2.25	2.37
50	1.22	1.34	1.46	1.58	1.62	1.70	1.81	1.93	2.05	2.17	2.29	2.41
46	1.23	1.35	1.47	1.59	1.63	1.71	1.83	1.95	2.07	2.19	2.31	2.43
42	1.24	1.37	1.49	1.61	1.65	1.73	1.85	1.97	2.09	2.22	2.34	2.46
38	1.26	1.38	1.50	1.63	1.67	1.75	1.87	2.00	2.12	2.24	2.37	2.49
34	1.28	1.40	1.53	1.65	1.69	1.78	1.90	2.03	2.15	2.28	2.40	2.53
30	1.30	1.43	1.56	1.68	1.72	1.81	1.94	2.06	2.19	2.32	2.44	2.57
28	1.32	1.44	1.57	1.70	1.74	1.83	1.96	2.08	2.21	2.34	2.47	2.60
26	1.33	1.46	1.59	1.72	1.76	1.85	1.98	2.11	2.24	2.37	2.50	2.63
24	1.35	1.48	1.61	1.74	1.79	1.87	2.01	2.14	2.27	2.40	2.53	2.66
22	1.37	1.51	1.64	1.77	1.82	1.90	2.04	2.17	2.30	2.44	2.57	2.71
20	1.40	1.53	1.67	1.80	1.85	1.94	2.08	2.21	2.35	2.48	2.62	2.76
19	1.41	1.55	1.69	1.82	1.87	1.96	2.10	2.23	2.37	2.51	2.65	2.78
18	1.43	1.57	1.71	1.84	1.89	1.98	2.12	2.26	2.40	2.54	2.68	2.82
17	1.45	1.59	1.73	1.87	1.91	2.01	2.15	2.29	2.43	2.57	2.71	2.85
16	1.47	1.61	1.75	1.89	1.94	2.04	2.18	2.32	2.46	2.61	2.75	2.89
15	1.49	1.64	1.78	1.92	1.97	2.07	2.21	2.36	2.50	2.65	2.79	2.94
14	1.52	1.67	1.81	1.96	2.01	2.10	2.25	2.40	2.55	2.69	2.84	2.99
13	1.55	1.70	1.85	2.00	2.05	2.15	2.30	2.45	2.60	2.75	2.90	3.05
12	1.59	1.74	1.89	2.04	2.10	2.20	2.35	2.50	2.66	2.81	2.97	3.12
11	1.63	1.79	1.94	2.10	2.15	2.26	2.42	2.57	2.73	2.89	3.05	3.21
10	1.69	1.85	2.01	2.17	2.22	2.33	2.50	2.66	2.82	2.98	3.15	3.31
9	1.75	1.92	2.09	2.26	2.31	2.43	2.60	2.77	2.93	3.10	3.27	3.44
8	1.84	2.02	2.20	2.37	2.43	2.55	2.73	2.90	3.08	3.26	3.44	3.62
7	1.96	2.15	2.34	2.53	2.59	2.72	2.91	3.10	3.29	3.48	3.67	3.86
6	2.14	2.35	2.55	2.76	2.83	2.96	3.17	3.38	3.58	3.79	4.00	4.21
5	2.43	2.66	2.90	3.13	3.21	3.36	3.60	3.83	4.07	4.30	4.54	4.77
4	2.99	3.27	3.56	3.85	3.94	4.14	4.42	4.71	5.00	5.29	5.58	5.87
3	4.49	4.92	5.36	5.80	5.94	6.23	6.67	7.11	7.55	7.99	8.43	8.87
2	15.9	17.5	19.1	20.7	21.3	22.3	23.9	25.5	27.1	28.7	30.3	31.9

Table 24.4: $\alpha = .01$ or 99% confidence

sample size	**Critical Values $C_\star$** **$\alpha = .01$** C_0											
	1.00	1.10	1.20	1.30	1.33	1.40	1.50	1.60	1.70	1.80	1.90	2.00
500	1.09	1.19	1.30	1.41	1.44	1.52	1.62	1.73	1.84	1.95	2.05	2.16
300	1.11	1.22	1.33	1.44	1.48	1.55	1.66	1.77	1.88	1.99	2.10	2.21
250	1.13	1.24	1.35	1.46	1.50	1.57	1.68	1.79	1.90	2.01	2.13	2.24
200	1.14	1.26	1.37	1.48	1.52	1.59	1.71	1.82	1.93	2.04	2.16	2.27
175	1.15	1.27	1.38	1.49	1.53	1.61	1.72	1.83	1.95	2.06	2.18	2.29
150	1.17	1.28	1.40	1.51	1.55	1.63	1.74	1.86	1.97	2.09	2.20	2.32
125	1.19	1.30	1.42	1.54	1.57	1.65	1.77	1.89	2.00	2.12	2.23	2.35
100	1.21	1.33	1.45	1.57	1.61	1.69	1.81	1.93	2.04	2.16	2.28	2.40
90	1.23	1.35	1.47	1.59	1.63	1.71	1.83	1.95	2.07	2.19	2.31	2.43
80	1.24	1.37	1.49	1.61	1.65	1.73	1.85	1.97	2.09	2.22	2.34	2.46
70	1.27	1.39	1.51	1.63	1.67	1.76	1.88	2.00	2.13	2.25	2.37	2.50
60	1.29	1.42	1.54	1.67	1.71	1.79	1.92	2.04	2.17	2.30	2.42	2.55
50	1.33	1.46	1.58	1.71	1.75	1.84	1.97	2.10	2.23	2.36	2.49	2.62
46	1.35	1.47	1.60	1.73	1.78	1.87	2.00	2.13	2.26	2.39	2.52	2.65
42	1.37	1.50	1.63	1.76	1.81	1.89	2.03	2.16	2.29	2.42	2.56	2.69
38	1.39	1.53	1.66	1.79	1.84	1.93	2.06	2.20	2.33	2.47	2.60	2.74
34	1.42	1.56	1.70	1.83	1.88	1.97	2.11	2.25	2.38	2.52	2.66	2.80
30	1.46	1.60	1.74	1.88	1.93	2.02	2.16	2.30	2.45	2.59	2.73	2.87
28	1.48	1.63	1.77	1.91	1.96	2.05	2.20	2.34	2.48	2.63	2.77	2.91
26	1.51	1.65	1.80	1.94	1.99	2.09	2.23	2.38	2.53	2.67	2.82	2.97
24	1.54	1.69	1.84	1.98	2.03	2.13	2.28	2.43	2.58	2.73	2.87	3.02
22	1.58	1.73	1.88	2.03	2.08	2.18	2.33	2.48	2.64	2.79	2.94	3.09
20	1.62	1.77	1.93	2.08	2.14	2.24	2.40	2.55	2.71	2.86	3.02	3.18
19	1.65	1.80	1.96	2.12	2.17	2.28	2.43	2.59	2.75	2.91	3.07	3.23
18	1.67	1.83	1.99	2.15	2.21	2.31	2.47	2.64	2.80	2.96	3.12	3.28
17	1.71	1.87	2.03	2.19	2.25	2.36	2.52	2.69	2.85	3.01	3.18	3.34
16	1.74	1.91	2.07	2.24	2.30	2.41	2.57	2.74	2.91	3.08	3.24	3.41
15	1.78	1.95	2.12	2.29	2.35	2.46	2.63	2.81	2.98	3.15	3.32	3.49
14	1.83	2.01	2.18	2.35	2.41	2.53	2.70	2.88	3.06	3.23	3.41	3.59
13	1.89	2.07	2.25	2.43	2.49	2.61	2.79	2.97	3.15	3.33	3.51	3.70
12	1.96	2.14	2.33	2.51	2.57	2.70	2.89	3.07	3.26	3.45	3.64	3.83
11	2.04	2.23	2.42	2.62	2.68	2.81	3.01	3.20	3.40	3.59	3.79	3.99
10	2.14	2.34	2.55	2.75	2.82	2.95	3.16	3.36	3.57	3.77	3.98	4.19
9	2.27	2.49	2.70	2.92	2.99	3.14	3.35	3.57	3.79	4.01	4.23	4.44
8	2.45	2.68	2.92	3.15	3.23	3.38	3.62	3.85	4.09	4.32	4.56	4.79
7	2.71	2.96	3.22	3.47	3.56	3.73	3.99	4.25	4.51	4.77	5.03	5.29
6	3.10	3.39	3.68	3.98	4.07	4.27	4.57	4.86	5.16	5.46	5.76	6.05
5	3.78	4.13	4.49	4.85	4.97	5.21	5.58	5.94	6.30	6.67	7.03	7.39
4	5.24	5.74	6.24	6.75	6.91	7.25	7.76	8.26	8.77	9.27	9.78	10.3
3	10.1	11.1	12.1	13.1	13.4	14.1	15.1	16.1	17.1	18.1	19.1	20.0
2	79.8	87.8	95.7	103.	106.	112.	120.	128.	136.	144.	152.	160.

24.5 Validation

For $\alpha = .10$ the above procedure was validated through simulation as follows. Normal batch random samples were generated according to the assumed model for various values of ρ, using $\mu = 0$, $\sigma^2 = 1 = \rho + (1 - \rho)$, i.e., $\sigma_b^2 = \rho$ and $\sigma_e^2 = 1 - \rho$. The known value of C_L was assumed to be 1, i.e., $L = -3$. For each generated collection of B batch samples we computed the observed $\widehat{C}_{pk}$ and compared it against the value $C_{\mathrm{Adj}\,\star}(N)$ and also against the value $C_\star$ that would be correct if $\sigma_b = 0$ were correct. Repeating this 1000 times for each configuration of ρ, B and $(n_1, \ldots, n_B)$ we recorded the observed rates of exceeding the respective critical points. This was done for $B = 10, 20, 30, 40$ batches of same size $n = 2, 3, 5$ each for $\rho = 0, .2, .4, .6, .8, 1$. To study the sample size imbalance effect we also simulated $B = 10, 20, 30, 40$ batches, half of the batch samples of size n_1 and half of size n_2 for $(n_1, n_2) = (2, 3), (2, 5)$ and $(3, 5)$.

The resulting observed confidence levels are summarized graphically in Figures 1-2. The three horizontal lines centered on $1 - \alpha = .9$ represent the target confidence for this simulation and 95% uncertainty limits for 1000 replications. The dotted curves represent the observed confidence for the proposed procedure whereas the dashed curves represent the observed confidence of the procedure that ignores batch effects, i.e., treats all observations as mutually independent.

It is quite obvious from the figures that the degradation of the latter procedure can be quite serious even with moderate batch effects, especially for few batches ($B = 10$) of "large" size ($n = 5$). The proposed procedure appears to hold its intended confidence level quite well, being slightly conservative when it is off by a small amount. Also given in these plots are the observed average effective sample sizes for each ρ.

24.6 Sample Calculation

The data Table 24.5 represent data on 21 batches of some composite material property data with lower specification value $L = 45$. From the data in this table we obtain:

$$\bar{X} = 49.638\,, \quad S = 1.320\,, \quad \text{thus} \quad \widehat{C}_L = 1.17\,.$$

Ignoring the batch effects and assuming that we deal with $N = 63$ independent observations we obtain from our Table 3 (using the column corresponding to $C_0 = 1$ and interpolating at the row corresponding to $N = 63$) the critical value 1.147, i.e., we would then conclude with 90% confidence that the true C_{pk} is at least 1 since $1.17 > 1.147$.

Figure 24.1: 1000 Simulations with Batches of Sizes 2 & 5

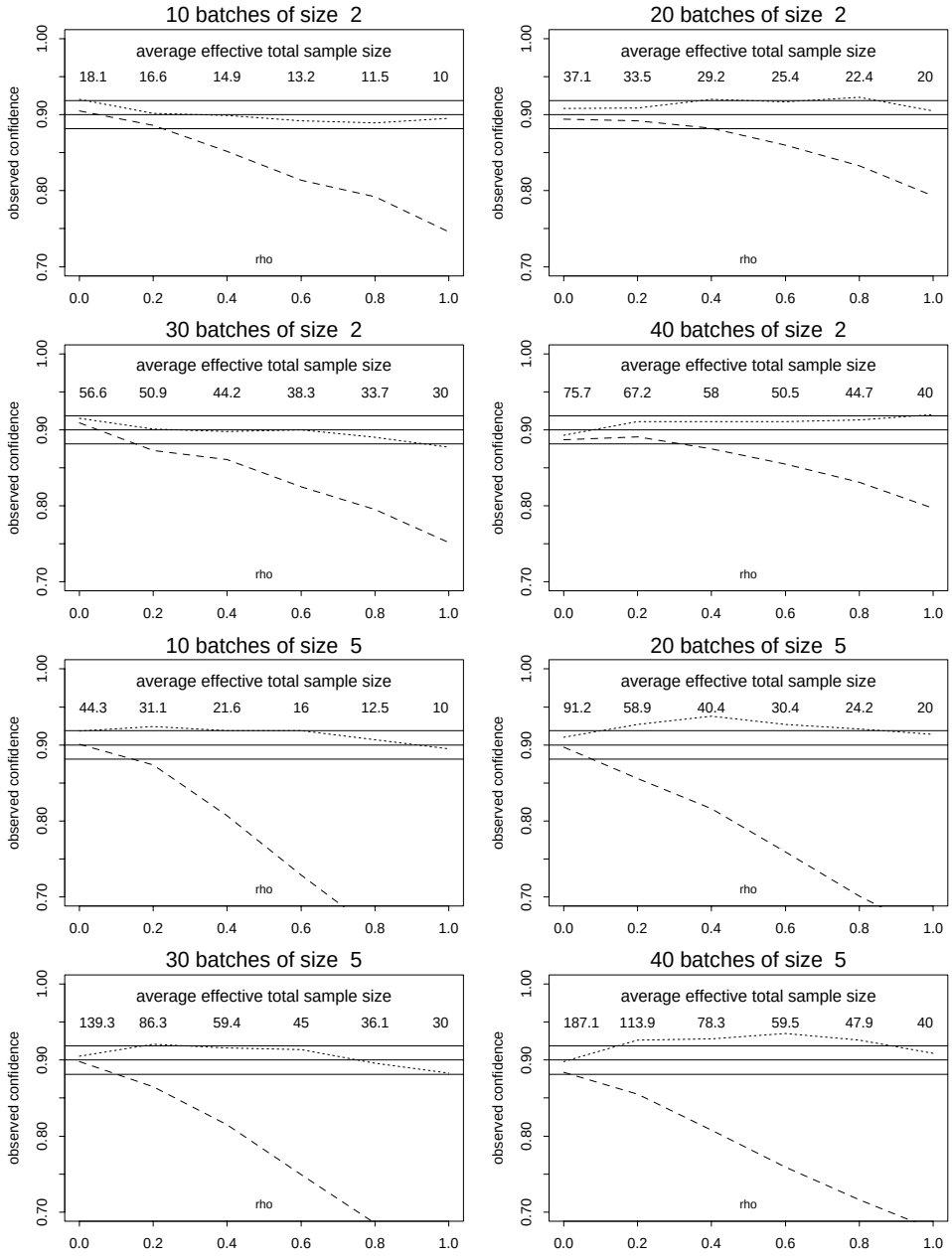

Figure 24.2: 1000 Simulations with Batches of Size 2 & 3 and 3 & 5

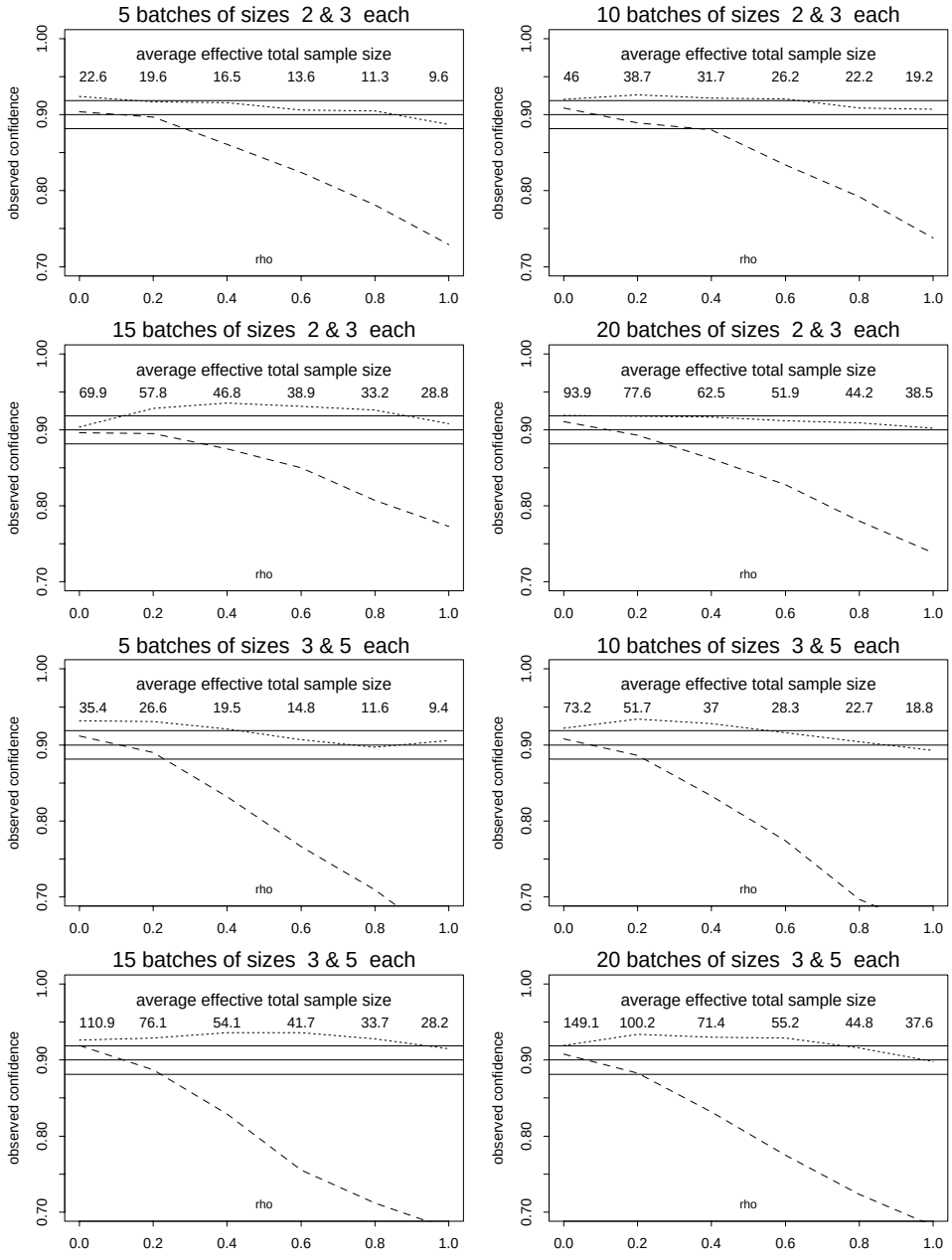

Table 24.5: Example Batch Data

batch	n_i	sample data	sample average
1	1	50.5	50.5
2	1	50.2	50.2
3	4	50.7, 50.8, 51.4, 51.3	51.05
4	1	49.3	49.3
5	3	51.0, 51.2, 53.4	51.867
6	3	50.9, 51.6, 51.8	51.433
7	1	49.3	49.3
8	3	48.6, 48.2, 46.6	47.8
9	2	50.4, 49.9	50.15
10	2	48.2, 47.5	47.85
11	3	50.5, 48.2, 49.5	49.4
12	3	49.7, 51.4, 50.6	50.567
13	4	49.6, 51.1, 51.1, 52.5	51.075
14	4	48.4, 50.2, 48.8, 49.1	49.125
15	4	48.8, 49.8, 50.0, 50.5	49.775
16	5	49.3, 50.2, 49.8, 48.9, 48.7	49.38
17	4	49.3, 47.5, 49.4, 48.4	48.65
18	4	47.8, 47.7, 48.8, 49.9	48.55
19	3	50.0, 49.5, 49.3	49.6
20	4	48.5, 49.2, 48.3, 47.8	48.45
21	4	47.9, 49.6, 49.8, 49.0	49.075

However, the given data show strong batch effects and the above conclusion is not warranted as we will see when adjusting by the "effective" sample size. For the data above we obtain

$$SS_b = 78.921\,, \quad SS_e = 29.148\,, \quad f = 17.123\,, \quad \hat{\sigma}_e^2 = .6939\,, \quad \hat{\sigma}_b^2 = 1.093$$

and thus $\hat{\rho} = .6116$ and $N^\star = 25.056$. Again interpolating from Table 3, this time at the row corresponding to $N^\star = 25.056$, we obtain

$$C_{\text{Table}\,\star}(25.056) = 1.255 \quad \text{and thus} \quad C_{\text{Adj}\,\star}(63) = 1.0122 \cdot 1.255 = 1.27$$

as the appropriately adjusted critical value. Since $1.17 < 1.27$ we cannot conclude with 90% confidence that the true C_{pk} is at least 1.

24.7 Concluding Remarks

The *effective sample size* device, in the two situations of tolerance and C_{pk} bounds examined here, provides a simple way of "approximating" the complex batch effect scenario by the better understood i.i.d. case. For C_{pk} confidence bounds the solution provided is apparently the first such treatment. Quality assurance practitioners, who tend to deal more often than not with such complications in data and who understand the statistical analysis solutions for the idealistic i.i.d. case, will welcome the simple modifications required by the approach presented here. From the simulations it appears that any error in coverage performance is mildly conservative. Whether this error can be reduced while maintaining the simplicity is not clear. It is more tempting to systematically examine many other data scenarios, well understood in the i.i.d. case but difficult in the context of random or batch effects, in the light of similar modifications using the *effective sample size* device. Our scheme of finding the effective sample size was based on matching the variances of sample means under the given data scenario and under the i.i.d. approximation. Other criteria for matching, such as information indices, may be examined. Much more broadly but also vaguely at this point, one could contemplate how much inferences of any type under some i.i.d. approximation may differ from the corresponding inferences under the given data situation. One simple but deficient approximation is to randomly select one representative from each batch. What could be gained by resampling this process?

24.8 Appendix

Here we present the rationale for the approximation $g \approx f$. Let $w_i = n_i/N$ and observe $\sum_{i=1}^{B} w_i = 1$. Further let

$$A = \sum_{i=1}^{B} w_i^2 \quad \text{and} \quad U = \sum_{i=1}^{B} w_i \left(w_i - A\right)^2 = \sum_{i=1}^{B} w_i^3 - A^2 \, .$$

Then

$$\frac{(1 - \sum w_i^2)^2}{\sum w_i^2 - 2\sum w_i^3 + (\sum w_i^2)^2} = \frac{1-A}{A} \, \frac{1-A}{1-A-2U/A} \approx \frac{1-A}{A} \, ,$$

where in the approximation step we assume that

$$\frac{U}{A} = \sum_{i=1}^{B} w_i \left(\frac{w_i}{A} - 1\right)^2 A \ll 1 \, , \quad \text{since} \quad w_i \approx \frac{1}{B} \, , \quad A \approx \frac{1}{B} \, , \quad \frac{w_i}{A} \approx 1 \, .$$

Note that $U/A = 0$, when the n_i are all the same. In that case the above approximation is exact.

24.9 Bibliography

Chou, Y.M., Owen, D.B., and Borego, S.A. (1990), "Lower confidence limits on process capability indices," *J. Qual. Technol.* **22**, 223-229.

Fisher, L. and Van Belle, G. (1993), *Biostatistics*, Wiley, New York.

Jennett, W.J. and Welch, B.L. (1939), "The control of proportion defective as judged by a single quality characteristic varying on a continuous scale," *J. Roy. Statist. Soc.-Suppl.* **6**, 80-88.

Johnson, N.L. and Welch, B.L. (1940), "Applications of the noncentral t-distribution," *Biometrika* **31**, 362-389.

Kane, V.E. (1986), "Process capability indices," *J. Qual. Technol.* **18**, 41-52

Kish, L. (1965), *Survey Sampling*, Wiley, New York.

Kotz, S. and Johnson, N.L. (1993), *Process Capability Indices*, Chapman & Hall, London.

Kushler, R. and Hurley, P. (1992), "Confidence bounds for capability indices," . *Qual. Technol.* **24**, 188-195.

Mee, R.W. and Owen, D.B. (1983), "Improved factors for one-sided tolerance limits for balanced one-way ANOVA random model," *J. Amer. Statist. Assoc.* **78**, 901-905.

Odeh, R.E. and Owen, D.B. (1980), *Tables for Normal Tolerance Limits, Sampling Plans, and Screening*, Marcel Dekker, Inc., New York.

Owen, D.B. (1968), "A survey of properties and applications of the noncentral t-distribution," *Technometrics* **10**, 445-478.

Owen, D.B. (1985), "Noncentral t-distribution," *Encyclopedia of Statistical Sciences, Vol. 6*, Wiley, New York.

Seeger, P. and Thorsson, U. (1972), "Two-sided tolerance limits with two-stage sampling from normal populations – Monte Carlo studies of the distribution of coverages", *Applied Statistics* **22**, 292-300.

Skinner, C.J., Holt, D., and Smith, T.M.F. (Ed) (1989), *Analysis of Complex Surveys*, Wiley, New York.

Vangel, M.G. (1995), "Design Allowables from Regression Models Using Data From Several Batches", *Composite Materials: Testing and Design (Twelfth Volume)*, ASTM STP 1274, R.B. Deo and C.R. Saff, Eds., American Socitey for Testing and Materials, Philadelphia, pp. 358-370.