J. Statist. Comput. Simul., 2002, Vol. 72(9), pp. 707–718

A PROCESS CAPABILITY INDEX THAT IS BASED ON THE PROPORTION OF CONFORMANCE

MICHAEL PERAKIS and EVDOKIA XEKALAKI*

Department of Statistics, Athens University of Economics and Business, 76 Patision St., 104 34, Athens, Greece

(Received 19 October 2001; In final form 20 March 2002)

In this paper a new process capability index is proposed, which is based on the proportion of conformance of the process and has several appealing features. This index is simple in its assessment and interpretation and is applicable to normally or non-normally distributed processes. Likewise, its value can be assessed for continuous or discrete processes, it can be used under either unilateral or bilateral tolerances and the assessment of confidence limits for its true value is not very involved, under specific distributional assumptions. Point estimators and confidence limits for this index are investigated, assuming two very common continuous distributions (normal and exponential).

Keywords: Process capability indices; Proportion of conformance; Confidence limits; Normal distribution; Exponential distribution

1 INTRODUCTION

Process capability indices (PCIs) aim to quantify the capability of a process (X) to meet some specifications that are related to a measurable characteristic of its produced items. These specifications are determined through the lower specification limit (L), the upper specification limit (U) and the target value (T). A variety of such indices has been developed in the last two decades. Unambiguously, the most prominent among them are C_p, C_{pk}, C_{pm} and C_{pmk}. A comprehensive discussion of these four indices (and many other indices too) is provided by Kotz and Johnson (1993), Kotz and Lovelace (1998) and Kotz and Johnson (2002).

A deficiency of the indices C_p; C_{pk}, C_{pm} and C_{pmk} is that their values can be assessed solely for processes for which both the lower and the upper specification limits have been specified. Nevertheless, sometimes one may be faced with a unilateral process, *i.e.* a process for which only a lower or an upper specification limit has been set. In such cases, the assessment of these four indices becomes impossible and we have to resort to the indices CPL and CPU suggested by Kane (1986). Therefore, it is apparent that there exists a limitation in the use of the four standard indices.

* Corresponding author. E-mail: exek@aueb.gr

ISSN 0094-9655 print; ISSN 1563-5163 online
DOI: 10.1080/00949650290031471

One further weakness of the four standard PCIs, is that their definition is based on the restrictive assumption that the distribution of the process is normal. Even though this assumption is not quite evident, it is reflected from the constants that appear in their denominators, whose selection is related to the properties of the normal distribution. This has the implication that their assessment and interpretation for processes not normally distributed may lead to utterly wrong conclusions. In order to overcome this deficiency, Clements (1989) proposed some generalizations of C_p and C_{pk}, while Pearn and Kotz (1994–1995) extended the initial idea of Clements to C_{pm} and C_{pmk}.

Despite the fact that these generalized indices can be used regardless of the form of the distribution of the process, they have some drawbacks. In particular, they do not have a direct relationship with the proportion of conformance (p) of the process, defined as the probability of producing within specifications (sometimes instead of this term the term yield is used). Besides, they cannot be assessed if the examined process is discrete and the derivation of the distributions of their estimators and the construction of confidence limits for them are very cumbersome without assuming particular distributional forms (see, *e.g.* Perakis and Xekalaki, 1999). However, even in this case the distributions of the estimators of the more involved generalized indices, such as C_{pm} and C_{pmk}, are intractable thus making the construction of confidence limits for them a rather difficult task. Borges and Ho (2000) derived the approximate distribution of the estimator of the generalized index C_{pk} making no specific distributional assumptions. They also proposed a method that leads to approximate conservative lower confidence limits for the true value of the same index.

Another important disadvantage of the four standard indices and their generalizations is that they are not applicable to situations where except for the specification limits, a minimum acceptable yield has been set. For instance, let us assume that we have two processes, say A and B. The yield of process A is supposed to be 0.996, while that of process B is 0.985. Moreover, the customer of the product produced from process A requires a yield at least equal to 0.997, while the customer of the product produced from process B is less strict and requires a yield at least equal to 0.980. Obviously, in such situations the capability of a process does not depend solely on the specification limits, while it is also dependent on the value of the minimum acceptable yield. This value may differ, depending on the nature of the process. Therefore, a yield that is regarded as high for a process may be regarded as low for another. Indeed, according to the requirements of the customer of process B, the process is capable since its yield exceeds the minimum allowable yield. On the other hand, even though the yield of process A is higher than that of process B, the customer will not regard it as capable because its yield is smaller than its minimum acceptable value. The use of the four standard PCIs is not appropriate in such situations, since their values do not take into account the minimum acceptable yield of the process.

In the next section, a new index is defined that overcomes the deficiencies of the standard PCIs described above and has several appealing features. Its basic properties are examined and a general estimator of it is proposed. Sections 3 and 4 deal with the new index in the case of normal and exponential processes. Under each distributional assumption, point estimators for the new index are suggested and lower confidence limits (lcls) are constructed. Upper confidence limits can also be obtained in a similar manner. However, only the case of lcls is considered as these are of greater interest (due to the fact that large values of PCIs are desirable) and have thus attracted a lot of attention in the literature (see, *e.g.* Chou *et al.*, 1990; Li *et al.*, 1990; Boyles, 1991; Kushler and Hurley, 1992). Finally, in Section 5 the obtained results are briefly discussed and some topics that deserve further investigation in connection with the new index are pointed out.

2 THE INDEX C_{pc}

In this section, a new PCI is introduced which overcomes the drawbacks of the standard indices discussed in the previous section. It is defined as the ratio

$$C_{pc} = \frac{1-p_0}{1-p},$$

where p_0 denotes the minimum allowable proportion of conformance (pc). The value of p_0 must be intuitively close to unity and depends on the nature of the examined process and the requirements of the customers.

It is worth noting at this point, that this index has some further appealing features:

- it is applicable even to situations of the type "quantify the capability of a process if the specification limits are L, U and the minimum allowable yield is p_0"
- it can be assessed under either unilateral or bilateral specifications
- its definition is not based on the usual assumption of normality
- it is directly associated with p and thus its interpretation in its terms is simple
- it satisfies (under some conditions) the property that all the standard PCI have, which is, to be equal to the unity if the process mean is located exactly at the midpoint of the specification area (*i.e.* $\mu = M = (L+U)/2$) and $p = 0.9973$
- its assessment is relatively simple
- it can be readily understood by practitioners
- assuming that the distribution of the process is normal, the assessment of confidence limits for its true value is fairly simple
- under specific non-normal distributional assumptions the construction of confidence limits for its true value is feasible
- it can be assessed for discrete processes as well

As mentioned before, the value of the minimum allowable proportion of conformance p_0 is intuitively required to be close to 1. Whenever the value of p_0 is not specified, the value 0.9973 seems to be a plausible choice. By this selection the index has the property of being equal to unity if $p = 0.9973$ and the process mean equals M – a property that all the standard PCIs possess. Note that despite the fact that the value 0.9973 is connected with the properties of the normal distribution, it plays an important role in statistical process control and is customarily regarded as a sufficiently large value of p.

In the sequel, for simplicity, it is assumed that $p_0 = 0.9973$. Nonetheless, the discussion given below can be modified easily assuming any other value of p_0. Selecting p_0 to be equal to 0.9973, the index C_{pc} can be rewritten in the form

$$C_{pc} = \frac{0.0027}{1-p}. \tag{2.1}$$

From the definition of C_{pc} it follows that the yield of the process can be expressed directly in terms of this index. Indeed, solving (2.1) for p, it follows that $p = 1 - 0.0027/C_{pc}$.

At this point, it would be interesting to remark that C_{pc} can be associated with Yeh and Bhattacharya's (1998) index defined as

$$C_f = \min\left\{\frac{p_0^L}{p_1}, \frac{p_0^U}{p_2}\right\},$$

where $p_1 = P(X < L)$, $p_2 = P(X > U)$, and p_0^L, p_0^U are the expected proportions of nonconforming products that the manufacturer can tolerate on the lower and upper specification limits, respectively. The value of C_{pc} coincides with that of C_f provided that the process is symmetric, its mean coincides with M, and $p_0^L = p_0^U$. However, separating p into two parts (p_1 and p_2) and taking their minimum leads (for any assumption on the distribution of the process) to a distribution of the estimator of the index with a quite involved form and thus the construction of confidence limits for the index C_f becomes extremely tough without resorting to the method of bootstrap. Moreover, the yield of the process cannot be expressed directly as a function of the index C_f-something that is possible in the case of the index C_{pc}.

Let us now examine the behavior of the index C_{pc} for different values of p. Obviously if $p = 0.9973$, then $C_{pc} = 1$. If the yield of the process is greater than 0.9973, then $C_{pc} > 1$ and as p approaches unity it tends to infinity. On the other hand, if $p < 0.9973$, then $C_{pc} < 1$ and the value of the index becomes negligible as p tends to zero. Obviously, the smallest value that p may take is zero and so the smallest possible value of this index is 0.0027 (or, generally, $1 - p_0$).

The only unknown parameter involved in the definition of the index C_{pc} is p. Its value depends on the form of the distribution of the process and can be estimated on the basis of a random sample collected from the examined process. If we denote the estimator of p by $\hat{p}$ (the functional form of $\hat{p}$ depends on the distribution of the process), then the index C_{pc} can be estimated by

$$\hat{C}_{pc} = \frac{0.0027}{1 - \hat{p}}. \tag{2.2}$$

In the next two sections the properties of estimator (2.2) are examined for two very common distributions: the normal and the exponential.

3 THE INDEX C_{pc} FOR NORMAL PROCESSES

Assuming normality of the examined process, its yield is given by

$$p = P(L < X < U) = \Phi\left(\frac{U - \mu}{\sigma}\right) - \Phi\left(\frac{L - \mu}{\sigma}\right),$$

where μ, σ denote the mean and the standard deviation of the process, respectively and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The value of p can be estimated using the estimator

$$\hat{p} = \Phi\left(\frac{U - \bar{X}}{S}\right) - \Phi\left(\frac{L - \bar{X}}{S}\right),$$

where $\bar{X}, S$ are the sample mean and the sample standard deviation respectively, obtained from a random sample. As Kotz and Johnson (1993) stress, the estimator $\hat{p}$ is biased. However, due to the complexity of the minimum variance unbiased estimator of p, they conclude that the use of $\hat{p}$ is in most of the times fairly adequate.

Three alternative techniques for constructing approximate lower confidence limits for the true value of p are described by Wang and Lam (1996), while Perakis and Xekalaki (2001) consider a modification of one of them which achieves a better coverage. In view of the fact that the index C_{pc} is a function of p, it becomes evident that a proper modification of these methods makes the construction of lcls for the index C_{pc} feasible. The four methods and their implementation for the construction of lcls for C_{pc} are discussed in the sequel.

The first two methods are based on the articles of Owen and Hua (1977) and Chou and Owen (1984), respectively. In particular, Owen and Hua (1977) sought confidence limits for the probabilities $p_1 = P(X < L)$ and $p_2 = P(X > U)$, while Chou and Owen (1984) constructed one-sided simultaneous confidence regions on the lower (p_1) and the upper (p_2) tail areas of the normal distribution. Extensive tables containing the values of the confidence limits obtained from these two methods are included in the corresponding articles. Denoting the upper confidence limits for p_1 and p_2 obtained by any of these two methods by p_1^* and p_2^*, respectively, Wang and Lam (1996) stress that a $100(1-\alpha)\%$ lcl for p is given by

$$1 - p_1^* - p_2^*. \tag{3.1}$$

This limit can be used as a basis for constructing a lcl for C_{pc}. Actually, starting from lcl (3.1) one may find that a $100(1-\alpha)\%$ approximate lcl for the index C_{pc} is given by

$$\frac{0.0027}{p_1^* + p_2^*}. \tag{3.2}$$

An obvious drawback of the two methods described above is that their implementation requires use of the entries of the tables provided in the corresponding articles for the confidence limits of p_1 and p_2. This restricts their use and thus Wang and Lam (1996) introduced a third method for constructing an approximate lcl for p which is much simpler and does not require any special tables. In particular, Wang and Lam (1996) sought a value p^* such that $P(p \geq p^*) = 1 - \alpha$. Substituting $P(L < X < U)$ for p they obtained

$$P[P(\bar{X} - K_1 S < X < \bar{X} + K_2 S) \geq p^*] = 1 - \alpha,$$

where $K_1 = (\bar{X} - L)/S$ and $K_2 = (U - \bar{X})/S$. Standardizing the terms, so as to result in the standard normal distribution function, the left hand side term in this equation reduces to

$$P\left[\Phi\left(\frac{Z}{\sqrt{n}} + K_2\sqrt{\frac{Y}{n-1}}\right) - \Phi\left(\frac{Z}{\sqrt{n}} - K_1\sqrt{\frac{Y}{n-1}}\right) \geq p^*\right],$$

where $Z = \sqrt{n}(\bar{X} - \mu)/\sigma$ follows the standard normal distribution and $Y = (n-1)S^2/\sigma^2$ follows the chi-square distribution with $n - 1$ degrees of freedom (df). Wang and Lam (1996) point out that if the probabilities that K_1 and K_2 are negative are negligible, this probability can be approximated rather accurately through

$$P\left[\Phi\left(\frac{1}{\sqrt{n}} + K_2\sqrt{\frac{Y}{n-1}}\right) - \Phi\left(\frac{1}{\sqrt{n}} - K_1\sqrt{\frac{Y}{n-1}}\right) \geq p^*\right]$$

and concluded that the desired $100(1-\alpha)\%$ approximate lcl for p is given by

$$p^* = \Phi\left(\frac{1}{\sqrt{n}} + \max(K_1, K_2)\sqrt{\frac{\chi^2_{n-1,\alpha}}{n-1}}\right) - \Phi\left(\frac{1}{\sqrt{n}} - \min(K_1, K_2)\sqrt{\frac{\chi^2_{n-1,\alpha}}{n-1}}\right), \tag{3.3}$$

where $\chi^2_{n-1,\alpha}$ denotes the α quantile of the chi-square distribution with $(n-1)$ df. Based on lcl (3.3) one can easily find that a $100(1-\alpha)\%$ approximate lcl for C_{pc} is given by

$$\frac{0.0027}{1-p^*}. \tag{3.4}$$

Perakis and Xekalaki (2001) showed that the coverage of lcl (3.3) can be improved substantially using a variant of it given by

$$\tilde{p} = \Phi\left(\frac{1}{\sqrt{n}} + C_1\right) - \Phi\left(\frac{1}{\sqrt{n}} - C_2\right), \tag{3.5}$$

where

$$C_1 = \max(K_1, K_2)\left(1 + \frac{1}{n}\right)\sqrt{\frac{\chi^2_{n-1,\alpha}}{n-1}}$$

and

$$C_2 = \min(K_1, K_2)\left(1 + \frac{1}{n}\right)\sqrt{\frac{\chi^2_{n-1,\alpha}}{n-1}}.$$

Taking advantage of this result, one can obtain an alternative lcl for C_{pc}. Indeed, from (3.5) it can be shown easily that a $100(1-\alpha)\%$ approximate lcl for C_{pc} is given by

$$\frac{0.0027}{1-\tilde{p}}. \tag{3.6}$$

In order to investigate the performance of lcls (3.4) and (3.6) an extensive simulation study was conducted. Specifically, random samples of various sizes were generated from the normal distribution for various parameter combinations. For every combination of parameters and sample size, the total number of the generated samples was 25,000. For each of these samples, lcls for the index C_{pc} were assessed using formulae (3.4) and (3.6) and the percentage of times at which the actual index value exceeded each of the two lcls was calculated.

TABLE I The Examined Processes and Their C_{pc} Values.

k_1	k_2	μ	σ	C_{pc}	k_1	k_2	μ	σ	C_{pc}
1	1	15	5	0.00850	3	4	14.285	1.428	1.95403
1	2	13.333	3.333	0.01488	3	5	13.75	1.25	1.99947
1	3	12.5	2.5	0.01687	3	6	13.333	1.111	1.99989
1	4	12	2	0.01701	3	7	13	1	2.0000
1	5	11.666	1.667	0.01701	4	4	15	1.25	42.6023
1	6	11.428	1.428	0.01701	4	5	14.444	1.111	84.4396
1	7	11.25	1.25	0.0170	4	6	14	1	85.2020
2	2	15	2.5	0.05933	4	7	13.636	0.909	85.2049
2	3	14	2	0.1120	5	5	15	1	4701.76
2	4	13.333	1.667	0.11850	5	6	14.545	0.909	9371.21
2	5	12.857	1.428	0.11867	5	7	14.166	0.833	943.09
2	6	12.5	1.25	0.11867	6	6	15	0.833	1363367.5
2	7	12.222	1.111	0.1186	6	7	14.615	0.769	2737338.6
3	3	15	1.667	0.99995	7	7	15	0.714	10536237

TABLE II The Observed Coverage of lcl (3.4) (Top Value) and lcl (3.6) (Bottom Value) for Confidence Level 0.9.

k_1 k_2	$n = 25$	$n = 50$	$n = 100$	$n = 200$	k_1 k_2	$n = 25$	$n = 50$	$n = 100$	$n = 200$
1, 1	0.95248	0.94492	0.93252	0.92392	3, 4	0.94248	0.94164	0.9432	0.94284
	0.92668	0.91664	0.92024	0.91004		0.91092	0.91604	0.92736	0.93004
1, 2	0.95288	0.95192	0.9538	0.95108	3, 5	0.93868	0.9416	0.94404	0.94168
	0.933	0.93636	0.93968	0.94484		0.90828	0.92052	0.9298	0.93236
1, 3	0.9512	0.948	0.94516	0.94788	3, 6	0.94052	0.93776	0.9418	0.94444
	0.92928	0.93116	0.93924	0.9368		0.9106	0.9228	0.92596	0.93424
1, 4	0.94404	0.9452	0.9438	0.94208	3, 7	0.942	0.94392	0.94076	0.94328
	0.9286	0.93104	0.9336	0.93848		0.90876	0.91684	0.9276	0.93196
1, 5	0.94944	0.9454	0.94536	0.94556	4, 4	0.94144	0.93648	0.93152	0.92628
	0.92988	0.93048	0.93664	0.93764		0.91016	0.91256	0.91288	0.91124
1, 6	0.94724	0.94612	0.94484	0.94356	4, 5	0.93492	0.93632	0.9378	0.93712
	0.9256	0.92988	0.93596	0.93548		0.89916	0.91484	0.91872	0.92608
1, 7	0.9462	0.94656	0.94396	0.9442	4, 6	0.9364	0.9348	0.9376	0.93836
	0.92568	0.929	0.93232	0.93616		0.89904	0.91056	0.91964	0.9286
2, 2	0.94732	0.94012	0.9322	0.92432	4, 7	0.93468	0.93608	0.93884	0.93792
	0.92812	0.93072	0.93448	0.93464		0.90408	0.91064	0.91588	0.92156
2, 3	0.94868	0.94776	0.95064	0.9466	5, 5	0.94028	0.93496	0.93104	0.92556
	0.92224	0.92792	0.93304	0.93456		0.90388	0.91056	0.908	0.9116
2, 4	0.944	0.94748	0.94788	0.95	5, 6	0.93208	0.93196	0.93356	0.932
	0.91572	0.92764	0.9354	0.93888		0.89352	0.9056	0.9136	0.91872
2, 5	0.949	0.94764	0.94624	0.94788	5, 7	0.9282	0.93108	0.93328	0.93396
	0.92052	0.92984	0.93368	0.93952		0.89452	0.90928	0.91408	0.91764
2, 6	0.94508	0.94804	0.94692	0.94968	6, 6	0.93564	0.9326	0.92628	0.92112
	0.91992	0.92684	0.93268	0.93992		0.90168	0.90712	0.90764	0.912
2, 7	0.94532	0.94768	0.94812	0.94752	6, 7	0.9254	0.9282	0.9284	0.92828
	0.92084	0.9286	0.93208	0.93712		0.8916	0.90184	0.90964	0.91732
3, 3	0.94516	0.934	0.9334	0.92588	7, 7	0.86356	0.93016	0.92712	0.92396
	0.91588	0.916	0.915	0.90996		0.86564	0.9044	0.91116	0.9072

TABLE III The Observed Coverage of lcl (3.4) (Top Value) and lcl (3.6) (Bottom Value) for Confidence Level 0.95.

k_1 k_2	$n = 25$	$n = 50$	$n = 100$	$n = 200$	k_1 k_2	$n = 25$	$n = 50$	$n = 100$	$n = 200$
1, 1	0.97904 0.96204	0.9722 0.95884	0.9674 0.95996	0.9652 0.95448	3, 4	0.97076 0.954	0.9706 0.9592	0.97312 0.96352	0.97132 0.96484
1, 2	0.97748 0.96556	0.97744 0.96684	0.97388 0.97008	0.9744 0.96736	3, 5	0.96956 0.95556	0.96972 0.95836	0.97332 0.96512	0.97148 0.96372
1, 3	0.97268 0.95888	0.9712 0.96332	0.97056 0.96208	0.97004 0.963	3, 6	0.96804 0.95464	0.97092 0.96164	0.972 0.9638	0.97304 0.96408
1, 4	0.96976 0.9574	0.96912 0.95848	0.96628 0.95844	0.9648 0.95956	3, 7	0.97188 0.95364	0.97188 0.95852	0.97192 0.96456	0.97296 0.9674
1, 5	0.96844 0.95644	0.96828 0.957	0.96644 0.95952	0.96456 0.96124	4, 4	0.973 0.95768	0.97148 0.95712	0.96848 0.95692	0.96464 0.95652
1, 6	0.97248 0.96036	0.96884 0.95756	0.96676 0.95996	0.96532 0.96072	4, 5	0.96908 0.94908	0.96816 0.95572	0.96752 0.9586	0.97096 0.9634
1, 7	0.96964 0.95504	0.96788 0.95872	0.96612 0.96052	0.96564 0.96196	4, 6	0.9672 0.95112	0.9682 0.95428	0.96956 0.95988	0.97036 0.96276
2, 2	0.97512 0.96152	0.97112 0.95664	0.96804 0.95708	0.96348 0.9572	4, 7	0.96576 0.95152	0.97004 0.9544	0.969 0.95852	0.96864 0.96336
2, 3	0.97464 0.95992	0.9744 0.96448	0.97584 0.9654	0.97296 0.96688	5, 5	0.97156 0.95352	0.96984 0.95648	0.96676 0.95372	0.96268 0.95728
2, 4	0.973 0.95796	0.97228 0.96392	0.97252 0.96564	0.97376 0.96752	5, 6	0.96704 0.946	0.96872 0.95216	0.96752 0.95764	0.96728 0.95852
2, 5	0.9728 0.95936	0.97428 0.96172	0.97312 0.96332	0.97164 0.96556	5, 7	0.9676 0.94704	0.9674 0.95448	0.96868 0.95784	0.96724 0.96076
2, 6	0.97108 0.9602	0.97388 0.96104	0.97556 0.9656	0.97118 0.9694	6, 6	0.9702 0.94952	0.97004 0.95356	0.96704 0.95516	0.96476 0.95312
2, 7	0.971 0.9568	0.97396 0.96376	0.97164 0.96556	0.97304 0.96852	6, 7	0.9624 0.94336	0.96548 0.94868	0.96696 0.95492	0.96616 0.9584
3, 3	0.97388 0.96092	0.97252 0.95888	0.96652 0.9594	0.96412 0.95756	7, 7	0.86424 0.86584	0.94796 0.94652	0.9656 0.95208	0.964 0.95296

The sample sizes selected were $n = 25$, 50, 100 and 200. It was assumed that $L = 10$ and $U = 20$ and various combinations of $k_1 = (\mu - L)/\sigma$ and $k_2 = (U - \mu)/\sigma$ were chosen. Due to the symmetry of the normal distribution, processes with $\mu > 15$ (*i.e.* $k_1 > k_2$) have not been considered. The examined processes and their C_{pc} values are summarized in Table I, while the results of the study are summarized in Tables II and III.

The top value in the entries of the latter two tables corresponds to the observed coverage (OC) of the lcl given by (3.4), while the bottom value corresponds to that of lcl (3.6). Table II refers to the OC for confidence level 0.90, while Table III refers to the OC for confidence level 0.95. One may observe that in almost all the cases the OC of lcl (3.4) is greater than the nominal coverage. This implies that lcl (3.4) is conservative, namely it includes the actual value at a higher percentage than the nominal. In addition, the OC seems to approach the nominal coverage as the sample size increases, only in the case where $k_1 = k_2$. In all the other cases, the improvement that is achieved increasing the sample size seems to be unimportant. Finally, it is obvious that in almost all the cases the OC of lcl (3.6) is closer to the nominal than that of lcl (3.4).

4 THE INDEX C_{pc} FOR EXPONENTIAL PROCESSES

In practical situations, one may often be faced with processes whose distributions are far from being normal. In this section, the index C_{pc} is considered under the assumption that the underlying distribution of the examined process is of a non-normal form and, in particular, exponential. Such a distributional assumption is generally valid for data that have a natural one-sided boundary (frequently located at zero) with a large probability mass concentrated near this boundary. Examples of cases where this distributional assumption seems to be reasonable are the time required until failure in a life test, the surface roughness and the coating thickness, as Gunter (1989) points out. It should also be remarked that, the fact that the exponential distribution arises frequently in industrial processes is also pointed out in the article by Yeh and Bhattacharya (1998).

In measuring the capability of a process whose distribution can be regarded to be the exponential distribution, it is reasonable to assume that in most of the times only one specification limit has been set. This may be either U (*e.g.* surface roughness), or L (*e.g.* time until failure). In the sequel, to avoid confusion, we adopt the notation C_{pcu} for the case of an upper specification limit and C_{pcl} for the case of a lower specification limit. In the rare case where both U and L have been assigned, the distribution of the estimator of the index C_{pc} becomes quite complicated and therefore this case is not considered here.

If only U has been set, the process yield is given by $P(X < U) = 1 - e^{-\lambda U}$ and the index C_{pcu} is defined as

$$C_{pcu} = \frac{0.0027}{e^{-\lambda U}}. \tag{4.1}$$

On the other hand, if only L has been specified, $p = P(X > L) = e^{-\lambda L}$ and thus the index C_{pcl} is defined as

$$C_{pcl} = \frac{0.0027}{1 - e^{-\lambda L}}. \tag{4.2}$$

The only unknown parameter that is involved in the expressions of the indices C_{pcu} and C_{pcl} is λ. In practice, the true value of λ is unknown, and hence the need for its estimation becomes evident. As is well known, if a random sample $X_1, \ldots, X_n$ from the exponential distribution with parameter λ is available, the maximum likelihood estimator (mle) $\hat{\lambda}$ of λ is the reciprocal of the sample mean, *i.e.*

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} X_i} = \bar{X}^{-1}.$$

Substituting $\hat{\lambda}$ for λ in formulae (4.1) and (4.2) results in the following estimators of C_{pcu} and C_{pcl}:

$$\hat{C}_{\text{pcu}} = \frac{0.0027}{e^{-\hat{\lambda}U}}, \qquad \hat{C}_{\text{pcl}} = \frac{0.0027}{1 - e^{-\hat{\lambda}L}}.$$

The probability density functions (pdfs) of $\hat{C}_{\text{pcu}}$ and $\hat{C}_{\text{pcl}}$ are given in Theorems 4.1 and 4.2, which are given below. The results stated in both these theorems can be verified readily.

THEOREM 4.1 *The pdf of $W = \hat{C}_{\text{pcu}}$ is given by*

$$f_w(w) = \frac{(Un\lambda)^n}{\Gamma(n)} (5.9145 + \ln w)^{-n-1} w^{-1} e^{(-\lambda Un)/(5.9145 + \ln w)}, \quad 0.0027 < w < \infty.$$

THEOREM 4.2 *The pdf of $W = \hat{C}_{\text{pcl}}$ is given by*

$$f_w(w) = \frac{(\lambda nL)^n}{\Gamma(n)} \left(-\ln\left(\frac{1 - 0.0027}{w} \right) \right)^{-n-1} e^{\lambda nL/\ln(1-0.0027/w)} \frac{0.0027}{w^2 - 0.0027w}, \quad 0.0027 < w < \infty.$$

The form of the pdf of $\hat{C}_{\text{pcu}}$ is depicted in Figure 1 for three different sample sizes ($n = 25$, 50 and 100) and for $U = 10$ and $\lambda = 0.6$. The vertical line corresponds to the actual value of

FIGURE 1 The probability density function of $W = \hat{C}_{\text{pcu}}$ for various sample sizes assuming that $U = 10$ and $\lambda = 0.6$.

C_{pcu}, which, for the selected combination of U and λ, equals 1.089. One may observe that as the value of n increases, the amount of mass concentrated near the actual value of the index increases and the mode of the distribution reaches the actual value of C_{pcu}.

The remaining part of this section deals with the problem of constructing lcls for the indices C_{pcu} and C_{pcl}. In order to construct a $100(1-\alpha)\%$ exact lcl for C_{pcu}, one first has to construct a $100(1-\alpha)\%$ exact lcl for λ. Considering that the distribution of $Y = \sum X_i$, is the Gamma, it can be proved that the distribution of the auxiliary random variable $Z = \lambda Y$ does not depend on λ. Actually, the pdf of Z is of the form $g(z) = [\Gamma(n)]^{-1} z^{n-1} e^{-z}$, $z > 0$. The assessment of a $100(1-\alpha)\%$ lcl for λ can be achieved using the α quantile of the distribution of Z. Denoting this quantile by c_1, we conclude that $P(c_1 < Z) = 1 - \alpha \Rightarrow P(c_1/Y < \lambda) = 1 - \alpha$. Hence, the quantity c_1/Y is a $100(1-\alpha)\%$ exact lcl for λ. Besides, since

$$P\left(\frac{c_1}{Y} < \lambda\right) = P\left(\exp\left(\frac{-Uc_1}{Y}\right) > e^{-\lambda U}\right) = P\left(\frac{0.0027}{\exp(-Uc_1/Y)} < C_{\text{pcu}}\right),$$

it follows that a $100(1-\alpha)\%$ exact lcl for the index C_{pcu} is given by

$$\frac{0.0027}{\exp(-Uc_1/Y)}. \tag{4.3}$$

Similarly, denoting the $1-\alpha$ quantile of Z by c_2 it can be shown that a $100(1-\alpha)\%$ lcl for C_{pcl} is given by

$$\frac{0.0027}{1 - \exp(-Lc_2/Y)}. \tag{4.4}$$

Note that the values of c_1 and c_2 that are involved in the assessment of (4.3) and (4.4) can be obtained from any statistical package. Finally, if the sample size is sufficiently large, one may obtain approximate lower confidence limits for C_{pcl} and C_{pcu} based on the large sample properties of Y.

5 DISCUSSION

In this paper, a new PCI is presented that overcomes many of the deficiencies of the standard PCIs. Some of the advantages of this index are that it can be used for processes with unilateral or bilateral tolerances, for continuous (normal or non-normal) or discrete processes, and can take into account the minimum acceptable process yield (if such a quantity has been set). Properties and estimation problems connected to this new index have been studied for normal and exponential processes. Because of its appealing features, examining its potential use in other types of processes often arising in connection with applications would be of practical value. Moreover, as already pointed out, one of the advantages of the new index is that its definition allows its use for discrete processes as well. It would therefore be interesting to examine the implementation of C_{pc} in the case of other types of processes, continuous or discrete (*e.g.* chi-square, Poisson, negative binomial).

References

Borges, W. and Ho, L. L. (2000). On the sampling distribution of Clement's capability index, *Communications in Statistics – Simulation and Computation*, **29**(1), 121–138.

Boyles, R. A. (1991). The Taguchi capability index, *Journal of Quality Technology*, **23**(1), 17–26.

Chou, Y.-M. and Owen, D. B. (1984). One-sided confidence regions on the upper and lower tail areas of the normal distribution, *Journal of Quality Technology*, **16**(3), 150–158.

Chou, Y. M., Owen, D. B. and Borrego, A. S. A. (1990). Lower confidence limits on process capability indices, *Journal of Quality Technology*, **22**(3), 223–229.

Clements, J. A. (1989). Process capability calculations for non-normal distributions, *Quality Progress*, **22**(9), 95–100.

Gunter, B. (1989). The use and abuse of C_{pk}: Part 2, *Quality Progress*, **22**(3), 108–109.

Kane, V. E. (1986). Process capability indices, *Journal of Quality Technology*, **18**(1), 41–52 (Corrigenda, **18**(4), 265).

Kotz, S. and Johnson, N. L. (1993). *Process capability indices*, Chapman and Hall.

Kotz, S. and Johnson, N. L. (2002). Process capability indices – a review, 1992–2000, *Journal of Quality Technology*, **34**(1), 2–19.

Kotz, S. and Lovelace, C. R. (1998). *Process capability indices in theory and practice*, Arnold.

Kushler, R. H. and Hurley, P. (1992). Confidence bounds for capability indices, *Journal of Quality Technology*, **24**(4), 188–195.

Li, H., Owen, D. B. and Borrego, A. S. A. (1990). Lower confidence limits on process capability indices based on the range, *Communications in Statistics – Simulation and Computation*, **19**(1), 1–24.

Owen, D. B. and Hua, T. A. (1977). Tables of confidence limits on the tail area of the normal distribution, *Communication in Statistics – Simulation and Computation*, **B6**(3), 285–311.

Pearn, W. L. and Kotz, S. (1994–1995). Application of Clements' method for calculating second- and third-generation process capability indices for non-normal Pearsonian populations, *Quality Engineering*, **7**(1), 139–145.

Perakis, M. and Xekalaki, E. (1999). Process capability indices for the exponential, the chi-square and the lognormal distribution, *Technical report no 90*. Department of Statistics, Athens University of Economics and Business, August 1999 (Submitted).

Perakis, M. and Xekalaki, E. (2001). On an improvement in confidence limits for proportion of conformance, *Technical report no 128*. Department of Statistics, Athens University of Economics and Business, March 2001.

Wang, C. M. and Lam, C. T. (1996). Confidence limits for proportion of conformance, *Journal of Quality Technology*, **28**(4), 439–445.

Yeh, A. B. and Bhattacharya, S. (1998). A robust process capability index, *Communications in Statistics – Simulation and Computation*, **27**(2), 565–589.