Probability in the Engineering and Informational Sciences, **13**, 1999, 237–250. Printed in the U.S.A.

ON THE ESTIMATION OF σ AND THE PROCESS CAPABILITY INDICES C_p AND C_{pm}

Ravindra Khattree
Department of Mathematics and Statistics
Oakland University
Rochester, Michigan 48309-4485

We consider the problem of estimation of the standard deviation σ and the process capability indices C_p and C_{pm} for a normally distributed process. The problem is addressed with the objective of obtaining optimal estimators under the classic criterion of minimum variance unbiased estimation, as well as under the Pitman measure of closeness.

1. INTRODUCTION

When setting up a control chart for any process, it is customary also to monitor the variability of the process using an *R*-chart or an *s*-chart. With today's computational ease and routine automation of such processes, *s*-charts are often preferred over *R*-charts, as the former uses the information available in the form of samples more effectively and to a much greater extent. If m is independent random samples of equal size n are assumed to have come from a normal process with mean μ and variance σ^2, then an *s*-chart plots the values of the sample standard deviation s_f for the future samples. Since $E(s_f) = c_{4,n}\sigma$, where $c_{4,n} = \sqrt{\frac{2}{n-1}}\,\frac{\Gamma[n/2]}{\Gamma[(n-1)/2]}$ (Montgomery [18]), the center line for the *s*-chart will be $c_{4,n}\sigma$ if σ was known. Since σ is usually unknown, it is often replaced by one of its unbiased estimates.

A related problem in quality control is the assessment of the capability of an industrial process. Given the upper and lower specification limits U and L, respectively, a measure of the process capability is defined as

$$C_p = \frac{U-L}{6\sigma} = \frac{U-L}{6}\,\frac{1}{\sigma},$$

 0269-9648/99 $12.50

https://doi.org/10.1017/S0269964899132108 Published online by Cambridge University Press

which is essentially a scaler multiple of σ^{-1}. When σ is unknown, it is often replaced in the above expression by either one of its unbiased estimators or the square root of that of σ^2, the process variance. However, in either scenario, the corresponding estimator of C_p does not preserve the property of unbiasedness, and, in fact, may not be optimal in any other meaningful sense. The objective of this article is to consider the problems of the estimation of σ and σ^{-1} (hence of C_p or C_{pm}—defined in Section 6) so as to provide certain optimal, thus superior, estimators.

The problems will be addressed in the classic framework as well as in the framework of the probabilistic criterion of Pitman measure of closeness. In the classic framework, although a weak optimality, the unbiasedness of the estimator is usually required to reduce the class of potential estimators to a smaller class of intuitive ones, out of which the one satisfying some other stronger optimality criterion, such as the minimum variance, is selected. An alternative approach may be to consider the estimators with minimum mean squared error (or more generally, minimum risk corresponding to a given loss function). However, finding such an estimator in the class of all possible estimators is usually impossible in most problems. This is often circumvented by first obtaining a suitable and intuitive unbiased estimator or a maximum likelihood estimator and shrinking it to obtain an appropriate minimum mean squared error (or minimum risk) estimator. It may, however, be noted that in certain cases this approach may fail in that the resulting expression for the estimator may involve the parameter itself.

The probabilistic criterion of the Pitman measure of closeness does not require unbiasedness. The criterion, however, is closely related to the concept of median unbiasedness (Ghosh and Sen [6] and Nayak [19]).

2. KKKDR AND ANGUS ESTIMATORS OF σ

The problem of estimating σ, as in the present framework, was recently considered by Kirmani, Kocherlakota, and Kocherlakota [17], Derman and Ross [5], and Angus [1]. Given m independent random samples $\{x_{i1}, \ldots, x_{in}\}$, of size n each, an unbiased estimator of σ can be given by $\hat{\sigma} = \bar{s}/c_{4,n}$, where $\bar{s} = m^{-1}\sum_{i=1}^{m} s_i$ is the mean of the m independently distributed sample standard deviations $s_i = [(n-1)^{-1} \times \sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2]^{1/2}$, $i = 1, \ldots, m$ corresponding to m independent samples.

Note that $\bar{s}$ above is obtained as the simple average of m independent sample standard deviations, which are square roots of corresponding variances. An alternative to this estimate, namely, $\hat{\sigma} = \bar{s}/c_{4,n}$, may be another estimate, say $\hat{\hat{\sigma}}$, obtained (after appropriate scaling adjustment) by interchanging the steps of averaging and square rooting [5,17], that is, by taking the square root of the average of m independent sample variances. This is meaningful because, in the process of computation as well as in a statistical sense, the variances are more basic quantities than the standard deviations. Also, the probability distribution of the weighted sum of the sample variances is more tractable due to the additivity property of independent chi-squared random variables. This property is especially convenient, as we shall see in Section 3, when our m independent samples were of different sizes, say $n_1, n_2, \ldots, n_m$, in

https://doi.org/10.1017/S0269964899132108 Published online by Cambridge University Press

which case the coefficients $c_{4,n_1}, c_{4,n_2}, \ldots, c_{4,n_m}$ will all be different, hence the simple averaging of $s_1, s_2, \ldots, s_m$ as done for the usual s-chart will not be significantly meaningful.

Specifically, as stated above, for the equal sample size case, an unbiased estimate of σ proposed in [5,17] is

$$\hat{\hat{\sigma}} = b_{4,n,m}^{-1} \sqrt{\frac{\sum_{i=1}^{m} s_i^2}{m}},$$

where

$$b_{4,n,m} = \sqrt{\frac{2}{m(n-1)}} \frac{\Gamma[(mn - m + 1)/2]}{\Gamma[(mn - m)/2]}.$$

In fact, it is evident from the above expression that $b_{4,n,m} = c_{4,mn-m+1}$. Thus, for the s-chart the center line will be set at $c_{4,n}\hat{\hat{\sigma}}$.

Why should one prefer $\hat{\hat{\sigma}}$ over $\hat{\sigma}$ as an estimate of process standard deviation in an s-chart? It is because, while $\hat{\sigma}$ and $\hat{\hat{\sigma}}$ are both unbiased for σ, $\hat{\hat{\sigma}}$ is the uniform minimum variance unbaised estimator (UMVUE) of σ [1], hence has the smaller variance among all unbiased estimates. It thus provides an improvement over $\hat{\sigma}$. Specifically,

$$\text{var}(\hat{\sigma}) = \frac{1 - c_{4,n}^2}{c_{4,n}^2} \frac{\sigma^2}{m}$$

and

$$\text{var}(\hat{\hat{\sigma}}) = \frac{1 - b_{4,n,m}^2}{b_{4,n,m}^2} \sigma^2.$$

The relative efficiency of $\hat{\hat{\sigma}}$ over $\hat{\sigma}$ may be judged by,

$$E = \frac{\text{var}(\hat{\sigma})}{\text{var}(\hat{\hat{\sigma}})} = \frac{1 - c_{4,n}^2}{mc_{4,n}^2} \frac{b_{4,n,m}^2}{1 - b_{4,n,m}^2}.$$

Numerical calculations (not presented here) show that, for various values of m and n, this relative efficiency is uniformly larger than 1 unless $m = 1$, in which case the two estimates are identical. What is more important is that when n is small, the efficiency values become considerably larger as m increases. In fact, in certain cases, a gain of over 13% can be realized when $\hat{\hat{\sigma}}$ instead of $\hat{\sigma}$ is used as the estimate of σ. The above situation of small n and moderate to large m is especially of interest when the sampling resources are limited. In such situations, it is often advisable to sample from the process at a number of time points, even though these samples may be small in size as opposed to sampling large samples at only a few time points.

Why should one be so particular about having an estimate of σ with smaller variance at the first place? The estimated value of σ not only determines the center line but also the upper and lower control limits of an s-chart as $B_{4,n}c_{4,n}\hat{\sigma}$ and $B_{3,n}c_{4,n}\hat{\sigma}$ (or $B_{4,n}c_{4,n}\hat{\hat{\sigma}}$ and $B_{3,n}c_{4,n}\hat{\hat{\sigma}}$, respectively), where the constants $B_{3,n}$ and $B_{4,n}$ have been tabulated in Appendix VI of [18]. In addition, this estimate is also used to determine the limits of $\overline{X}$-charts. Thus, a lot depends on the estimate of σ, and one will naturally prefer as precise an estimate as possible of σ. An estimator with higher variance will also have a higher probability of being observed as too large or too small, thereby resulting in a higher probability of the center line and control limits of the s-chart being mislocated. Additionally, the effect of this excessively high or low estimate of σ will also affect the control limits of the corresponding $\overline{X}$-chart as being too wide or too narrow. The estimator $\hat{\hat{\sigma}}$ provides a definite and uniform advantage over $\hat{\sigma}$ in both of these respects.

3. UNBALANCED DATA

If the m independent samples were of sizes $n_1, n_2, \ldots, n_m$, respectively, then both $\hat{\sigma}$ and $\hat{\hat{\sigma}}$ need to be redefined. Specifically, $\hat{\sigma}$ can be generalized in two ways, namely,

$$\hat{\sigma}^{(1)} = \frac{\sum_{i=1}^{m} s_i}{\sum_{i=1}^{m} c_{4,n_i}}$$

or

$$\hat{\sigma}^{(2)} = m^{-1} \sum_{i=1}^{m} s_i / c_{4,n_i},$$

and $\hat{\hat{\sigma}}$ will now be defined as

$$\hat{\hat{\sigma}} = c^{-1}_{4,\sum_{i=1}^{m} n_i - m + 1} \sqrt{\frac{\sum_{i=1}^{m} (n_i - 1) s_i^2}{\sum_{i=1}^{m} n_i - m}}.$$

It readily follows from the argument given in [1] that, even in this case, $\hat{\hat{\sigma}}$ is UMVUE hence has smaller variance than $\hat{\sigma}^{(1)}$ as well as $\hat{\sigma}^{(2)}$. However, no universal preference can be established between $\hat{\sigma}^{(1)}$ and $\hat{\sigma}^{(2)}$. Specifically, $\hat{\sigma}^{(1)}$ has smaller variance than $\hat{\sigma}^{(2)}$ if and only if

$$\left(\sum_{i=1}^{m} c_{4,n_i} \right)^{-2} \sum_{i=1}^{m} (1 - c^2_{4,n_i}) < m^{-2} \sum_{i=1}^{m} \frac{1 - c^2_{4,n_i}}{c^2_{4,n_i}}.$$

However, for most datasets, $\hat{\sigma}^{(1)}$ and $\hat{\sigma}^{(2)}$ are likely to be close to each other.

Returning back to the discussion of $\hat{\hat{\sigma}}$, in general, each coefficient c_{4,n_i} involves the computation of two gamma functions. In that sense, in addition to its superiority with respect to smaller variance, the computation of $\hat{\hat{\sigma}}$ can be done more efficiently because only one such coefficient, namely, $c_{4,\sum_{i=1}^m n_i - m+1}$, is involved. Further, if $\nu = c_{4,\sum_{i=1}^m n_i - m+1}$ is moderately large (say, greater than 25), then $c_{4,\nu}$ can be approximated by $c_{4,\nu} = 4(4\nu - 3)^{-1}(\nu - 1)$ [18].

4. ESTIMATION OF PROCESS CAPABILITY

As indicated earlier, to estimate the process capability index $C_p = \frac{U-L}{6}\frac{1}{\sigma}$, the common practice has been to either substitute for σ, the square root of the pooled sample variance, namely s_p or the unbiased estimate $\hat{\sigma}$ (or $\hat{\sigma}^{(1)}$ or $\hat{\sigma}^{(2)}$) defined earlier. One may also consider using $\hat{\hat{\sigma}}$ in the above expression. However, none of these resulting estimators of C_p are unbiased for C_p. Our interest in this section will be to provide certain unbiased estimator of C_p. One of these estimators will be shown to have the minimum possible variance among all unbiased estimators. Trivially, the problem is equivalent to that of the unbiased estimation of $\eta = \sigma^{-1}$, hence for the purpose of our discussion from now on our parameter of interest will be η rather than C_p.

For the unbalanced data, the following three estimators are unbiased for η,

$$\hat{\eta}_1 = \left\{\sum_{i=1}^m \sqrt{\frac{n_i - 1}{2}} \cdot \frac{\Gamma[(n_i-2)/2]}{\Gamma[(n_i-1)/2]}\right\}^{-1} \sum_{i=1}^m s_i^{-1} = \left(\sum_{i=1}^m g_i\right)^{-1} \sum_{i=1}^m s_i^{-1}$$

$$\hat{\eta}_2 = m^{-1} \sum_{i=1}^m \left\{\sqrt{\frac{n_i - 1}{2}} \cdot \frac{\Gamma[(n_i-2)/2]}{\Gamma[(n_i-1)/2]}\right\}^{-1} s_i^{-1} = m^{-1} \sum_{i=1}^m g_i^{-1} s_i^{-1}$$

$$\hat{\eta}_3 = \sqrt{2}\, \frac{\Gamma\left[\sum_{i=1}^m (n_i-1)/2\right]}{\Gamma\left[\sum_{i=1}^m (n_i-2)/2\right]} \cdot \left\{\sum_{i=1}^m (n_i - 1)s_i^2\right\}^{-1/2}$$

$$= \sqrt{2}\, \frac{\Gamma\left[\sum_{i=1}^m (n_i-1)/2\right]}{\Gamma\left[\sum_{i=1}^m (n_i-2)/2\right]} \cdot \frac{1}{\sqrt{\sum_{i=1}^n n_i - m}} \cdot s_p^{-1} = h^{-1} s_p^{-1},$$

where s_p^2 is the pooled variance given by $s_p^2 = \left(\sum_{i=1}^m n_i - m\right)^{-1} \sum_{i=1}^m (n_i - 1)s_i^2$. Since the unbiased estimator $\hat{\eta}_3$ given above depends on s_p^2, which is complete-sufficient, it follows from the Rao–Blackwell Theorem that $\hat{\eta}_3$ is also the uniformly minimum variance unbiased estimator of η.

When data are balanced, that is, when $n_i = n, i = 1, \ldots, m$, $\hat{\eta}_1$ and $\hat{\eta}_2$ are identical and, in this case,

$$\hat{\eta}_1 = \hat{\eta}_2 = \sqrt{\frac{2}{n-1}} \cdot \frac{\Gamma[(n-1)/2]}{\Gamma[(n-2)/2]} \cdot \bar{s}_H^{-1} = g^{-1} \bar{s}_H^{-1},$$

where $\bar{s}_H$ is the harmonic mean of $s_1, s_2 \ldots s_m$ and

$$g = \left[\sqrt{\frac{2}{n-1}} \cdot \frac{\Gamma[(n-1)/2]}{\Gamma[(n-2)/2]} \right]^{-1}.$$

Further, $\hat{\eta}_3$ simplifies to

$$\hat{\eta}_3 = \sqrt{\frac{2}{nm-m}} \cdot \frac{\Gamma[(nm-m)/2]}{\Gamma[(nm-m-1)/2]} \frac{1}{s_p} = h^{-1} s_p^{-1},$$

where h simplifies to

$$h = \left[\sqrt{\frac{2}{nm-m}} \cdot \frac{\Gamma[(nm-m)/2]}{\Gamma[(nm-m-1)/2]} \right]^{-1}.$$

The variance–covariance expressions for these estimators can be obtained from Chou and Owen [4] and Kirmani, Kocherlakota, and Kocherlakota [17].

5. ESTIMATION UNDER THE PITMAN MEASURE OF CLOSENESS

Given two estimators T_1 and T_2 of a real valued parameter θ, T_1 is said to be better than T_2 in the sense of the Pitman closeness under loss function $L(T, \theta)$ if

$$P_\theta[L(T_1, \theta) < L(T_2, \theta)] \geq \tfrac{1}{2} \quad \text{for all } \theta. \tag{5.1}$$

An informal interpretation of the above criterion can be given as: in a pair of two competing estimators, prefer the one which is closer to the parameter (in terms of loss function $L(\cdot, \cdot)$) more than 50% of the time.

For a general review of this criterion, see Keating, Mason, and Sen [11]. Various estimation of variances and covariances related problems under this criterion have been dealt with by Rao [21,22], Keating and Gupta [10], Keating [9], Khattree [12,13], Ghosh and Sen [6], Nayak [19,20], Sen, Nayak, and Khattree [24], and Gupta and Khattree [7,8].

Rao [21,22] in his two seminal works pointed out that Pitman closeness is an intrinsic measure to compare estimators and illustrated in many different estimation problems that the shrinking of an unbiased estimator to obtain a minimum mean square error (MMSE) estimator does not necessarily yield an estimator, which is superior in the sense of Eq. (5.1), when the loss function $L(T, \theta) = |T - \theta|$ is the absolute error loss function. In our specific context, this implies that the shrinking of the UMVUE of σ or σ^{-1} to obtain an MMSE estimator may not necessarily result in an estimator which is better in the sense of Eq. (5.1). Intuitively, it is because the MMSE criterion places too much emphasis on large values which may be observed with very small probabilities.

Given pooled sample variance s_p^2 from m independent subsamples of sizes $n_1, \ldots, n_m$, respectively, it may thus be desirable to obtain estimators of σ and $\eta = \sigma^{-1}$ within the respective classes

$$\mathcal{C} = \{cs_p : c > 0\}$$

and

$$\mathcal{D} = \{d \cdot s_p^{-1} : d > 0\},$$

which are closer to σ and η than any other estimator within $\mathcal{C}$ and $\mathcal{D}$, respectively, in the sense of Eq. (5.1). Fortunately, this can be done for a fairly general class of loss functions as shown by Nayak [20]. Specifically, the loss function $L(T,\theta)$ is assumed to satisfy the following conditions:

i. $L(t,\theta) = 0$ whenever $t = \theta$.
ii. For any fixed t, $L(t,\theta)$ is strictly increasing in θ for all $\theta < t$ and strictly decreasing in θ for all $\theta > t$.
iii. For any fixed θ, $L(t,\theta)$ is strictly decreasing in t for all $t < \theta$ and strictly increasing in t for all $t > \theta$.

These conditions are considerably mild in that they are satisfied by many commonly used loss functions such as the absolute loss $L_0(T,\theta) = |T - \theta|$ and the entropy loss $L_1(T,\theta) = \ln(T/\theta) - T/\theta - 1(T/\theta > 0)$.

Some of the other loss functions that satisfy the above requirements are

$$L_2(T,\theta) = \left(\frac{T}{\theta} - 1\right)^2$$

$$L_3(T,\theta) = \frac{\theta}{T} - \ln\frac{\theta}{T} - 1$$

$$L_4(T,\theta) = \left(\frac{\theta}{T} - 1\right)^2$$

$$L_5(T,\theta) = e^{b|T-\theta|} - b|T-\theta| - 1, \qquad b > 0 \text{ is a known constant.}$$

See Khattree [12], Khattree and Gill [14], Sen, Nayak, and Khattree [24], Gupta and Khattree [7,8], and Sen and Khattree [23], where such loss functions have been used in the context of Pitman measure of closeness. The following theorem provides the closest estimators of σ and $\eta = \sigma^{-1}$ in the respective classes $\mathcal{C}$ and $\mathcal{D}$.

THEOREM 5.1: *Let the loss function $L(T,\theta)$ satisfy the requirements i–iii stated above. Further, assume that $L(t,\theta)$ is a function of $t/\theta > 0$. Then, the best esti-*

https://doi.org/10.1017/S0269964899132108 Published online by Cambridge University Press

mators in the sense of Eq. (5.1) *of* σ *and* $\eta = \sigma^{-1}$ *in the respective classes* $\mathcal{C}$ *and* $\mathcal{D}$ *are given by,*

$$\hat{\sigma}_{pc} = (Q/M)^{1/2}$$

and

$$\hat{\eta}_{pc} = (M/Q)^{1/2},$$

where $Q = \sum_{i=1}^{m} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 = \left(\sum_{i=1}^{m} n_i - m\right) s_p^2$, *and* M *is the median of a chi-squared variable with degrees of freedom* $\left(\sum_{i=1}^{m} n_i - m\right)$.

Proof of the above theorem is straightforward and essentially uses the same argument as in Nayak [20]. We thus skip the proof.

It may be pointed out that the m independent subsamples can be viewed as the data in a one-way classification when there is no treatment effect. That being the case, in order to estimate σ and $\eta = \sigma^{-1}$, the (more) appropriate estimators may be obtained as certain suitable scalar multiples of $\sqrt{Q^*}$ and $\sqrt{1/Q^*}$, respectively, where $Q^* = \sum_{i=1}^{m} \sum_{j=1}^{n_i} (x_{ij} - \bar{\bar{x}})^2$, where $\bar{\bar{x}} = \left(\sum_{i=1}^{m} n_i\right)^{-1} \sum_{i=1}^{m} \sum_{j=1}^{n_i} x_{ij}$. In particular, we may consider the classes

$$\mathcal{C}^* = \{c\sqrt{Q^*} : c > 0\}$$

and

$$\mathcal{D}^* = \{d\sqrt{1/Q^*} : d > 0\},$$

and seek the corresponding optimal estimators in the sense of Eq. (5.1). The answer is immediate and the best estimators are now analogously given by

$$\hat{\sigma}_{pc}^* = \sqrt{Q^*/M^*}$$

and

$$\hat{\eta}_{pc}^* = \sqrt{M^*/Q^*},$$

where M^* is the median of a chi-squared random variable with degrees of freedom $\sum_{i=1}^{m} n_i - 1$.

Is $\hat{\sigma}_{pc}^*$ better than $\hat{\sigma}_{pc}$ in the sense of Eq. (5.1)? The answer to this question is affirmative, and a similar statement can be made about the comparison between $\hat{\eta}_{pc}^*$ and $\hat{\eta}_{pc}$. In fact, both of these assertions follow from a result about the estimation of σ^2 in the similar context. These are all stated as the following theorem.

THEOREM 5.2: *Under the criterion in Eq.* (5.1), *and under the assumptions of Theorem* 5.1,

i. *for estimation of* σ^2, Q^*/M^* *is better than* Q/M,
ii. *for estimation of* σ, $\sqrt{Q^*/M^*}$ *is better than* $\sqrt{Q/M}$,
iii. *for estimation of* σ^{-1}, $\sqrt{M^*/Q^*}$ *is better than* $\sqrt{M/Q}$.

https://doi.org/10.1017/S0269964899132108 Published online by Cambridge University Press

The proof of i (essentially due to [20]) is outlined here. The other two parts trivially follow from that.

PROOF: It suffices to show that Q^*/M^* is better than bQ for any $b > 0$. As $Q^* = Q + \sum_{i=1}^{m} n_i(\bar{x}_i - \bar{\bar{x}})^2$, and random variables

$$\frac{Q}{\sigma^2} \sim \chi^2_{\sum_{i=1}^{m} n_i - m} \quad \text{and} \quad \frac{\sum_{i=1}^{m} n_i(\bar{x}_i - \bar{\bar{x}})^2}{\sigma^2} \sim \chi^2_{m-1}$$

are independently distributed, it follows that the random variables $W = Q/Q^* \sim$ Beta $\left(\sum_{i=1}^{m} n_i - m, m - 1\right)$ and $Q^*/\sigma^2 \sim \chi^2_{\sum_{i=1}^{m} n_i - 1}$ are independently distributed. Now consider

$$P_{\sigma^2}[L(Q^*/M^*, \sigma^2) < L(bQ, \sigma^2)] = P_{\sigma^2}\left[h\left(\frac{Q^*/M^*}{\sigma^2}\right) < h\left(\frac{bQ}{\sigma^2}\right)\right]$$

and observe that the distributions of Q/σ^2 and Q^*/σ^2 are free from the parameter σ^2. Because the events

$$\left\{\frac{bQ}{\sigma^2} < \frac{Q^*/M^*}{\sigma^2} < 1\right\} \quad \text{and} \quad \left\{1 < \frac{Q^*/M^*}{\sigma^2} < \frac{bQ}{\sigma^2}\right\}$$

are mutually exclusive and each of these implies the event

$$\left\{h\left(\frac{Q^*/M^*}{\sigma^2}\right) < h\left(\frac{bQ}{\sigma^2}\right)\right\},$$

we must have

$$P\left[h\left(\frac{Q^*/M^*}{\sigma^2}\right) < h\left(\frac{bQ}{\sigma^2}\right)\right] \geq P\left[\frac{bQ}{\sigma^2} < \frac{Q^*/M^*}{\sigma^2} < 1\right] + P\left[1 < \frac{Q^*/M^*}{\sigma^2} < \frac{bQ}{\sigma^2}\right]$$

$$= P\left[\left\{\frac{Q}{Q^*} < \frac{1}{bM^*}\right\} \cap \left\{\frac{Q^*}{\sigma^2} < M^*\right\}\right]$$

$$+ P\left[\left\{\frac{Q^*}{\sigma^2} > M^*\right\} \cap \left\{\frac{Q}{Q^*} > \frac{1}{bM^*}\right\}\right].$$

Due to the independence of $W = Q/Q^*$ and Q^*/σ^2, it follows that the above expression can be written as

$$P\left[W < \frac{1}{bM^*}\right]P\left[\frac{Q^*}{\sigma^2} < M^*\right] + P\left[W > \frac{1}{bM^*}\right]P\left[\frac{Q^*}{\sigma^2} > M^*\right],$$

and, since M^* is the median of Q^*/σ^2, this is also equal to

$$\frac{1}{2}P\left[W < \frac{1}{bM^*}\right] + \frac{1}{2}P\left[W > \frac{1}{bM^*}\right] = \frac{1}{2},$$

which proves i. Now to prove ii and iii, observe that any loss function for estimating σ or $\eta = \sigma^{-1}$ can be written as some loss function for σ^2. Thus ii and iii can be reformulated in terms of i. Specifically, for illustration let us consider iii, which concerns the estimation of σ^{-1}. With $L(t,\eta) = h(t/\eta)$, define $H(u) = h(1/\sqrt{u})$ and note that $H(u)$ also satisfies the conditions stated in the theorem. Thus,

$$\begin{aligned}
&P_\eta[L(T_1,\eta) < L(T_2,\eta)] \\
&\quad = P_\eta[h(T_1/\eta) < h(T_2/\eta)] \\
&\quad = P_\eta[H(\eta^2/T_1^2) < H(\eta^2/T_2^2)] \\
&\quad = P_{\sigma^2}\left[H\left(\frac{1/T_1^2}{\sigma^2}\right) < H\left(\frac{1/T_2^2}{\sigma^2}\right)\right] \\
&\quad = P_{\sigma^2}\left[L^*\left(\frac{1}{T_1^2},\sigma^2\right) < L^*\left(\frac{1}{T_2^2},\sigma^2\right)\right]
\end{aligned}$$

for some appropriate loss function $L^*(\cdot,\cdot)$ also satisfying the stated conditions. Since $1/T^2 = Q^*/M^*$ is the best estimator of σ^2 in the sense of Eq. (5.1), it follows that the optimum estimator of $\eta = \sigma^{-1}$ is given by $T = \sqrt{M^*/Q^*}$.

6. CLASSIC ESTIMATION OF C_{pm}

Chan, Cheng, and Spring [3] defined the index C_{pm} as

$$C_{pm} = \frac{U-L}{6}\frac{1}{\sqrt{E(X-\tau)^2}}, \tag{6.1}$$

where τ is a known target value of the process. They indicated that C_{pm} may be a better measure of capability of the process when the target value is *known*. To estimate C_{pm} from a normal random sample $x_1,\ldots,x_n$, they propose to estimate $\sqrt{E(X-\tau)^2}$ by $\sqrt{\sum_{i=1}^n (x_i-\tau)^2/(n-1)}$ and use it in Eq. (6.1). Let the corresponding estimate be $\hat{C}_{pm}$. It may be pointed out that a more appropriate estimate of $\sqrt{E(X-\tau)^2}$ to use in Eq. (6.1) may have been $\sqrt{\sum_{i=1}^n (x_i-\tau)^2/n}$. It is because, even though none of the estimate between $\left(\sqrt{\sum_{i=1}^n (x_i-\tau)^2/(n-1)}\right)^{-1}$ and $\left(\sqrt{\sum_{i=1}^n (x_i-\tau)^2/n}\right)^{-1}$ is unbiased for $\left(\sqrt{E(X-\tau)^2}\right)^{-1}$, the latter at least accounts for the appropriate degree of freedom and is also the maximum likelihood estimator. However, use of the former estimator cannot be justified on any optimality ground. The choice made by Chan et al. [3] was perhaps more for want of mathematical convenience, since their proposed estimator was more easily comparable (with respect to bias and mean squared error) with their corresponding estimate of C_p, when $\tau = \mu$, the true process mean.

We will consider the problem of estimation of C_{pm} in a more formal context. As in [3], we will also make the assumption that $\tau = \mu$.

When $\tau = \mu$, the indices C_p and C_{pm} are identical. However, with μ known, the estimators of C_p as proposed earlier in Section 5 will not be necessarily optimum, since none of the corresponding estimation techniques accounts for the extra information available in the form of known μ. It is worth pointing out that there exist situations (for example, for the estimation of σ^2) when, contrary to our intuition, it is better (in the sense of MSE) to ignore the information available in the form of known μ and estimate it by $\bar{x}$ (Birch and Robertson [2], Khattree [13], Gupta and Khattree [7,8], and Nayak [20]).

For simplicity, we will initially consider the situation where a single random sample $x_1, \ldots, x_n$ of size n from a normally distributed process is available, and the process mean μ is known to be equal to τ, the target value. When $\tau = \mu$, $E(X-\tau)^2 = E(X-\mu)^2 = \sigma^2$, and, to estimate $(E(X-\mu)^2)^{-1/2}$, a reasonable starting point is to consider a scalar multiple of $\left(\frac{1}{n}\sum_{i=1}^n (x_i-\mu)^2\right)^{-1/2}$ or of $\left(\sum_{i=1}^n (x_i-\mu)^2\right)^{-1/2}$. Since $\sum_{i=1}^n (x_i-\mu)^2$ is complete-sufficient for σ^2, it follows from the Rao–Blackwell Theorem that an unbiased estimator of $\sigma^{-1} = (E(X-\tau)^2)^{-1/2}$ based on $\left(\sum_{i=1}^n (x_i-\mu)^2\right)^{1/2}$ will also be the minimum variance unbiased estimator. Now, because $\sum_{i=1}^n (x_i-\mu)^2/\sigma^2 \sim \chi_n^2$, we have $E\left(\sum_{i=1}^n (x_i-\mu)^2\right)^{-1/2} = A^{-1}\sigma^{-1}$, where $A = \sqrt{2}\frac{\Gamma[n/2]}{\Gamma[(n-1)/2]} = c_{4,n}\sqrt{(n-1)}$. Thus, $A\left(\sum_{i=1}^n (x_i-\mu)^2\right)^{-1/2}$ is UMVUE of σ^{-1}. Consequently, UMVUE of C_{pm} when $\tau = \mu$ is given by

$$\hat{C}^*_{pm} = A\left(\frac{U-L}{6}\right)\frac{1}{\sqrt{\sum_{i=1}^n (x_i-\tau)^2}}.$$

When the process is not centered at the target τ but the process mean μ is known, then $\frac{1}{\sigma^2}\sum_{i=1}^n (x_i-\tau)^2 \sim \chi_n^2(\lambda)$, the noncentral chi-squared distribution with degrees of freedom n and the noncentrality parameter $\lambda = \frac{n(\tau-\mu)^2}{\sigma^2}$. An argument similar to that given above will result in the UMVUE of C_{pm} as

$$\hat{C}^*_{pm,\tau\neq\mu} = \frac{A^*(U-L)}{6}\frac{1}{\sqrt{\sum_{i=1}^n (x_i-\tau)^2}}$$

where

$$A^* = \sqrt{2}e^{\lambda}\left[\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}\frac{\Gamma[(n-1+2j)/2]}{\Gamma[(n+2j)/2]}\right]^{-1},$$

and

$$\lambda = n(\mu-\tau)^2/\sigma^2.$$

Unfortunately, $\hat{C}^*_{pm,\tau\neq\mu}$ is not a valid estimator, as it involves the unknown parameter σ^2 through λ. One may substitute $n^{-1}\sum_{i=1}^n (x_i-\mu)^2$ as an estimate of σ^2 to obtain an approximate value of λ and use it in the expression $\hat{C}^*_{pm,\tau\neq\mu}$ to obtain an estimator in this case. From the quality control point of view, this may not, however, be a correct approach. Instead, one should perhaps reset the process to bring the process mean μ at the intended target τ. In view of this, the case $\tau \neq \mu$ will not be considered in the next section.

7. PITMAN CLOSEST ESTIMATOR OF C_{pm} WHEN $\tau = \mu$

As earlier, it may be of interest to find the estimator of C_{pm} which is closest in the sense of the Pitman measure of closeness. Alternatively, we can concentrate on the Pitman closest estimator of $(E(X-\tau)^2)^{-1/2}$. The specific classes of estimators which are of interest are

$$\mathcal{A} = \left\{ a \cdot \left[\sum_{i=1}^{n} (x_i - \bar{x})^2 \right]^{-1/2}, \qquad a > 0 \right\}$$

and

$$\mathcal{B} = \left\{ b \cdot \left[\sum_{i=1}^{n} (x_i - \tau)^2 \right]^{-1/2}, \qquad b > 0 \right\}.$$

Note that the estimator proposed in [3] belongs to class $\mathcal{A}$. The following theorem, proof of which follows from [13], provides the answer.

THEOREM 7.1: *The best estimators of $\eta = (E(X-\tau)^2)^{-1/2}$ in classes $\mathcal{A}$ and $\mathcal{B}$, in the sense of Eq.* (5.1) *and under the same assumptions on $L(T,\theta)$, as stated in Theorem* 5.1, *are respectively given by*

$$\hat{\eta}^{**}_{pm} = \left[M^{**} \Big/ \sum_{i=1}^{n} (x_i - \bar{x})^2 \right]^{1/2}$$

$$\hat{\eta}^{\#\#}_{pm} = \left[M^{\#\#} \Big/ \sum_{i=1}^{n} (x_i - \tau)^2 \right]^{1/2},$$

where M^{**} and $M^{\#\#}$ are the medians of χ^2_{n-1} and χ^2_n, respectively. The corresponding estimates of C_{pm} are given by $\hat{C}^{**}_{pm} = (U - L/6)\,\hat{\eta}^{**}_{pm}$ and $\hat{C}^{\#\#}_{pm} = (U - L/6)\,\hat{\eta}^{\#\#}_{pm}$.

While $\hat{\eta}^{**}_{pm}$ and $\hat{\eta}^{\#\#}_{pm}$ are optimal in the respective classes $\mathcal{A}$ and $\mathcal{B}$, one may still ask, of the two, which is preferred? The following theorem provides the answer to this inquiry.

THEOREM 7.2: *Under the same assumptions on $L(T,\theta)$ as in Theorem* 5.1, *$\hat{\eta}^{\#\#}_{pm}$ is closer to η than $\hat{\eta}^{**}_{pm}$ is to η. Thus, $\hat{C}^{\#\#}_{pm}$ is closer to C_{pm} than $\hat{C}^{**}_{pm}$ is to C_{pm}.*

Proof of the above theorem is along the same lines as that for Theorem 5.2, hence it is not repeated here.

The case where there are m random subsamples of sizes $n_1, \ldots, n_m$ does not require any special treatment because, with known mean $\mu = \tau$, these subsamples can be collectively treated as a random sample of size $N = \Sigma n_i$. A class of estimators analogous to $\mathcal{A}$, say $\mathcal{A}_{\text{pooled}}$ of η which consists of scalar multiples of $\left[\sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2\right]^{-1/2}$, can still be defined. However, it follows from Theorem 7.2 that the optimal (and in fact, any) estimator in that class will be dominated in the sense of Eq. (5.1), by $\hat{\eta}_{pm}^{\#\#} = \left[M_N^{\#\#}/\sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \tau)^2\right]^{1/2}$, where $M_N^{\#\#}$ is the median of a χ_N^2. Thus, no special treatment of this scenario is needed.

We have not explicitly made any mention of maximum likelihood estimation in our discussion. Since the maximum likelihood estimators (MLE) of σ^2 under the situations discussed here are well known, the MLE of C_p and C_{pm} are trivially obtained by the appropriate substitutions. It may further be observed that the classes $\mathcal{C}$, $\mathcal{D}$, $\mathcal{C}^*$, $\mathcal{D}^*$, and $\mathcal{B}$ do indeed contain the maximum likelihood estimators (under the appropriate context) which in each case are dominated by the corresponding optimal estimator within that class, in the sense of the Pitman measure of closeness. Thus, a detailed comparison with maximum likelihood estimators is really not necessary. Similarly, the MMSE estimators (and in general, the minimum risk estimators under the loss functions L_1–L_5) of C_p and C_{pm} within these classes can be obtained by applying the routine calculus steps. However, as in the case of MLE, the domination of these estimators under the criterion of Pitman measure of closeness trivially follows.

Acknowledgment

The author would like to thank a referee for helpful comments.

References

1. Angus, J.E. (1997). Improved estimation of σ in quality control, Revisited. *Probability in the Engineering and Informational Sciences* 9: 37–42.
2. Birch, J.J. & Robertson, T. (1983). A classroom note on the sample variance and the second moment. *American Mathematical Monthly* 90: 703–705.
3. Chan, L.K., Cheng, S.W., & Spring, F.A. (1988). A new measure of process capability: C_{pm}. *Journal of Quality Technology* 20: 162–175.
4. Chou, Y.-M. & Owen, D.B. (1989). On the distributions of the estimated process capability indices. *Communications in Statistics, Theory and Methods* 18: 4549–4560.
5. Derman, C. & Ross, S. (1995). An improved estimator of σ in quality control. *Probability in the Engineering and Informational Sciences* 9: 411–415.
6. Ghosh, M. & Sen, P.K. (1989). Median unbiasedness and Pitman closeness. *Journal of the American Statistical Association* 84: 1089–1091.
7. Gupta, R.D. & Khattree, R. (1993). A comparison of some estimators of the variance covariance matrix when the population mean is known. *Statistics* 24: 331–345.
8. Gupta, R.D. & Khattree, R. (1994). Some comments on maximum likelihood estimators of scale parameter or its square. *Journal of Indian Statistical Association* 32: 123–131.
9. Keating, J.P. (1985). More on Rao's phenomenon. *Sankhyā* B47: 18–21.
10. Keating, J.P. & Gupta, R.C. (1984). Simultaneous comparison of scale estimators. *Sankhyā* B46: 275–280.
11. Keating, J.P., Mason, R.L., & Sen, P.K. (1993). Pitman's measure of closeness: A comparison of statistical estimators. Philadelphia: Society for Industrial and Applied Mathematics.
12. Khattree, R. (1987). On comparison of estimates of dispersion using generalized Pitman nearness criterion. *Communications in Statistics, Theory and Methods* 16: 263–274.
13. Khattree, R. (1992). Comparing estimators for population variance using Pitman nearness. *American Statistician* 46: 214–217.

14. Khattree, R. & Gill, D.S. (1987). Estimation of signal to noise ratio using Mahalanobis distance. *Communications in Statistics, Theory and Methods* 16: 897–907.
15. Khattree, R. & Gupta, R.D. (1989). Estimation of matrix valued signal to noise ratio. *Journal of Multivariate Analysis* 30: 312–327.
16. Khattree, R. & Gupta, R.D. (1990). Estimation of realized signal to noise ratio for a pair of multivariate signals. *Australian Journal of Statistics* 32: 239–246.
17. Kirmani, S.N.U.A., Kocherlakota, K., & Kocherlakota, S. (1991). Estimators of σ and the process capability index based on subsamples. *Communications in Statistics, Theory and Methods* 20: 275–291.
18. Montgomery, D.C. (1996). *Introduction to statistical quality control*. New York: John Wiley.
19. Nayak, T.K. (1990). Estimation of location and scale parameters using generalized Pitman nearness criterion. *Journal of Statistical Planning and Inference* 24: 259–268.
20. Nayak, T.K. (1994). Pitman nearness comparison of some estimators of population variance. *American Statistician* 48: 99–103.
21. Rao, C.R. (1980). Discussion of "Minimum chi-square, not maximum likelihood!" by J. Berkson. *The Annals of Statistics* 8: 482–485.
22. Rao, C.R. (1981). Some comments on the minimum mean square error as a criterion in estimation. In *Statistics and related topics*. Amsterdam: North Holland, pp. 123–143.
23. Sen, A. & Khattree, R. (1999). On estimating the current intensity of failure for the power law process. *Journal of Statistical Planning and Inference* (to appear).
24. Sen, P.K., Nayak, T.K., & Khattree, R. (1991). Comparison of estimators of a dispersion matrix under generalized Pitman nearness criterion. *Communications in Statistics, Theory and Methods* 20: 3473–3486.