Estimating the Standard Deviation in Quality-Control Applications

MAHMOUD A. MAHMOUD

Cairo University, Cairo, Egypt

G. ROBIN HENDERSON

Halcro Consultancy, 20 Mayburn Loan, Loanhead, EH20 9EN, UK

EUGENIO K. EPPRECHT

PUC-Rio, R. Marquês de São Vicente 225, 22451-900, Rio de Janeiro, Brazil

WILLIAM H. WOODALL

Virginia Tech, Blacksburg, VA 24061-0439, USA

In estimating the standard deviation of a normally distributed random variable, a multiple of the sample range is often used instead of the sample standard deviation in view of the range's computational simplicity. Although it is well known that use of the sample standard deviation is more efficient if the sample size exceeds 2, many statistical quality-control textbooks argue that the loss in efficiency when using the sample range to estimate the process standard deviation is very small with relatively small sample sizes. In this paper, we show that this loss in efficiency can be relatively large even for very small sample sizes and thus strongly advise against using range-based methods. We found that some previously published tables of relative efficiencies were either mislabeled or inaccurate. We also make some recommendations when a number of samples have been taken over time.

Key Words: Control Charts; Mean-Squared Error; Process Capability; Relative Efficiency; Statistical Process Control.

Introduction

USE OF SOME statistical quality-control tools, such as the Shewhart variable control charts and process-capability indices, require estimation of both the process mean and the process standard deviation, often from m independent samples each of size n. In their review paper, Jensen et al. (2006) reported that research has shown that estimation error in Phase I of control charting can lead to degraded performance in Phase II. Thus, it is important to use efficient estimators of the process parameters.

Use of the sample ranges is often recommended for estimating the standard deviation σ of a normally distributed quality characteristic when the sample size is very small; see, for example, Montgomery (2009) and Ott (1975). Grant and Leavenworth (1996, p. 125) stated, "In practical control-chart work in industry, R rather than s should nearly always be used as a measure of subgroup dispersion. R is easier to explain; almost everyone can understand range, whereas people with little background in statistics have difficulty understanding standard

Dr. Mahmoud is an Associate Professor in the Department of Statistics, Faculty of Economics and Political Science. His e-mail address is mamahmou@vt.edu.

Mr. Henderson is a Statistical Consultant. His e-mail address is halcro.robin@googlemail.com.

Dr. Epprecht is an Associate Professor in the Department of Industrial Engineering. His e-mail address is eke@puc-rio.br.

Dr. Woodall is a Professor in the Department of Statistics. He is a Fellow of ASQ. His e-mail address is bwoodall@vt.edu.

deviation." Woodall and Montgomery (2000) noted that the use of the sample standard deviation is more efficient if $n > 2$. They concluded, however, that the loss in efficiency when using the range method instead of the standard deviation to estimate σ is very small with the relatively small sample sizes typically used in variables control charts and other statistical process control (SPC) applications.

We first consider the case of a single sample (i.e., $m = 1$) and show that the use of the sample range can be more inefficient than has been recognized. The sample range for random variables $X_1, X_2, \ldots, X_n$ ($n > 1$) is defined as

$$R = \underset{i}{\mathrm{Max}}(X_i) - \underset{i}{\mathrm{Min}}(X_i). \tag{1}$$

The use of the sample range in estimating the Shewhart $\overline{X}$-chart limits and process-capability indices requires values of the mean and standard deviation of R/σ as a function of the sample size n. We focus on the case of independent, identically distributed (i.i.d.) normal random variables. It is well known that both the mean and standard deviation of R, μ_R and σ_R, respectively, are multiples of σ in the form

$$\mu_R = d_2\sigma \quad \text{and} \quad \sigma_R = d_3\sigma, \tag{2}$$

where d_2 and d_3 are functions of the sample size n, known as "control-chart constants" in the SPC literature. The control chart constants d_2 and d_3 corresponding to different values of the sample size n have been tabulated for a normal population in many statistical quality-control textbooks; see, for example, Montgomery (2009, p.702). Also, see our Table 2.

As mentioned above, the sample standard deviation S, where

$$S = \sqrt{\frac{\sum(X_i - \overline{X})^2}{n-1}}$$

and $\overline{X}$ is the sample mean, is a more efficient estimator for σ than the usual estimator based on the sample range for samples of size $n > 2$. Using the fact that the quantity $(n-1)S^2/\sigma^2$ follows a chi-squared distribution with $n-1$ degrees of freedom, it can be shown that the mean and standard deviation of S, μ_S and σ_S, respectively, are also multiples of σ in the form

$$\mu_S = c_4\sigma \quad \text{and} \quad \sigma_S = \sqrt{1-c_4^2}\,\sigma, \tag{3}$$

where

$$c_4 = \frac{\sqrt{\dfrac{2}{n-1}} \times \Gamma[n/2]}{\Gamma[(n-1)/2]}. \tag{4}$$

Montgomery (2009, p. 702) stated that the constant c_4 can be closely approximated by

$$c_4 \approx \frac{4n-4}{4n-3}, \tag{5}$$

for $n > 25$. This approximation leads to an approximate variance for S of

$$\sigma_S^2 \approx \frac{(8n-7)\sigma^2}{(4n-3)^2}. \tag{6}$$

When one has only a single sample, the commonly used estimators for the process standard deviation are

$$\hat{\sigma}_1 = \frac{R}{d_2}, \quad \hat{\sigma}_2 = \frac{S}{c_4}, \quad \text{and} \quad \hat{\sigma}_3 = S. \tag{7}$$

Note that, when $n = 2$, $\hat{\sigma}_1 = \hat{\sigma}_2 = R/1.128$ while $\hat{\sigma}_3 = R/1.414$. The standard deviations of the unbiased estimators $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are

$$\frac{d_3}{d_2}\sigma \quad \text{and} \quad \frac{\sqrt{1-c_4^2}}{c_4}\sigma, \tag{8}$$

respectively, while the standard deviation of the biased estimator $\hat{\sigma}_3$ is

$$\sqrt{1-c_4^2}\,\sigma. \tag{9}$$

The measure of relative efficiency commonly used to compare one estimator with another is the reciprocal of the ratio of the respective mean-squared errors. The mean-squared error (MSE) of an estimator $\hat{\sigma}$ is defined as

$$\mathrm{MSE}(\hat{\sigma}) = \mathrm{E}(\hat{\sigma} - \sigma)^2 = \mathrm{Var}(\hat{\sigma}) + [B(\hat{\sigma})]^2, \tag{10}$$

where $\mathrm{E}(\cdot)$ and $\mathrm{Var}(\cdot)$ denote "expectation" and "variance," respectively, and $B(\hat{\sigma}) = \mathrm{E}(\hat{\sigma}) - \sigma$ is the bias of $\hat{\sigma}$. Vardeman (1999) used the square root of the mean-squared errors to compare the efficiencies of some proposed estimators of the population standard deviation.

The tables in Montgomery (2009, p. 112) and in the NIST/SEMATECH Handbook (Section 6.3.2.1) purport to contain the relative efficiencies of $\hat{\sigma}_1$ to $\hat{\sigma}_3$, but actually contain the relative efficiencies of $\hat{\sigma}_1$ to $\hat{\sigma}_2$. Their values evidently come from Ott (1975, p. 200), who did not clearly specify the estimator being compared with $\hat{\sigma}_1$, although this point was clarified in Ott et al. (2000). Table 1 shows the relative efficiency values given in these three sources. The results in Table 1 have been used to justify using the estimator $\hat{\sigma}_1$ to estimate the standard deviation for small sample sizes instead of $\hat{\sigma}_3$. Montgomery (2009, p. 96) stated that the loss of efficiency resulted from

TABLE 1. The Relative Efficiency Values of $\hat{\sigma}_1$ Compared with $\hat{\sigma}_2$

n	Relative efficiency
2	1.000
3	0.992
4	0.975
5	0.955
6	0.930
10	0.850

using $\hat{\sigma}_1$ is negligible when the sample size $n \leq 6$ and that this estimator "works very well and is entirely satisfactory."

The issue of which estimator of the process standard deviation to use based on relative efficiency comparisons goes back to Shewhart (1931) and Davies and Pearson (1934). Shewhart (1931, p. 287) compared the relative efficiencies of S/c_4 to R/d_2, i.e., of $\hat{\sigma}_2$ with $\hat{\sigma}_1$, in a figure and remarked that "the very rapid decrease in the efficiency of the estimate derived from the range is striking." Davies and Pearson (1934) compared the relative efficiencies of these two estimators in a more complete version of Table 1. These authors stated that the efficiency of the range method falls off rapidly for values of $n > 5$, and that it should not be used for $n > 10$. One should note, however, that this sample-size recommendation only applies if the standard deviation estimator $\hat{\sigma}_2$ is being considered versus the range method.

In our paper, we show that Table 1 cannot be used to justify the use of the range method to estimate σ for small sample sizes. We also give accurate values for the relative efficiencies of $\hat{\sigma}_1$ to estimators $\hat{\sigma}_2$ and $\hat{\sigma}_3$, along with relevant formulas used to calculate them. In addition, we consider the relative efficiencies of some estimators in the multiple sample case. For the convenience of the reader, a table defining the 12 estimators we consider, along with some of their properties, is given in the Appendix.

Relative Efficiency of Estimators of σ

Single-Sample Case

The relative efficiency of $\hat{\sigma}_1$ to $\hat{\sigma}_2$ is

$$\mathrm{RE}(\hat{\sigma}_1/\hat{\sigma}_2) = \frac{\mathrm{MSE}(\hat{\sigma}_2)}{\mathrm{MSE}(\hat{\sigma}_1)} = \frac{d_2^2(1-c_4^2)}{d_3^2 c_4^2}. \quad (11)$$

The relative efficiency of $\hat{\sigma}_1$ to $\hat{\sigma}_3$ is

$$\mathrm{RE}(\hat{\sigma}_1/\hat{\sigma}_3) = \frac{2(1-c_4)d_2^2}{d_3^2}, \quad (12)$$

while the relative efficiency of $\hat{\sigma}_2$ to $\hat{\sigma}_3$ is

$$\mathrm{RE}(\hat{\sigma}_2/\hat{\sigma}_3) = \frac{2c_4^2}{1+c_4}. \quad (13)$$

The fifth and sixth columns of Table 2 show the values of $\mathrm{RE}(\hat{\sigma}_1/\hat{\sigma}_2)$ and $\mathrm{RE}(\hat{\sigma}_1/\hat{\sigma}_3)$, respectively. One can observe that the values of $\mathrm{RE}(\hat{\sigma}_1/\hat{\sigma}_2)$ given in column 5 of Table 2 match the relative efficiency values shown in Table 1, although the value for $n = 6$ doesn't quite match the value of 0.930 given by Montgomery (2009, p. 112) and in the NIST/SEMATECH Handbook. Ott (1975, p. 200) only gave two significant digits for $n = 6$, i.e., a value of 0.93.

Kenett and Zacks (1998, p. 352) argued that "despite the wide use of the sample ranges to estimate the process standard deviation, the method is neither very efficient nor robust. It is popular only because the sample range is easier to compute than the sample standard deviation. However, because many hand calculators now have built-in programs for computing the sample standard deviation, the computational advantage of the range should not be considered." Our results support Kenett and Zacks' (1998) argument. It can be observed from the values of $\mathrm{RE}(\hat{\sigma}_1/\hat{\sigma}_3)$ in column six of Table 2 that one is better off using S to estimate σ. For example, for $n = 2$, the estimator $\hat{\sigma}_1$ is only 71% efficient compared with $\hat{\sigma}_3$. Thus, Table 1 cannot be used to justify the use of $\hat{\sigma}_1$ to estimate the process standard deviation for small sample sizes. (We have included some of the relevant control-chart constants in Table 2 for the convenience of the reader.)

Bissell (1990, p. 339) gave the relative efficiency values $\mathrm{RE}(\hat{\sigma}_1/\hat{\sigma}_3)$ and $\mathrm{RE}(\hat{\sigma}_2/\hat{\sigma}_3)$. To do this, he used the following approximation for the variance of $\hat{\sigma}_3$, which was given in Hald (1967, p.300):

$$\sigma_S^2 \approx \frac{\sigma^2}{2(n-1)}. \quad (14)$$

The approximation in Equation (14) does not match the one given by Montgomery (2009) presented in Equation (6). For even moderately large values of n, however, the square roots of both of these expressions give a very close approximation to the exact value given in Equation (9). To calculate his two sets of relative efficiencies, Bissell (1990) used the approximation in Equation (14) as the denominator of the

TABLE 2. Control-Chart Constants and Relative Efficiencies

n	c_4	d_2	d_3	$\mathrm{RE}(\hat\sigma_1/\hat\sigma_2)$	$\mathrm{RE}(\hat\sigma_1/\hat\sigma_3)$	$\mathrm{RE}(\hat\sigma_2/\hat\sigma_3)$	$\mathrm{RE}(\hat\sigma_3/\hat\sigma_4)$	$\mathrm{RE}(\hat\sigma_3/\hat\sigma_5)$	$\mathrm{RE}(\hat\sigma_6/\hat\sigma_5)$
2	0.79788	1.1284	0.85250	1.000	0.708	0.708	0.899	0.899	0.978
3	0.88623	1.6926	0.88837	0.992	0.826	0.833	0.949	0.943	0.978
4	0.92132	2.0588	0.87981	0.975	0.862	0.884	0.981	0.961	0.980
5	0.93999	2.3259	0.86408	0.955	0.870	0.911	1.010	0.970	0.982
6	0.95153	2.5344	0.84804	0.933	0.866	0.928	1.039	0.976	0.984
7	0.95937	2.7044	0.83321	0.911	0.856	0.939	1.067	0.980	0.986
8	0.96503	2.8472	0.81983	0.890	0.844	0.948	1.095	0.983	0.987
9	0.96931	2.9700	0.80783	0.869	0.830	0.954	1.122	0.985	0.989
10	0.97266	3.0775	0.79705	0.850	0.815	0.959	1.150	0.986	0.989
11	0.97535	3.1729	0.78731	0.831	0.801	0.963	1.177	0.988	0.990
12	0.97756	3.2585	0.77848	0.814	0.786	0.966	1.203	0.989	0.991
13	0.97941	3.3360	0.77042	0.797	0.772	0.969	1.229	0.990	0.992
14	0.98097	3.4068	0.76302	0.781	0.759	0.972	1.255	0.990	0.992
15	0.98232	3.4718	0.75621	0.766	0.745	0.974	1.281	0.991	0.993
16	0.98348	3.5320	0.74991	0.751	0.733	0.975	1.306	0.992	0.993
17	0.98451	3.5879	0.74405	0.738	0.721	0.977	1.331	0.992	0.993
18	0.98541	3.6401	0.73859	0.725	0.709	0.978	1.355	0.993	0.994
19	0.98621	3.6890	0.73348	0.712	0.697	0.979	1.379	0.993	0.994
20	0.98693	3.7350	0.72869	0.700	0.687	0.980	1.403	0.993	0.994
21	0.98758	3.7783	0.72417	0.689	0.676	0.981	1.427	0.994	0.995
22	0.98817	3.8194	0.71991	0.678	0.666	0.982	1.450	0.994	0.995
23	0.98870	3.8583	0.71589	0.667	0.656	0.983	1.473	0.994	0.995
24	0.98919	3.8953	0.71207	0.657	0.647	0.984	1.496	0.995	0.995
25	0.98964	3.9306	0.70844	0.648	0.638	0.984	1.519	0.995	0.995

relative efficiency ratio. There are two problems with Bissell's (1990) approach. First, the approximation in Equation (14) is only accurate for moderately large sample sizes. Second, Bissell (1990) should have considered the MSE for S, not the variance, because S is a biased estimator of σ. For $n = 2$, for example, he gave $\mathrm{RE}(\hat\sigma_1/\hat\sigma_3) = \mathrm{RE}(\hat\sigma_2/\hat\sigma_3) = 0.876$, instead of the exact value 0.708.

Vardeman (1999) studied the efficiency of alternative estimators for σ based on multiples of R and S. In particular, Vardeman (1999) studied the efficiency of the following two estimators:

$$\hat\sigma_4 = \frac{d_2}{d_2^2 + d_3^2} R \quad \text{and} \quad \hat\sigma_5 = c_4 S.$$

Vardeman (1999) showed that the biased estimators $\hat\sigma_4$ and $\hat\sigma_5$ are optimal among estimators of their respective forms in terms of the mean-squared error measure defined in Equation (10). The mean-squared errors of these estimators are

$$\mathrm{MSE}(\hat\sigma_4) = \frac{d_3^2}{d_2^2 + d_3^2}\sigma^2$$

and

$$\mathrm{MSE}(\hat\sigma_5) = (1 - c_4^2)\sigma^2, \tag{15}$$

respectively. Based on these MSE values, the relative efficiency of $\hat\sigma_3$ to $\hat\sigma_4$ is

$$\mathrm{RE}(\hat\sigma_3/\hat\sigma_4) = \frac{d_3^2}{2(d_3^2 + d_2^2)(1 - c_4)}, \tag{16}$$

while the relative efficiency of $\hat\sigma_3$ to $\hat\sigma_5$ is

$$\mathrm{RE}(\hat\sigma_3/\hat\sigma_5) = \frac{1 + c_4}{2}. \tag{17}$$

The eighth and ninth columns of Table 2 give the relative efficiency values $\mathrm{RE}(\hat\sigma_3/\hat\sigma_4)$ and $\mathrm{RE}(\hat\sigma_3/\hat\sigma_5)$. Two important conclusions are obtained from the results in these two columns. First, the loss in efficiency if one uses the sample standard deviation S to estimate σ instead of the optimal estimators $\hat\sigma_4$ and $\hat\sigma_5$ is relatively small. For example, when $n = 2$, S is

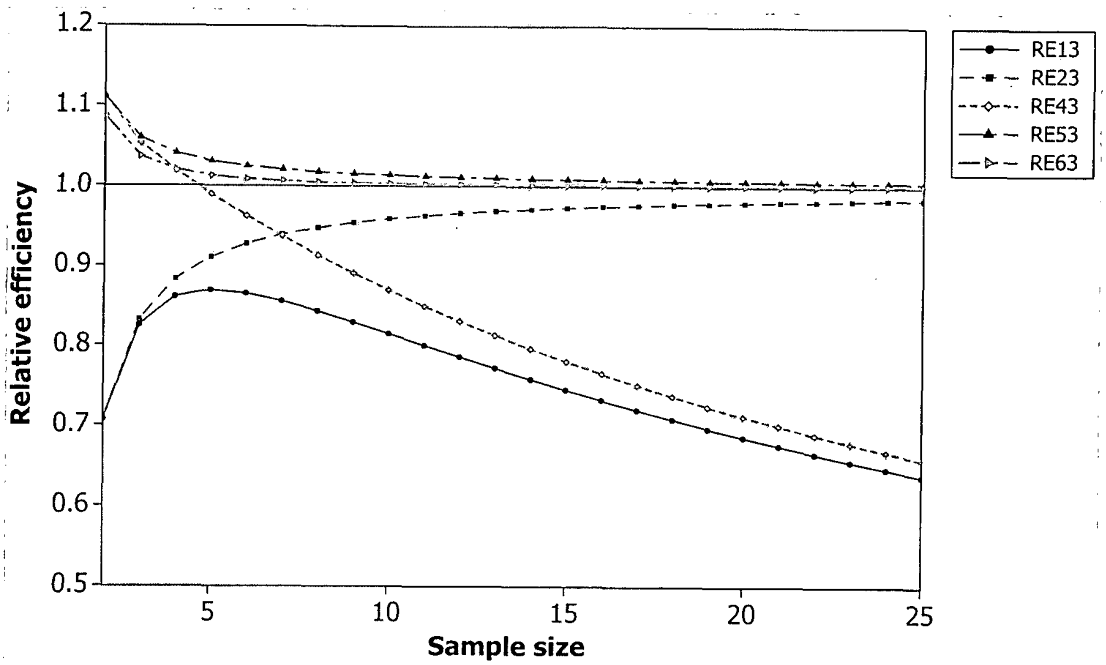

FIGURE 1. The Relative Efficiency Values of $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\sigma}_4$, $\hat{\sigma}_5$, and $\hat{\sigma}_6$ to $\hat{\sigma}_3$.

90% efficient compared with the optimal estimators $\hat{\sigma}_4$ and $\hat{\sigma}_5$. Second, for $n > 4$, S is more efficient than the optimal estimator among all estimators calculated as a multiple of the sample range R, i.e., $\hat{\sigma}_4$.

For the sake of completeness, we now consider the maximum likelihood estimator (MLE) of σ, which is $\hat{\sigma}_6 = (\sqrt{(n-1)/n})S$. The MSE of this estimator is

$$\text{MSE}(\hat{\sigma}_6) = \left(2 - 2\sqrt{\frac{n-1}{n}}c_4 - \frac{1}{n}\right)\sigma^2. \quad (18)$$

Comparison of the last two columns of Table 2 indicates that the efficiency of the MLE estimator $\hat{\sigma}_6$ lies between that of the optimal estimator $\hat{\sigma}_5$ and that of the estimator $\hat{\sigma}_3$.

Figure 1 shows the relative efficiency values of estimators $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\sigma}_4$, $\hat{\sigma}_5$, and $\hat{\sigma}_6$ to estimator $\hat{\sigma}_3$. As shown in this figure, only $\hat{\sigma}_5$ and $\hat{\sigma}_6$ are uniformly more efficient than $\hat{\sigma}_3$ for all values of the sample size n. Estimator $\hat{\sigma}_4$ is more efficient than estimator $\hat{\sigma}_3$ only for $n \leq 4$. Estimator $\hat{\sigma}_3$ is uniformly more efficient than the two commonly used estimators $\hat{\sigma}_1$ and $\hat{\sigma}_2$. An important conclusion from Figure 1 is the very rapid decrease in the relative efficiency (with respect to $\hat{\sigma}_3$) of the estimators derived from the sample range, i.e., $\hat{\sigma}_1$ and $\hat{\sigma}_4$.

Multiple-Sample Case

Where $m > 1$ independent random samples of a common size $n > 1$ lead to sample ranges R_1, R_2, ..., R_m or sample standard deviations S_1, S_2, ..., S_m, there is more than one way of combining these into a single estimator. We will only consider the case where the subgroup size is greater than one. For the $n = 1$ case (i.e., multiple individual observations), the reader is referred to Braun and Park (2008). For $n > 1$, standard SPC texts recommend the estimators $\overline{R}/d_2$ and $\overline{S}/c_4$, which are equivalent to averaging the m independent values of $\hat{\sigma}_1$ or $\hat{\sigma}_2$, respectively, corresponding to each sample.

Denoting the bias and the variance of a single sample estimator by B and V, respectively, the MSE of the multiple-sample estimator given by the average of m such estimators has MSE given by

$$\text{MSE}_{\text{averaged estimator}} = V/m + B^2, \quad (19)$$

which converges to B^2 as m increases.

In particular, when the individual sample estimators are unbiased, the MSE of the averaged estimator will decrease asymptotically to zero as m increases. This means that, although in the single-sample context the biased estimators $\hat{\sigma}_3$ and $\hat{\sigma}_5$ always have

TABLE 3. Relative Efficiency of $\hat{\sigma}_7$ with Respect to $\hat{\sigma}_8$

	n							
m	2	3	4	5	6	10	15	20
15	3.88	3.15	2.65	2.32	2.10	1.66	1.44	1.33
20	5.03	3.99	3.29	2.84	2.53	1.92	1.61	1.45
25	6.19	4.83	3.93	3.35	2.96	2.17	1.78	1.58
30	7.35	5.67	4.57	3.87	3.39	2.43	1.95	1.71
50	11.97	9.04	7.14	5.92	5.10	3.45	2.62	2.21
100	23.54	17.47	13.55	11.07	9.38	6.00	4.31	3.48
300	69.81	51.18	39.22	31.64	26.51	16.20	11.08	8.53

smaller MSE values than the unbiased estimator $\hat{\sigma}_2$, for sufficiently large values of m, the situation is reversed and the average of m independent values of $\hat{\sigma}_2$ (i.e., the usual unbiased estimator $\overline{S}/c_4$) will have smaller MSE than the average of m independent values of either $\hat{\sigma}_3$ or $\hat{\sigma}_5$. It is important therefore to compare the relative efficiencies of the averages of m values of the estimators $\hat{\sigma}_2$ (unbiased) and $\hat{\sigma}_5$ (minimum MSE estimator in the case of a single sample) for different values of m and n. We let $\hat{\sigma}_7$ and $\hat{\sigma}_8$ represent, respectively, these estimators; that is, $\hat{\sigma}_7 = \overline{S}/c_4$ and $\hat{\sigma}_8 = c_4\overline{S}$.

From Equations (8), (15), and (19), given m samples of size n,

$$\text{MSE}(\hat{\sigma}_7) = \frac{1-c_4^2}{c_4^2}\frac{\sigma^2}{m} \tag{20}$$

and

$$\text{MSE}(\hat{\sigma}_8) = c_4^2(1-c_4^2)\frac{\sigma^2}{m} + (1-c_4^2)^2\sigma^2. \tag{21}$$

This gives

$$\text{RE}(\hat{\sigma}_7/\hat{\sigma}_8) = c_4^2[m(1-c_4^2)+c_4^2], \tag{22}$$

which increases monotonically with m.

Table 3 shows values of $\text{RE}(\hat{\sigma}_7/\hat{\sigma}_8)$ for several values of m and n. The advantage of $\hat{\sigma}_7$ over $\hat{\sigma}_8$ is remarkable.

Of course we could also calculate $\text{MSE}(\overline{S})$, but because $\overline{S}$ is the geometric mean of $\hat{\sigma}_7 = \overline{S}/c_4$ and $\hat{\sigma}_8 = c_4\overline{S}$, its MSE will lie somewhere between their MSEs. Therefore, there is no need to consider this estimator in detail to conclude that it is not to be recommended either, in face of the greater efficiency of $\hat{\sigma}_7$. Neither do we need to consider the average of m independent values of $\hat{\sigma}_6$, because the bias of $\hat{\sigma}_6$ is larger than the bias of S.

The results in Table 3 illustrate our statement that, in the multiple-sample case, the average of unbiased estimators can be much more efficient than the average of biased estimators (even when the latter are optimal in the single-sample case). This leads to the conclusion that averaging m values of the optimal single-sample range-based estimator $\hat{\sigma}_4$ is not advisable. For completeness, however, we examine the classic unbiased multiple-sample range-based estimator $\hat{\sigma}_9 = \overline{R}/d_2$ (which is the average of m values of $\hat{\sigma}_1$) and, analogous to what was done in the single-sample case regarding $\hat{\sigma}_1$ and $\hat{\sigma}_2$, compare its efficiency with that of $\hat{\sigma}_7$. Because $\hat{\sigma}_9$ and $\hat{\sigma}_7$ are averages of m values of $\hat{\sigma}_1$ and $\hat{\sigma}_2$, respectively, and both $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are unbiased, it becomes evident from Equation (19) that $\text{RE}(\hat{\sigma}_9/\hat{\sigma}_7) = \text{RE}(\hat{\sigma}_1/\hat{\sigma}_2)$ (the latter is given in the fifth column of Table 2). The estimator $\hat{\sigma}_1$ is not to be recommended in the single-sample case and the corresponding estimator in the multiple-sample case, $\hat{\sigma}_9$, is not to be recommended either.

Another method for combining the individual standard deviation estimators from m samples into a single estimator is using the pooled sample standard deviation, defined as

$$S_{\text{pooled}} = \sqrt{\frac{\sum_{i=1}^{m}(n_i-1)S_i^2}{\sum_{i=1}^{m}(n_i-1)}}, \tag{23}$$

where S_i^2 is the ith sample variance and n_i is the ith sample size. (Note that, when $m = 1$, S_{pooled} reduces to $\hat{\sigma}_3$, the sample standard deviation.) Vardeman (1999) noted that "even a poor multiple of S_{pooled} is better than the best linear combination of sample ranges" in terms of the mean-squared error measure

TABLE 4. Values of $\mathrm{RE}(\hat{\sigma}_7/\hat{\sigma}_{10})$, $\mathrm{RE}(\hat{\sigma}_{11}/\hat{\sigma}_{10})$, and $\mathrm{RE}(\hat{\sigma}_{12}/\hat{\sigma}_{10})$

n	m	$\nu+1$	$c_4(n)$	$c_4(\nu+1)$	$\mathrm{RE}(\hat{\sigma}_7/\hat{\sigma}_{10})$	$\mathrm{RE}(\hat{\sigma}_{11}/\hat{\sigma}_{10})$	$\mathrm{RE}(\hat{\sigma}_{12}/\hat{\sigma}_{10})$
2	20	21	0.7979	0.9876	0.865	0.994	0.975
2	25	26	0.7979	0.9901	0.867	0.995	0.980
2	30	31	0.7979	0.9917	0.869	0.996	0.983
2	50	51	0.7979	0.9950	0.872	0.998	0.990
3	20	41	0.8862	0.9938	0.909	0.997	0.988
3	25	51	0.8862	0.9950	0.910	0.998	0.990
3	30	61	0.8862	0.9958	0.911	0.998	0.992
3	50	101	0.8862	0.9975	0.913	0.999	0.995
4	20	61	0.9213	0.9958	0.932	0.998	0.992
4	25	76	0.9213	0.9967	0.933	0.998	0.993
4	30	91	0.9213	0.9972	0.933	0.999	0.994
4	50	151	0.9213	0.9983	0.934	0.999	0.997
5	20	81	0.9400	0.9969	0.946	0.998	0.994
5	25	101	0.9400	0.9975	0.946	0.999	0.995
5	30	121	0.9400	0.9979	0.947	0.999	0.996
5	50	201	0.9400	0.9988	0.947	0.999	0.998

defined in Equation (10). Vardeman (1999) showed that the optimal standard deviation-based estimator of σ in terms of MSE measure is given by

$$c_4 S_{\text{pooled}}, \tag{24}$$

where the function c_4 has argument $\nu + 1$ and $\nu = \sum_{i=1}^{m}(n_i - 1)$. The optimal estimator given in Equation (24) equals $\hat{\sigma}_5$ when $m = 1$. We denote this estimator by $\hat{\sigma}_{10}$ and compare its efficiency with the efficiency of $\hat{\sigma}_7$. To avoid ambiguity, we will henceforth use the notation $c_4(\cdot)$ to include the argument of the function. Vardeman (1999) showed that

$$\mathrm{MSE}(\hat{\sigma}_{10}) = [1 - c_4^2(\nu+1)]\sigma^2. \tag{25}$$

When all samples are of the same size n, $\nu + 1 = mn - m + 1$ for the determination of c_4. Given Equations (20) and (25), the relative efficiency of $\hat{\sigma}_{10}$ with respect to $\hat{\sigma}_7$ is

$$\mathrm{RE}(\hat{\sigma}_{10}/\hat{\sigma}_7) = \frac{[1 - c_4^2(n)]}{mc_4^2(n)[1 - c_4^2(\nu+1)]}, \tag{26}$$

which is greater than 1 for any values of m and n. This means that $\hat{\sigma}_{10}$ is uniformly better than $\hat{\sigma}_7$ in terms of MSE.

Table 4 contains $\mathrm{RE}(\hat{\sigma}_7/\hat{\sigma}_{10})$ values for a number of values of n and m, along with values of $c_4(n)$ and $c_4(\nu+1)$ for convenience of the reader. These results show that $\hat{\sigma}_{10}$ is preferable to $\hat{\sigma}_7$, although not commonly used in SPC practice. The smaller the sample size n, the greater the advantage of $\hat{\sigma}_{10}$ over $\hat{\sigma}_7$. Note that $\hat{\sigma}_{10}$ is biased, but the statement made before that, in the multiple-sample case, the combination of unbiased estimators is generally more efficient than the combination of biased estimators applies only to the case where the estimators are combined through averaging them, where Equation (19) holds.

Finally, it should be noted that, even for moderate values of m and n, $c_4(\nu + 1)$ is quite close to 1, so the estimators $\hat{\sigma}_{10}$ (biased, but optimal), $\hat{\sigma}_{11} = S_{\text{pooled}}$ (neither unbiased nor optimal, but most common), and $\hat{\sigma}_{12} = S_{\text{pooled}}/c_4(\nu+1)$ (unbiased) have nearly the same MSE. To show this, Table 4 also contains the relative efficiencies $\mathrm{RE}(\hat{\sigma}_{11}/\hat{\sigma}_{10})$ and $\mathrm{RE}(\hat{\sigma}_{12}/\hat{\sigma}_{10})$. For the computation of these values, the MSEs of $\hat{\sigma}_{11}$ and $\hat{\sigma}_{12}$ are required and are given by

$$\mathrm{MSE}(\hat{\sigma}_{11}) = 2[1 - c_4(\nu+1)]\sigma^2 \tag{27}$$

and

$$\mathrm{MSE}(\hat{\sigma}_{12}) = \frac{1 - c_4^2(\nu+1)}{c_4^2(\nu+1)}\sigma^2, \tag{28}$$

respectively.

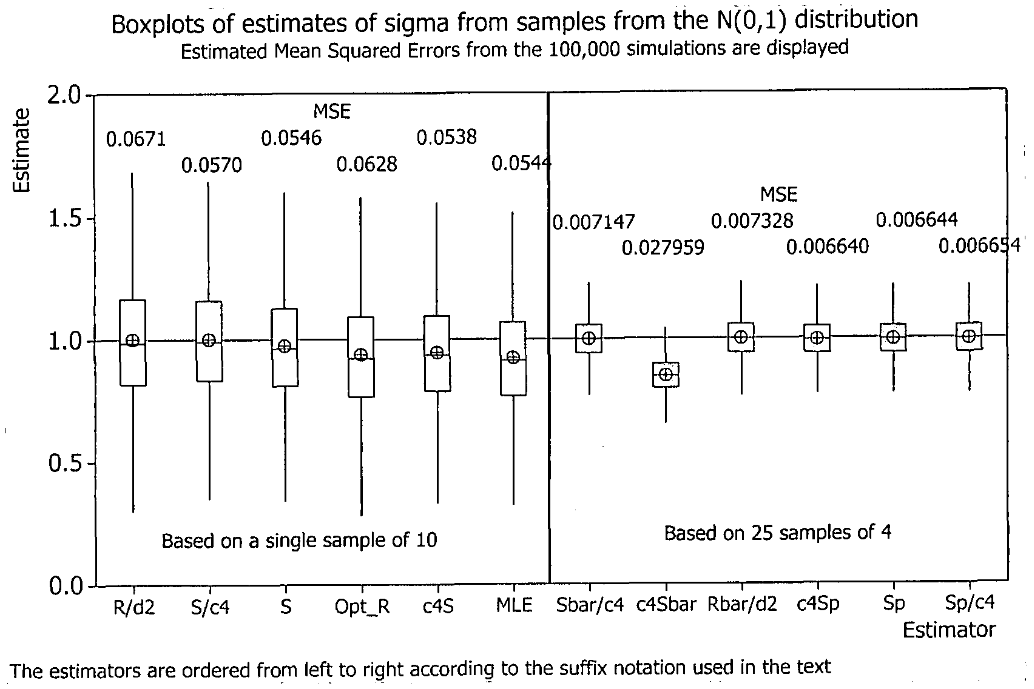

FIGURE 2. Boxplots of the Estimators Considered.

In the one-sample context, $\hat{\sigma}_2 = S/c_4$, $\hat{\sigma}_3 = S$, and $\hat{\sigma}_5 = c_4 S$ have reasonably different MSEs, which clearly indicate $\hat{\sigma}_5 = c_4 S$ as the best of these estimators. In the multiple-sample context, however, the differences between the MSEs of $\hat{\sigma}_{10} = c_4(\nu + 1)S_{\text{pooled}}$, $\hat{\sigma}_{11} = S_{\text{pooled}}$ and $\hat{\sigma}_{12} = S_{\text{pooled}}/c_4(\nu + 1)$ are much less important and, although $\hat{\sigma}_{10} = c_4(\nu + 1)S_{\text{pooled}}$ is uniformly better, it will not make much difference in practice if one prefers to use the unbiased estimator $\hat{\sigma}_{12} = S_{\text{pooled}}/c_4(\nu + 1)$ or the simpler and more familiar $\hat{\sigma}_{11} = S_{\text{pooled}}$. Clearly, however, any of these three estimators is to be preferred to $\hat{\sigma}_7 = \overline{S}/c_4(n)$ and Vardeman (1999) has previously shown the range-based methods to be very inefficient.

To provide an illustration, Figure 2 gives the boxplots of all the estimators considered, for the cases of single samples of size $n = 10$ and of $m = 25$ samples of size $n = 4$. The boxplots were each obtained through 100,000 simulations using Minitab® macros.

Concluding Remarks

In this paper, we have compared the relative efficiencies of different estimators of the standard deviation of a normally distributed random variable in two cases: a single sample of size n and m independent samples of size $n > 1$. (We do not consider the case $m > 1$ and $n = 1$ in our paper.) Our results show that the typically recommended use of the estimator R/d_2 to estimate σ instead of the sample standard deviation S when the sample size is small has not been always based on accurate relative efficiencies. The relative efficiency values given in several statistical quality-control sources, e.g. Montgomery (2009), Ott (1975), and the NIST/SEMATECH e-Handbook of Statistical Methods, compared the efficiency of R/d_2 with the unbiased estimator S/c_4, not to S.

For the single-sample case, we gave formulas and values that compare the efficiency of R/d_2 with S. Based on the mean-squared error criterion, the biased estimator S is definitely preferable to the commonly recommended estimator R/d_2. We also compared the relative efficiency of the sample standard deviation S and of the maximum likelihood estimator $(\sqrt{(n-1)/n})S$ (MLE) with the two optimal estimators of σ based on the minimum mean-squared error criterion studied by Vardeman (1999). Our results show that the loss in efficiency if one uses the sample standard deviation S or the MLE to estimate σ instead of the optimal estimators given in Vardeman (1999) is relatively small. However, one can use the optimal estimator $c_4 S$ more efficiently to estimate the process standard deviation.

In the multiple-sample case, where $m > 1$ individual estimates from samples of size $n > 1$ are available, as in Phase I of SPC, our results show that averaging the unbiased estimators leads to much more efficient multiple-sample estimators than averaging biased (albeit optimal in the single-sample case) estimators. The relative efficiency of $\overline{R}/d_2$ to $\overline{S}/c_4(n)$ equals the relative efficiency of R/d_2 to S/c_4, so use of $\overline{S}/c_4(n)$ is preferable to $\overline{R}/d_2$. However, the estimators based on S_{pooled} (namely, S_{pooled}, the minimum-MSE estimator $c_4 S_{\text{pooled}}$, and the unbiased estimator S_{pooled}/c_4, where c_4 has argument $mn-m+1$) are virtually equivalent for combinations of values of m and n typically encountered in practice, and any of these should be preferred to the commonly used estimator $\overline{S}/c_4(n)$. Our results support the conclusions of Vardeman (1999) in the multiple-sample case.

Our results are based on the assumption of normality. If the underlying distribution is more heavy-tailed, then use of the subgroup standard deviations might not be the best option. It is likely, however, that the use of the subgroup ranges would be even more adversely affected in this situation.

Overall, based on relative efficiency performance, we strongly recommend the use of the sample standard deviations in estimating the process standard deviation σ when calculating Shewhart $\overline{X}$-chart limits and when estimating process capability indices with subgrouped data (i.e., $n > 1$). We see no reason for the continued use of the range-based methods, especially because calculations are likely to be performed by software.

Acknowledgments

The authors would like to thank Dr. D. E. Roberts, Visiting Research Fellow at Edinburgh Napier University, who computed the values of the control-chart constants d_2 and d_3 to 9 decimal places using Mathcad. These values were used in the calculation of the relative efficiencies reported and displayed in this paper and are listed, rounded to five significant digits, in Table 2.

The research of Professor Woodall was partially supported by NSF grant CMMI-0927323. The research of Professor Epprecht was partially supported by CNPq, Brazil, PQ grant 302326/2008-1

Appendix
Table of Estimators Considered

Notation	Formula	Biased/unbiased	MSE
	Single sample		
$\hat{\sigma}_1$	R/d_2	U	$(d_3^2/d_2^2)\sigma^2$
$\hat{\sigma}_2$	S/c_4	U	$[(1-c_4^2)/c_4^2]\sigma^2$
$\hat{\sigma}_3$	S	B	$2(1-c_4)\sigma^2$
$\hat{\sigma}_4$	$[d_2/(d_2^2+d_3^2)]R$	B	$[d_3^2/(d_2^2+d_3^2)]\sigma^2$
$\hat{\sigma}_5$	$c_4 S$	B	$(1-c_4^2)\sigma^2$
$\hat{\sigma}_6$	$\sqrt{(n-1)/n}S$	B	$[2-2\sqrt{(n-1)/n}c_4 - 1/n]\sigma^2$
	Multiple samples		
Notation	Formula	Biased/unbiased	MSE
$\hat{\sigma}_7$	$\overline{S}/c_4(n)$	U	$[(1-c_4^2)/c_4^2](\sigma^2/m)$
$\hat{\sigma}_8$	$c_4(n)\overline{S}$	B	$c_4^2(1-c_4^2)(\sigma^2/m)+(1-c_4^2)^2\sigma^2$
$\hat{\sigma}_9$	$\overline{R}/d_2$	U	$(d_3^2/d_2^2)(\sigma^2/m)$
$\hat{\sigma}_{10}$	$c_4(\nu+1)S_{\text{pooled}}$	B	$[1-c_4^2(\nu+1)]\sigma^2$
$\hat{\sigma}_{11}$	S_{pooled}	B	$2[1-c_4(\nu+1)]\sigma^2$
$\hat{\sigma}_{12}$	$S_{\text{pooled}}/c_4(\nu+1)$	U	$[(1-c_4^2(\nu+1))/c_4^2(\nu+1)]\sigma^2$

Note: $\nu = \sum_{i=1}^{m}(n_i-1) = m(n-1)$ when all samples are of the same size.

References

BISSELL, A. F. (1990). "How Reliable Is Your Capability Index?" *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 39, pp. 331–340.

BRAUN, W. J. and PARK, D. (2008), "Estimation of σ for Individuals Charts". *Journal of Quality Technology* 40(3), pp. 332–344.

DAVIES, O. L. and PEARSON, E. S. (1934). "Methods of Estimating from Samples the Population Standard Deviation". Journal of the Royal Statistical Society 1, pp. 76–93.

GRANT, E. L. and LEAVENWORTH, R. S. (1996). *Statistical Quality Control*, 7th edition. New York, NY: McGraw-Hill Book Companies.

HALD, A. (1967). *Statistical Theory with Engineering Applications.* New York, NY: John Wiley & Sons.

JENSEN, W. A.; JONES-FARMER, L. A.; CHAMP, C. W.; and WOODALL, W. H. (2006). "Effects of Parameter Estimation on Control Chart Properties: A Literature Review". *Journal of Quality Technology* 38(4), pp. 349–364.

KENETT, R. S. and ZACKS, S. (1998). *Modern Industrial Statistics: Design and Control of Quality and Reliability.* Pacific Grove, CA: Brooks/Cole Publishing Company.

MONTGOMERY, D. C. (2009). *Introduction to Statistical Quality Control*, 6th edition. New York, NY: John Wiley & Sons.

NIST/SEMATECH *e-Handbook of Statistical Methods*, at http://www.itl.nist.gov/div898/handbook/index.htm, accessed 12/21/09.

OTT, E. R. (1975). *Process Quality Control: Troubleshooting and Interpretation of Data.* New York, NY: McGraw-Hill.

OTT, E. R.; SCHILLING, E.G.; and NEUBAUER, D. V. (2000). *Process Quality Control: Troubleshooting and Interpretation of Data*, 3rd edition. New York, NY: McGraw-Hill, New York.

SHEWHART, W. A. (1931). *Economic Control of Quality of Manufactured Product*, republished edition, 50th Anniversary Commemorative Reissue by American Society for Quality Control, Milwaukee, WI.

VARDEMAN, S. B. (1999). "A Brief Tutorial on the Estimation of the Process Standard Deviation". *IIE Transactions* 31, pp. 503–507.

WOODALL, W. H. and MONTGOMERY, D. C. (2000). "Using Ranges to Estimate Variability". *Quality Engineering* 13, pp. 211-217.