INT. J. PROD. RES., 1993, VOL. 31, NO. 7, 1621–1635

Process capability analysis—a robustness study

J. R. ENGLISH† and G. D. TAYLOR

The robustness of two popular process capability ratios, C_p and C_{pk}, when the random process being observed departs from normality is analysed. The distributions of estimated process capability ratios are derived and used as a basis for validation of large-scale simulation studies in an examination of departures from normality. The simulation studies and analytical results provide a basis for recommended procedures. It is recommended that the impact of process distributions be considered before using popular process capability indexes, due to the lack of robustness when departures from normality are ecountered. As an extension of our findings, we consider the Taguchi loss function as a generalization to process capability analysis regardless of the underlying population distribution.

1. Introduction

Historically, manufacturing organizations have maintained a production orientation rather than a quality focus. During the previous decade, this orientation has changed dramatically. Pioneers such as Dr Edward Deming (Deming 1986) have convinced many firms to support a quality focus by convincing managers to 'adopt the new philosophy' of making products right the first time. In support of these dynamic events, technology advances have been made and are documented in many statistical publications. Quality-related work is also appearing in the production literature at a fast rate. Much of this literature simply addresses quality issues with a production bias. See for example Chengalur *et al.* (1992), Kelton *et al.* (1990), Koo and Case (1990), and Barad (1990). Other authors address quality management issues in production settings (see Son and Hsu 1991 and Lascelles and Dale 1990). Swamidass and Majerus (1991) apply known quality control methods in innovative ways to solve production problems. Some authors address the difficult trade-offs between production and quality issues. Goyal and Gunasekaran (1990) study the effect of investment in quality on the economics of production. Hsu and Tapiero (1990) develop an economic model for determining optimal quality and production control policies when both quality and production quantity issues are considered. Murthy and Djamaludin (1990) formulated a model which examines open- and closed-loop policies for carrying out maintenance thus affecting quality.

One idea that is receiving a great deal of attention in both quality and production settings is process capability analysis (PCA). As a result, process capability analysis has become widely adopted as the ultimate measure of performance to evaluate the ability of a process to satisfy customers (in the form of specifications). Within the statistical publications, a reasonable framework is evolving for an inferential approach to PCA.

Received October 1992.
† To whom correspondence should be addressed.
Department of Industrial Engineering, University of Arkansas, Fayetteville, AS 72701, USA.

0020-7543/93 $10.00 © 1993 Taylor & Francis Ltd.

Unfortunately, basic research in measuring robustness of such measures in realistic environments are missing in the literature. In this work, we attempt to bridge the gap between production and quality researchers by evaluating classical PCA in the realistic situation where departures from normality are encountered. PCA is the baseline from which process performance must be measured. In fact, in our work we broaden PCA to include the widely accepted Taguchi loss function in departures from normality (Taguchi *et al.* 1989).

PCA as presented in this work is described as the means of measuring the ability of a process to satisfy a customer's described 'fitness for use' (Juran and Gryna 1980). The common measures of performance used in PCA are based on the ability of the process to satisfy customer specification limits. It is the assumption of this work (and of many other applications in PCA) that customer specifications are descriptive of the customers' 'fitness for use'. Therefore, the measures of performance utilized are indices which relate the natural tolerance limits of a process to the specification limits. Two of the more popular indices are C_p and C_{pk}; these are shown as follows:

$$C_p = \frac{USL - LSL}{6\sigma} \tag{1}$$

$$C_{pk} = \min(C_{pu}, C_{pl}) \tag{2}$$

where

$$C_{pu} = \frac{USL - \mu}{3\sigma}$$

$$C_{pl} = \frac{\mu - LSL}{3\sigma}$$

USL = upper specification limit
LSL = lower specification limit
$USL \geqslant LSL$
σ = process standard deviation
μ = process mean

Estimates of these indices are made by the appropriate estimates of σ and μ. Specifically

$$\hat{\sigma}^2 = s^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1} \tag{3}$$

$$\hat{\mu} = \bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} \tag{4}$$

where

X_i = process observation
n = sample size

An additional form of C_{pk} has been recommended (Chan *et al.* 1988), and is presented as follows:

$$C_{pk} = C_p(1 - k) \tag{5}$$

where

$$k = \frac{2|T-\mu|}{USL-LSL}$$

$0 \leqslant k \leqslant 1$
T = target

In most applications, the target, T, is the centre of the specification limits. Usually, only point estimates of indices (1) and (2) are utilized to measure the ability of a process to meet specifications. That is, the sampled estimates of μ and σ are determined from a sample of n observations as shown in equations (3) and (4), and single-point estimates of C_p and C_{pk} are made. If sampled values of C_p and/or C_{pk} exceed a minimum threshold value (usually $\geqslant 1$), the process is deemed capable. It is important to note that for the alternate form of C_{pk} (equation (5)) to be valid, k must be greater than or equal to zero and less than or equal to 1. In estimating C_{pk} using the alternative form, the parameter k is a random variable; therefore, when a process is not centred at the middle of the specification limits, the probability of observing a sample average greater than (or less than) the $USL(LSL)$ is certainly not zero, and depending on the particular situation, could be significantly greater than zero. Therefore, the alternative form for C_{pk} should be used with caution, realizing that the estimate of random variable k may exceed the limitations necessary for equivalence to C_{pk} by definition (equation (2)). In this work, the alternative form of C_{pk} (equation (5)) is not used due to the limitations as previously described.

Historically, PCA has been based on point estimates, and no inferential statistics have been applied. Fortunately, recent advances in PCA have allowed analysis to utilize hypothesis testing and confidence intervals on C_p and C_{pk} as the basis for establishing process capability (Boyles 1991, Chou *et al.* 1990 and Kane, 1986).

It is assumed throughout the existing literature that the process measurments are independent and identically distributed as normal variables [IIDN(μ, σ^2)]. It is the objective of this paper to evaluate the effectiveness of these estimates of C_p and C_{pk} to measure process capability given that the assumption of normality is violated. In particular, the analytical distributions for C_p and C_{pu} are derived under the assumptions of independence and normality, the sample distributions of C_p and C_{pk} under the assumptions of normality with departures from normality collected from large- scale simulation results (validation based upon the analytical distributions), and conclusions are made regarding the findings. The derivations of the distributions of the PCA ratios are an alternative to the derivations provided by other researchers (Boyles 1991, Chan *et al.* 1988, Chou *et al.* 1990, Kane 1986). More importantly, we provide an alternative form which accommodates our validation needs by deriving the distribution of the estimated PCA ratio directly. It is believed that quality control researchers will value these alternative derivations. The details of derivations are found in the appendixes.

2. Derivation of the probability density functions for $\hat{C}_p$ and $\hat{C}_{pk}$

The probability density function for the estimate of C_p is given as follows assuming the X_is are normally distributed with mean μ and variance σ^2:

$$f_{\hat{C}_p}(\hat{C}_p) = \frac{(k/\hat{C}_p^2)^{(n-1)/2} \exp(-k/2\hat{C}_p)}{\hat{C}_p^2 (2)^{(n-1)/2-1} \Gamma((n-1)/2)} \tag{6}$$

where $0 < \hat{C}_p < \infty$

$$k = C_p^2(n-1), \qquad \text{where } C_p \text{ is the population value.}$$

The proof of this is given in appendix A. The advantage of the pdf described in equation (6) over those offered in Kane (1986) or Chou *et al.* (1990) is that of the tolerance limit ($USL-LSL$) and a specific process [$N(\mu, \sigma^2)$] is used. For a particular process, unique specifications, mean, and standard deviation exist. This result is of particular importance in the evaluation of the robustness of process capability analysis since it provides a direct validation point when normality is assumed.

At this point, it would be beneficial to derive the pdf for the estimate of C_{pk} when the underlying process is composed of data which are independent and identically distributed as normal random variables. But, as cited in Chou *et al.* (1990), the relationship cannot be established explicitly.

The probability that the estimate of C_{pu} is greater than some value c may be determined through the integral relationship shown in expression (7) assuming the X_is are normally distributed with mean μ and variance σ^2.

$$\Pr[\hat{C}_{pu} \geqslant c] = \Pr[V - T \geqslant 0] = \int_{-\infty}^{\infty} f_V(v)\left(\int_0^v f_T(t)\,dt\right)dv \qquad (7)$$

This probability is completely described in Appendix B. The resulting probability is easily determined using double numerical integration. This derivation represents analytical advances over existing published research results and provides a basis for validation of the necessary simulation to evaluate the robustness of process capability indexes as departures from normality are observed. The equivalent derivation for C_{pl} is made similarly.

4. Relaxation of normality in process capability analysis

In many processes, the assumption of normality is common in PCA and is often not valid. For the purposes of this paper, various distributional assumptions are considered. Specifically, process observations are assumed to follow either the normal, triangular, uniform, or truncated exponential distributions. In an effort to examine the various distributions on equal footing, the parameters of the distributions are designed so that the population values of C_p and C_{pk} are exactly one. In other words, the mean is centred between the specifications limits, and the standard deviation of the process is such that the USL and LSL are exactly three standard deviations from the mean. This aspect provides the element of control for our analysis.

In classical PCA, it is traditionally accepted that a process is capable of meeting specifications if the designated process capability ratio is greater than or equal to one. Typically, point estimates of PCA ratios are made without regard to the underlying distribution of the process in question. Based upon the traditional approach to PCA, we have elected to control our experimentation based upon commonly accepted industrial practice. Arguably, one could consider trends as recommended by Juran (1988). Specifically, many organizations attempt to achieve capability at a level of 1·33. Distributions could be utilized which fix the PCA ratios at 1·33, but similar results would be observed. Furthermore, such an analysis could be controlled to fix the non-conforming percentage. In other words, the process parameters could be fixed such that the percentage non-conforming is at a pre-established value. With an approach such as this, there exists an infinite set of possible parameter values since there are effectively two unknowns (the density function parameters) and one known (the percentage non-conforming). Additionally, the practical implications of the research findings would be less informative than directly simulating the actual use of process capability ratios under 'ideal' (C_p and C_{pk} equal to one) situations. For these reasons, we have chosen to fix the true PCA ratios which serves as the control element in our experimentation.

The probability density functions (pdfs) and specification limits considered are as follows:

$$USL = 3$$
$$LSL = -3$$
$$\mu = 0$$
$$\sigma = 1$$

The resulting pdfs are shown in Table 1.

The percentage of non-conforming products for each of these distributions are shown in Table 2.

5. Simulation methodology and results

The primary methodology used to examine the effects of deviations from normality is that of simulation using the SIMNET language. It is desirable, from an experimental design standpoint, to isolate the effects of two key variables in these simulation studies. First, the underlying distribution of the data collected from the process being analysed

Process	PDF	
Normal	$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(x)^2)$	$-\infty < X < \infty$
Triangular	$f_X(x) = \frac{x}{\sqrt{6}} + \frac{1}{\sqrt{6}}$	$-\sqrt{6} \leqslant x < 0$
	$\frac{-x}{\sqrt{6}} + \frac{1}{\sqrt{6}}$	$0 \leqslant x < \sqrt{6}$
	0	Else
Uniform	$f_X(x) = \frac{1}{2\sqrt{3}}$	$-\sqrt{3} \leqslant x \leqslant \sqrt{3}$
	0	Else
Truncated exponential	$f_X(x) = \exp(-(X+1))$	$-1 < X < \infty$

Table 1. PDFs for non-normal processes.

Distribution	Percentage non-conforming
Normal	0·27
Triangular	0·00
Uniform	0·00
Truncated exponential	1·83

Table 2. Actual percentage non-conforming.

may very well violate the assumption of normality. For this reason we consider three possible scenarios other than the normal case. Secondly, we will examine the effect of increasing the sample size on process capability ratios under the various distribution assumptions. Each simulation run specifies a unique combination of sample size and assumed distribution of data.

As stated in the previous section, normal, triangular, uniform and truncated exponential distributions are considered. These distributions consider a relatively wide range of potential process capabilities. The normal distribution is the 'usual' distribution associated with PCA and is included for obvious reasons. The triangular distribution helps to demonstrate the possible process capability improvements that may be obtained when the tails of the normal distribution can be eliminated. The uniform distribution also demonstrates the possibilities for improvement with tight process control to eliminate extreme deviation from the process mean, yet also represents a fairly extreme departure from normality. Finally, the exponential distribution offers a look at a process distribution that differs greatly from the normal distribution yet has extreme outlying data points. The simulation parameters associated with each distribution are those outlined in Table 1. The parameters are established such that C_p and C_{pk} should be exactly one as an additional experimental control consideration.

The sample size also has an effect on measures of process capability. To examine the effects of sample size, the simulation studies consider sample sizes of 5, 10, 20, 30, and 50, for each of the four distributions considered. The selection of sample sizes are based upon the quantities typically collected. For example, sample sizes of 5–10 are used for control purposes, while samples of size 20–30 are used for process potential studies, and

Distribution	Batch	0·50	0·75	1·00	1·25	1·50	1·75	2·00	2·25	2·50	2·75
Normal	5	0·996	0·873	0·600	0·373	0·223	0·142	0·089	0·061	0·041	0·030
Triangular	5	1·000	0·881	0·561	0·320	0·193	0·122	0·079	0·054	0·039	0·027
Uniform	5	1·000	0·924	0·522	0·267	0·152	0·093	0·058	0·039	0·025	0·018
Exponential	5	0·969	0·846	0·683	0·527	0·404	0·309	0·240	0·186	0·150	0·118
Normal	10	1·000	0·933	0·568	0·243	0·091	0·036	0·013	0·005	0·002	0·001
Triangular	10	1·000	0·959	0·529	0·189	0·064	0·024	0·011	0·005	0·003	0·001
Uniform	10	1·000	0·989	0·501	0·127	0·035	0·012	0·005	0·002	0·001	0·000
Exponential	10	0·985	0·865	0·644	0·423	0·259	0·156	0·097	0·060	0·039	0·025
Normal	20	1·000	0·979	0·548	0·129	0·017	0·002	0·001	0·000	0·000	0·000
Triangular	20	1·000	0·994	0·523	0·087	0·011	0·001	0·000	0·000	0·000	0·000
Uniform	20	1·000	1·000	0·508	0·041	0·002	0·000	0·000	0·000	0·000	0·000
Exponential	20	0·996	0·900	0·617	0·315	0·135	0·055	0·022	0·008	0·003	0·002
Normal	30	1·000	0·994	0·534	0·070	0·003	0·000	0·000	0·000	0·000	0·000
Triangular	30	1·000	0·999	0·514	0·039	0·002	0·000	0·000	0·000	0·000	0·000
Uniform	30	1·000	1·000	0·498	0·013	0·000	0·000	0·000	0·000	0·000	0·000
Exponential	30	0·999	0·925	0·600	0·249	0·076	0·021	0·005	0·001	0·000	0·000
Normal	50	1·000	0·999	0·537	0·022	0·000	0·000	0·000	0·000	0·000	0·000
Triangular	50	1·000	1·000	0·507	0·012	0·000	0·000	0·000	0·000	0·000	0·000
Uniform	50	1·000	1·000	0·497	0·002	0·000	0·000	0·000	0·000	0·000	
Exponential	50	1·000	0·955	0·577	0·165	0·026	0·004	0·001	0·000	0·000	0·000

Table 3. Proportion of $\hat{C}_p$ greater than X.

Distribution	Batch	0·50	0·75	1·00	1·25	1·50	1·75	2·00	2·25	2·50	2·75
Normal	5	0·976	0·752	0·465	0·269	0·159	0·094	0·059	0·040	0·027	0·020
Triangular	5	0·994	0·749	0·408	0·223	0·131	0·081	0·053	0·037	0·024	0·017
Uniform	5	1·000	0·782	0·358	0·163	0·082	0·047	0·029	0·018	0·012	0·009
Exponential	5	0·919	0·792	0·630	0·476	0·346	0·251	0·183	0·138	0·104	0·078
Normal	10	0·999	0·839	0·430	0·162	0·056	0·020	0·008	0·003	0·001	0·001
Triangular	10	1·000	0·869	0·365	0·116	0·038	0·014	0·006	0·003	0·001	0·000
Uniform	10	1·000	0·929	0·304	0·060	0·014	0·004	0·002	0·000	0·000	0·000
Exponential	10	0·954	0·810	0·589	0·369	0·202	0·106	0·058	0·033	0·019	0·010
Normal	20	1·000	0·934	0·412	0·077	0·009	0·001	0·000	0·000	0·000	0·000
Triangular	20	1·000	0·963	0·360	0·047	0·005	0·001	0·000	0·000	0·000	0·000
Uniform	20	1·000	0·991	0·298	0·017	0·001	0·000	0·000	0·000	0·000	0·000
Exponential	20	0·985	0·849	0·563	0·260	0·089	0·026	0·008	0·003	0·001	0·001
Normal	30	1·000	0·973	0·391	0·038	0·002	0·000	0·000	0·000	0·000	0·000
Triangular	30	1·000	0·989	0·351	0·021	0·001	0·000	0·000	0·000	0·000	0·000
Uniform	30	1·000	0·999	0·298	0·004	0·000	0·000	0·000	0·000	0·000	0·000
Exponential	30	0·995	0·880	0·545	0·190	0·041	0·007	0·001	0·000	0·000	0·000
Normal	50	1·000	0·994	0·392	0·011	0·000	0·000	0·000	0·000	0·000	0·000
Triangular	50	1·000	0·999	0·342	0·005	0·000	0·000	0·000	0·000	0·000	0·000
Uniform	50	1·000	1·000	0·291	0·001	0·000	0·000	0·000	0·000	0·000	0·000
Exponential	50	1·000	0·919	0·524	0·113	0·010	0·001	0·000	0·000	0·000	0·000

Table 4. Proportion of $\hat{C}_{pk}$ greater than X.

samples of size 40–50 are typical of process capability studies. In all, 20 scenarios are simulated to examine each combination of four distributions and five sample sizes.

The simulations are run on a large scale to obtain greater confidence in the results. For each of the 20 scenarios, 10 000 samples are considered. The procedure followed is to generate an appropriate number of simulation entities as specified by the sample size and distribution parameters, to calculate the sample mean and standard deviation for the sample, and use this information to calculate the estimates of C_p and C_{pk}. The final results and conclusions are based on aggregate information from all 10 000 samples. For the normal case, the simulation results are carefully validated using the probability density functions for the estimates of C_p and C_{pk} as derived previously.

With samples of 10 000, the variation in the resulting probability estimates is minimal. For example, a 95% confidence interval on the average probability that the estimate of C_{pk} with sample size of 5 is greater than 1 spanned less than 0·008 (for all possible distributions) across the sample average of 10 observations of 10 000 samples each.

Results associated with the estimates of C_p and C_{pk} appear in Tables 3 and 4, respectively. The tables show, for each combination of distribution and sample size, the percentage of observed values that are greater than a specified value. For example, for the triangular distribution at a sample size of 20, the percentage of estimated C_{pk} values greater than 1·25 is 4·7%. Additional summary information will be presented subsequently.

The histograms of the estimates of C_p and C_{pk} are shown for sample sizes of 5 and 50 in Figs 1–4.

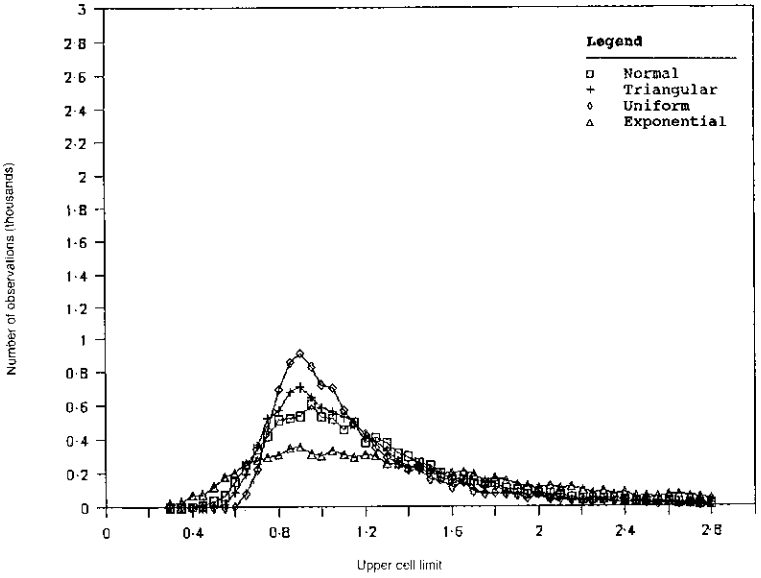

Figure 1. Histograms for $\hat{C}_p$ at $n=5$.

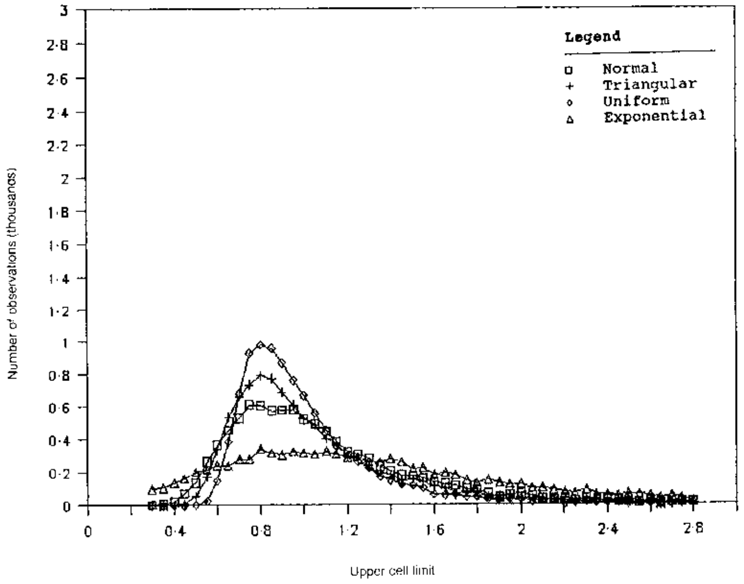

Figure 2. Histograms for $\hat{C}_{pk}$ at $n=5$.

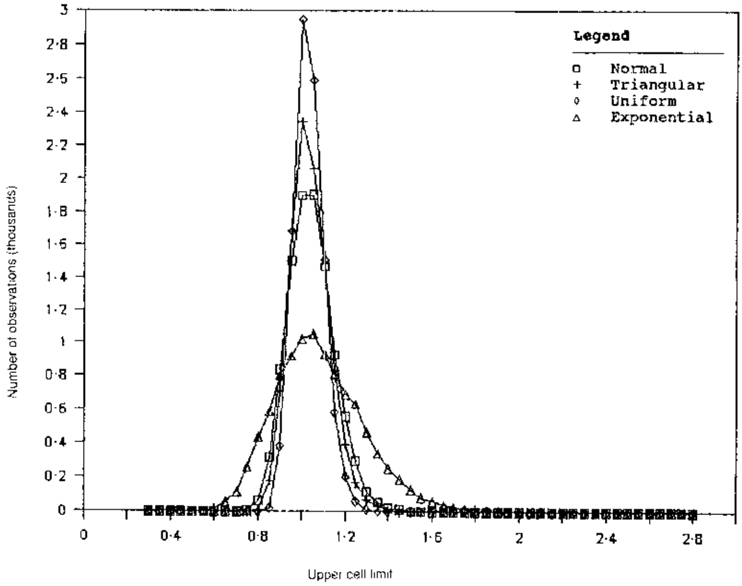

Figure 3. Histograms for $\hat{C}_p$ at $n=50$.

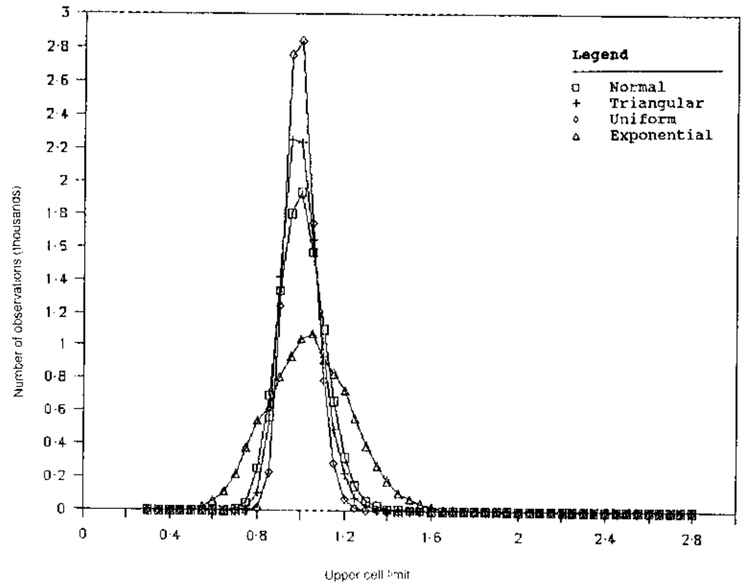

Figure 4. Histograms for $\hat{C}_{pk}$ at $n=50$.

6. Loss function

As previously stated, PCA as presented in this work is described as the means of measuring the ability of a process to satisfy a customer's described 'fitness for use' (Juran and Gryna 1980). The common measures of performance used in PCA are based on the ability of the process to satisfy customer specification limits which are descriptive of the customers' 'fitness for use'. Therefore, the measures of performance utilized are indices which relate the natural tolerance limits of a process to the specification limits, namely C_p and C_{pk}.

In what follows, the distributions considered are the same as those considered in previous sections. Specifically, the triangular, uniform and exponential distributions are considered as non-normal populations. In each case, the parameters of the distributions are established so that the true values of C_p and C_{pk} are exactly one (or exactly capable).

PCA indicates the ability of a process to satisfy customer specification limits, but it is perhaps more interesting to examine the costs associated with process variation. One means of examining these costs is the Taguchi quadratic loss function (Taguchi *et al.* 1989).

The basis of this function is that loss is incurred when the quality characteristics of a product deviate from target or nominal values, even if the deviation is very small. The quadratic loss function may be stated as follows:

$$L = k(Y - T)^2 \tag{8}$$

where

L = loss function
k = a proportionality constant
Y = observed value of quality characteristic
T = target value of quality characteristic

The loss imparted to society may be estimated by this loss function. Current work with the Taguchi loss function has utilized the function as the basis for building expected

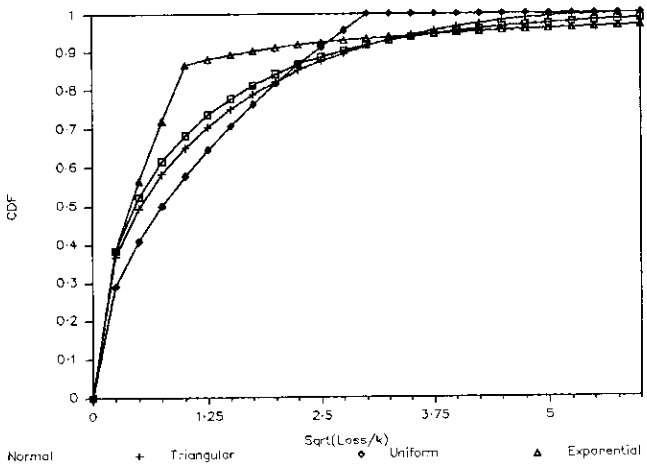

Figure 5. Distribution of loss.

cost models. In fact, the loss function represents a transformation from one random variable to another and should be treated appropriately. In Fig. 5, the resulting distributions of loss are shown for the four underlying parent populations assumed in our analysis. The derivations of the distributions are based upon elementary probability transformations found in any beginning textbook in probability theory.

It is interesting to note that the exponential distribution (which represents an extreme departure from normality) performs very well in terms that the probability that loss is less than small values k, but poorly as k increases.

7. Conclusions

The simulation results lead to several observations and conclusions. Specifically, the application of C_p and C_{pk} for process capability analysis when departures from normality are experienced must be made carefully. In what follows, a type I error (false alarm) is defined as the incorrect conclusion that the true process capability ratio (C_p or C_{pk}) is less than 1 when the true ratio is 1, or capable of meeting specifications.

In viewing the data for $\hat{C}_p$, several observations are made. Table 3 indicates that the number of type I errors when the underlying population is symmetric but not normal (uniform or triangular) and become less pronounced as the sample size increases. This conclusion is based upon the fact that the percentage of observations being greater than one decreases under small departures from normality. As the underlying process becomes less symmetric (exponential) the type I error rate decreases (Table 3). On the surface, this observation is alarming. This result stems from the skewness of the exponential distribution as compared to the normal distribution (Table 2). Since the exponential distribution is skewed further right than the normal distribution, more observations are seen to the extreme right resulting in higher estimates of the process standard deviation when compared to the normal distribution. Clearly, the disadvantage of the exponential case is observed when the process is truly not capable (i.e. $C_p < 1$ or $C_{pk} < 1$), since the type II error rate will obviously be higher for the non-normal case (based on the results presented in Table 3 and considering the impact of C_p being less than 1).

The triangular and uniform distributions have no tails beyond the three sigma range (Table 2). Therefore, as one would expect, fewer observations are seen to the extreme right resulting in lower estimates of the process standard deviation when compared with the hypothesized normal population. Therefore, an increasing number of estimated C_p s will fall to less than 1 when compared with a normally distributed process.

As demonstrated in Figs 1 and 3, the distributions of the estimated C_p for the non-normal processes approach that of the normal process as the sample size is increased. Clearly this is observed for the symmetric and non-normal cases (triangular and uniform) and to a lesser extent for the exponential case. Additionally, the relative increase in type I errors due to departures from normality decreases as the sample sizes increase as seen in the number of observations falling above one (contrast various sample sizes in Table 3). As a rule of thumb, if the sample size, n, is between 30 and 50, it is reasonable that the use of C_p for process capability is fairly robust to departures from normality.

The C_{pk} ratio is more sensitive to departures from normality. As shown in Table 4, a more dramatic increase in type I errors is realized. This observation stems from the dependence of C_{pk} on the estimate of the process mean when the process is skewed right (exponential). When comparing the distribution of the estimates of C_{pk} with C_p, Fig. 1

versus Fig. 2, or Fig. 3 versus Fig. 4, the resulting distribution for C_{pk} is skewed further right when the sample is small but the mode moves to the right as the sample size is increased. This observation leads to the conclusion that the increase of the sample size adds weight to the concern of small sample sizes in process capability analysis. Specifically, samples of size 30 should be considered minimal for reasonable inferential conclusions. These observations are indicative of the increase of observations less than one (type I errors) and predictive of the excessive increase in type II errors. Even as the sample sizes increase (Figs 2 and 4), the severe departures from the hypothesized normal process (the exponential case) result in relatively poor performance of C_{pk} for process capability analysis.

It is recommended that the analyst be sensitive to the process behaviour before using popular process capability indexes due to the lack of robustness to departures from normality. As a rule of thumb, C_p may be used with reasonable assurance when the sample size is large (30–50), but the use of C_{pk} should be avoided when severe departures from normality are observed. It is also important to note that the processes considered here are stationary (or in a state of statistical control). If doubt of such a conclusion exists, PCA is invalid.

Based upon what is presented above, several interesting observations may be made. In particular, one readily observes that for the exponential case, smaller cumulative losses are encountered more frequently than observed for the normal, uniform, and triangular cases. This provides an interesting conclusion since the exponential distribution represents the greatest departure from normality when compared to the other distributions considered. As presented in our previous work, classical process capability analysis is misleading (i.e. C_p, C_{pk}, etc.). Specifically, the exponentially distributed process observations result in distributions for $\hat{C}_p$ and $\hat{C}_{pk}$ which will often misdirect the conclusion of the ability of the process to satisfy specifications (deflated values of indexes). Therefore, behaviour of the underlying process provides the direction of the analysis. In particular, a more optimal change in process mean may provide less loss than would have been realized with a normal population.

The placement of the target value for the quality parameter is often a design variable and may be easily adjusted. Therefore, there is a strong incentive for manufacturing process design engineers to modify the idea of the utilization of specification and utilize more optimal target values supported by known process behaviour. As has been identified by the pioneering works of Taguchi *et al.* (1989), when 'nominal is best' is assumed, the expected value of loss is minimized regardless of the distribution, when the target value is set equal to the expected value of the underlying process. To our knowledge, there have been no advances which consider minimization of cumulative loss. It is the opinion of these researchers that this design issue should be the object for future research and should be fully exploited.

Acknowledgments

The authors wish to express their gratitude to Mr Tom Sargent of the Industrial Engineering Department at the University of Arkansas for his assistance in validating our results.

Appendix A—Distribution of the estimate of C_p

In what follows, a slight modification of the proofs in Kane 1986 and Chou *et al.* 1990 is presented to aid in the robustness analysis of PCA using C_p and C_{pk}. As previously mentioned, practically all process capability analysis assumes that the data,

X_i, resulting from a manufacturing process is independent and identically distributed as normal random variables with mean μ and variance σ^2. Specifically

$$X_i \sim \mathrm{IIDN}(\mu, \sigma^2)$$

It is well known that,

$$\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{\sigma^2} \sim X^2_{n-1} \tag{9}$$

where X^2_{n-1} is distributed as a chi-squared random variable with $(n-1)$ degrees of freedom. The random variable may be rewritten as

$$\frac{(n-1)s^2}{\sigma^2} \sim X^2_{n-1} \tag{10}$$

where

$$f_{X^2, n-1}(x) = \frac{1}{2^{(n-1)/2}\Gamma((n-1)/2)} x^{((n-1)/2)-1} \exp(-x/2) \qquad 0 < X < \infty \tag{11}$$

From process observations, estimates of C_p can be made. Defining this new random variable

$$\hat{C}_p = \frac{USL - LSL}{6s} \tag{12}$$

from (12), it is easily seen that

$$\hat{C}_p = \frac{(USL - LSL)\sigma}{6s(n-1)^{1/2}} \frac{(n-1)^{1/2}}{\sigma}$$

Rearranging

$$g(x) = \hat{C}_p = \frac{(USL - LSL)(n-1)^{1/2}}{6\sigma\sqrt{x}} \tag{13}$$

From Parzen 1960, it is known that

$$f_{\hat{C}_p}(\hat{C}_p) = f_{X^2, n-1}(g^{-1}(\hat{C}_p)) \left| \frac{dg^{-1}}{dC_p}(\hat{C}_p) \right| \tag{14}$$

where

$$g^{-1}(\hat{C}_p) = \frac{(USL - LSL)^2(n-1)^2}{36\sigma^2 C_p^2} \tag{15}$$

From equations (11), (14), and (15)

$$f_{\hat{C}_p}(\hat{C}_p) = \frac{(k/\hat{C}_p^2)^{(n-1)/2} \exp(-(k/2\hat{C}_p))}{\hat{C}_p^2(2)^{(n-1)/2-1}\Gamma((n-1)/2} \tag{16}$$

where $0 < \hat{C}_p < \infty$ and $k = C_p^2(n-1)$, where, C_p is the population value.

Appendix B—Distribution for estimate of C_{pk}

Since it is difficult to explicitly derive the probability density function for estimate of C_{pk}, alternative derivations are considered. Specifically, probability calculations may be determined for C_{pu} or C_{pl} alone. These developments are described as follows.

Consider following probability statement (one minus the distribution function)

$$\Pr[\hat{C}_{pu} \geqslant c] = \theta \tag{17}$$

where c = constant and $0 < \theta < 1$.

The random variable C_{pu} (described in equation (2)) may be rewritten as follows:

$$\hat{C}_{pu} = \hat{C}_p \frac{2(USL - \bar{X})}{USL - LSL} \tag{18}$$

Therefore

$$\Pr[\hat{C}_{pu} \geqslant c] = \Pr\left[2\frac{(USL - \bar{X})}{USL - LSL} - c/\hat{C}_p \geqslant 0\right] = \theta \tag{19}$$

Define the random variables V and T as follows:

$$V = \frac{2(USL - \bar{X})}{USL - LSL}$$

and

$$T = c/\hat{C}_p \tag{20}$$

V and T are dependent only upon the mean variance, respectively; therefore if normality is assumed, they are independent random variables (Lin and Mudholkar 1980 and Nelson 1981). Based on this observation, the probability described in equation (19) may be easily determined by considering the interference region in which V exceeds T. Such calculations are common in many areas of applied probability theory. Specifically, one sees such applications in reliability (Kapur and Lamberson 1977).

The resulting probability (described in equation (17)) may be determined through the following integral relationship:

$$\Pr[\hat{C}_{pu} \geqslant c] = \Pr[V - T \geqslant 0] = \int_{-\infty}^{\infty} f_V(v)\left(\int_0^v f_T(t)\,dt\right)dv \tag{21}$$

The necessary pdfs for V and T are determined as follows. V is defined as

$$V = \frac{2(USL - \bar{X})}{USL - LSL} \tag{22}$$

If the observations making up $\bar{X}$ are assumed to be independent and normally distributed with mean μ and variance σ^2, the resulting distribution may be determined using equation (14) from Parzen (1960). Specifically

$$g^{-1}(v) = USL - \frac{V(USL - LSL)}{2} \tag{23}$$

therefore

$$f_V(v) = \frac{USL - LSL}{2} \frac{1}{(\sigma/n)\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{USL - v(USL - LSL)/2 - \mu}{\sigma/n}\right)^2\right) \tag{24}$$

where USL, LSL, σ, μ, and n are as previously defined, and $-\infty < v < \infty$.

T is defined as follows:

$$T = c/\hat{C}_p$$

Recognizing that the pdf for the estimate of C_p is previously determined (equation (16)), the pdf may also be easily determined using simple transformation of random variables (equation (14), Parzen 1960). Specifically

$$g^{-1}(t) = c/t$$

therefore

$$f_T(t) = \frac{1}{t2^{(n-1)/2}\Gamma((n-1)/2} \left(\frac{kt^2}{c^2}\right)^{((n-1)/2)-1} \exp\left(\frac{-kt^2}{2c^2}\right) \qquad (25)$$

where n, k, c, are as previously defined and $0 < t < \infty$.

References

BARAD, M., 1990, A break-even quality level approach to location of inspection stations in a multi-stage production process. *International Journal of Production Research*, **28**, 29–45.

BOYLES, R. A., 1991, The Taguchi capability index. *Journal of Quality Technology*, **23**, 17–26.

CHAN, L. K., CHENG, S. W., and SPRING, F. A., 1988, A new measure of process capability: C_{pm}. *Journal of Quality Technology*, **20**, 162–175.

CHENGALUR, I. N., BALLOU, D. P., and PAZER, H. L., 1992, Dynamically determined optimal inspection strategies for serial production processes. *International Journal of Production Research*, **30**, 169–187.

CHOU, Y., OWEN, D. B., and BORREGO, A., S. A., 1990, Lower confidence limits on process capability indices. *Journal of Quality Technology*, **22**, 223–229.

DEMING, W. E., 1986, *Out of the Crisis* (Cambridge, Mass: MIT CAES).

GOYAL, S. K., and GUNASEKARAN, A., 1990, Effect of dynamic process quality control on the economics of production. *International Journal of Operations and Production Management*, **10**, 69–77.

HSU, L. F., and TAPIERO, C. S., 1990, An economic model for determining the optimal quality and process control policy in a queue-like production system. *International Journal of Production Research*, **28**, 1447–1457.

JURAN, J. M., 1988, *Juran on Planning for Quality* (New York: Free Press).

JURAN, J. M., and GRYNA, F. M., 1980, *Quality Planning and Analysis*, second edition (New York: Mcgraw-Hill).

KANE, V. E., 1986, Process capability indices. *Journal of Quality Technology*, **18**, 41–52.

KAPUR, K. C., and LAMBERSON, L. R., 1977, *Reliability in Engineering Design* (New York: John Wiley & Sons).

KELTON, W. D., HANCOCK, W. M., and BISCHAK, D. P., 1990, Adjustment rules based on quality control charts. *International Journal of Production Research*, **28**, 385–400.

KOO, T. Y., and CASE, K. E., 1990, Economic design of X-bar control charts for use in monitoring continuous flow processes. *International Journal of Production Research*, **28**, 2001–2011.

LASCELLES, D. M., and DALE, B. G., 1990, The key issues of a quality improvement process. *International Journal of Production Research*, **28**, 131–143.

LIN, C.-C., and MUDHOLKAR, G. S., 1980, A simple test for normality against asymmetric alternatives. *Biometrika*, **67**, 455–461.

MURTHY, D. N. P., and DJAMALUDIN, I., 1990, Quality control in a single state production system: open and closed loop policies. *International Journal of Production Research*, **28**, 2219–2242.

NELSON, L. S., 1981, A simple test for normality. *Journal of Quality Technology*, **13**, 76–77.

PARZEN, E., 1960, *Modern Probability and Its Applications* (New York: John Wiley & Sons).

SON, Y. K., and HSU, L. F., 1991, A method of measuring quality cost. *International Journal of Production Research*, **29**, 1785–1794.

SWAMIDASS, P. M., and MAJERUS, C., 1991, Statistical control of manufacturing cycle time and project time: lessons from statistical process control. *International Journal of Production Research*, **29**, 551–563.

TAGUCHI, G., ELSAYED, E. A., and HSIANG, T., 1989, *Quality Engineering in Production Systems* (New York: McGraw-Hill).