AN APPLICATION OF NON-NORMAL PROCESS CAPABILITY INDICES

k. s. chen and w. l. pearn*
Department of Industrial Engineering & Management, National Chin-Yi Institute of Technology, Taichung, Taiwan ROC
Department of Industrial Engineering & Management, National Chiao Tung University, Hsinchu, Taiwan 30050 ROC

SUMMARY
Numerous process capability indices, including $C_p$, $C_{pk}$, $C_{pm}$, and $C_{pmk}$, have been proposed to provide measures of process potential and performance. In this paper, we consider some generalizations of these four basic indices to cover non-normal distributions. The proposed generalizations are compared with the four basic indices. The results show that the proposed generalizations are more accurate than those basic indices and other generalizations in measuring process capability. We also consider an estimation method based on sample percentiles to calculate the proposed generalizations, and give an example to illustrate how we apply the proposed generalizations to actual data collected from the factory. © 1997 John Wiley & Sons, Ltd.

1. INTRODUCTION
Process capability indices (PCIs) have been widely used in the manufacturing industry, to provide a numerical measure on whether a process is capable of producing items meeting the quality requirement preset in the factory. Numerous capability indices have been proposed to measure process potential and performance. Examples include the two most commonly used indices, $C_p$ and $C_{pk}$ discussed in Kane, and the two more-advanced indices $C_{pm}$ and $C_{pmk}$ developed by Chan et al. and Pearn et al. There are many other indices but they can be viewed as modifications of the above four basic indices (see Boyles, Pearn and Chen and Zwick).

Discussions and analyses of these indices on point estimation and construction of confidence intervals have been the focus of many statistician and quality researchers including Kane, Chan et al., Chou et al., Pearn et al., Kotz et al., Vannman, Pearn and Chen, and many others. Most of the investigations, however, depend heavily on the assumption of normal variability. If the underlying distributions are non-normal, then the capability calculations are highly unreliable since the conventional estimator of $\sigma^2$ is sensitive to departures from normality, and estimators of those indices are calculated using $S^2$ (see Chang et al., Gunter, and Somerville and Montgomery). Therefore, those basic indices are inappropriate for processes with non-normal distributions. In this paper, we consider some generalizations of those basic indices to cover non-normal distributions. Comparisons on accuracy of the capability measurement between the basic indices and the proposed generalizations are provided.

2. THE INDICES $C_p(u,v)$
Vannman constructed a superstructure for the four basic indices, $C_p$, $C_{pk}$, $C_{pm}$, and $C_{pmk}$. The superstructure has been referred to as $C_p(u,v)$, which can be defined as the following:

$$C_p(u,v) = \frac{d - u|m - m|}{3\sqrt{\sigma^2 + v(\mu - T)^2}},$$

where $\mu$ is the process mean, $\sigma$ is the process standard deviation, $d = (USL - LSL)/2$ is half of the length of the specification interval, $m = (USL + LSL)/2$ is the mid-point between the upper and the lower specification limits, $T$ is the target value, and $u, v \geq 0$. It is easy to verify that $C_p(0,0) = C_p$, $C_p(1,0) = C_{pk}$, $C_p(0,1) = C_{pm}$, and $C_p(1,1) = C_{pmk}$ which have been defined explicitly as:

$$C_p = \frac{USL - LSL}{6\sigma},$$
$$C_{pk} = \min \left\{ \frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma} \right\}.$$
\[ C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}. \]

\[ C_{pmk} = \min \left\{ \frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}} \right\}. \]

The index \( C_p \) only considers the process variability \( \sigma \) and thus provides no sensitivity on process departure at all. The index \( C_{pk} \) takes the process mean into consideration but it can fail to distinguish between on-target processes from off-target processes (Pearn et al.). The index \( C_{pm} \) takes the proximity of process mean from the target value into account, and is more sensitive to process departure than \( C_p \) and \( C_{pk} \). The index \( C_{pmk} \) adds an additional term \((\mu - T)^2\) in the definition, as a penalty to the process quality due to the departure of process mean from the target value. This additional penalty ensures that \( C_{pmk} \) will be more sensitive to departure than \( C_{pk} \) and \( C_{pm} \), and therefore is able to distinguish better between off-target and on-target processes. Clearly without the term \((\mu - T)^2\) in the denominator, the index \( C_{pmk} \) becomes \( C_{pk} \). The ranking of the four basic indices, in terms of sensitivity to departure of process mean from the target value, from the most sensitive one up to the least sensitive are (1) \( C_{pmk} \), (2) \( C_{pm} \), (3) \( C_{pk} \), and (4) \( C_p \).

Estimators of the indices \( C_p(u,v) \) may be obtained by replacing \( \mu \) by the sample mean \( \bar{X} = (\sum x_i)/n \), and \( \sigma^2 \) by the sample variance \( S^2 = (n-1)^{-1} \sum(x_i-\bar{X})^2 \) in definition (1). For normal distributions, those estimators based on \( \bar{X} \) and \( S^2 \), are quite stable and reliable. But, for non-normal distributions, those estimates become highly unstable since the distribution of the sample variance, \( S^2 \), is sensitive to departures from normality, and estimators of those indices are calculated using \( S^2 \), as pointed out by Chan et al. Gunter, and Somerville and Montgomery demonstrated the strong impact this has on the sampling distribution of \( C_{pk} \).

3. THE GENERALIZATIONS \( C_{Np}(u,v) \)

To accommodate cases where the underlying distributions may not be normal, we consider the following generalizations of \( C_p(u,v) \), which we refer to as \( C_{Np}(u,v) \). The generalizations \( C_{Np}(u,v) \) can be defined as (in superstructure form):

\[ C_{Np}(u,v) = \frac{\left( \int F_{99.865} - F_{0.135} \right)}{6} \cdot \sqrt{\frac{m-M}{3}}. \]

where \( F_a \) is the \( a \)th percentile, \( M \) is the median of the distribution, \( m = (USL+LSL)/2 \) is the mid-point between the upper and the lower specification limits, and \( u, v \geq 0 \). Thus, in developing the generalizations we have replaced the process mean, \( \mu \), by the process median, \( M \), and the process standard deviation \( \sigma \) by \((F_{99.865} - F_{0.135})/6\) in the definition of the basic indices \( C_p(u,v) \). The idea behind such replacements is to mimic the property of the normal distribution for which the tail probability outside the \( \pm 3\sigma \) limits from \( \mu \) is 0.27%, thus assuring that if the calculated value of \( C_{Np}(u,v) \) = 1 (assuming the process is well-centred, or on-target) the probability that process is outside the specification interval \((LSL, USL)\) will be negligibly small. It should be noted that the median \( M \) is a more robust measure of central tendency than the mean \( \mu \), particularly for skewed distributions with long-tails.

By setting \((u,v) = (0,0), (0,1), (1,0), \) and \((1,1)\), we obtain the following generalizations of the four basic indices for non-normal distributions, which we refer to as \( C_{Np}, C_{Npk}, C_{Npnm}, \) and \( C_{Npnmk} \):

\[ C_{Np} = \frac{USL - LSL}{F_{99.865} - F_{0.135}}. \]

\[ C_{Npk} = \min \left\{ \frac{USL-M}{F_{99.865} - F_{0.135}}, \frac{M-LSL}{F_{99.865} - F_{0.135}} \right\}. \]

\[ C_{Npm} = \frac{USL-LSL}{6 \sqrt{\left[ F_{99.865} - F_{0.135} \right]^2 + (M-T)^2}}. \]

\[ C_{Npnmk} = \min \left\{ \frac{USL-M}{3 \sqrt{\left[ F_{99.865} - F_{0.135} \right]^2 + (M-T)^2}}, \frac{M-LSL}{3 \sqrt{\left[ F_{99.865} - F_{0.135} \right]^2 + (M-T)^2}} \right\}. \]

The ranking of the four generalized indices (when applied to non-normal distributions) in terms of sensitivity to departure of process median from the target value, from the most sensitive one up to the least sensitive turns out to be the same. They are (1) \( C_{Npnmk} \), (2) \( C_{Npnm} \), (3) \( C_{Npk} \), and (4) \( C_{Np} \). In the special case where the underlying distribution is normal, then \( M = \mu \), and \( F_{99.865} - F_{0.135} = 6\sigma \). Clearly, the generalizations \( C_{Np}(u,v) \) reduce to the basic indices \( C_p(u,v) \), and so \( C_{Np} = C_p \), \( C_{Npk} = C_{pk} \), \( C_{Npm} = C_{pm} \), and \( C_{Npnmk} = C_{pmk} \).

Recently, Zwick and Schneider et al. considered two generalizations of \( C_p \) and \( C_{pk} \), which are similar to \( C_{Np} \) and \( C_{Npk} \) but with process mean \( \mu \) rather than process median \( M \) in the definitions. Extending their definitions to include the other two basic indices, \( C_{pm} \) and \( C_{pmk} \), a superstructure can be constructed in the following, which we refer to as \( C_{Np}(u,v) \):
Table I. Characteristics of processes A, B, and C

<table>
<thead>
<tr>
<th>Process</th>
<th>μ</th>
<th>M</th>
<th>σ</th>
<th>χ_{0.135}</th>
<th>σ_{99.865}</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10.00</td>
<td>9.37</td>
<td>2.45</td>
<td>7.03</td>
<td>22.63</td>
</tr>
<tr>
<td>B</td>
<td>17.80</td>
<td>17.70</td>
<td>2.45</td>
<td>14.83</td>
<td>30.43</td>
</tr>
<tr>
<td>C</td>
<td>25.60</td>
<td>24.97</td>
<td>2.45</td>
<td>22.63</td>
<td>38.23</td>
</tr>
</tbody>
</table>

\[
C_{Np}(u,v) = \frac{d-u[\mu-m]}{3\sqrt{\frac{F_{99.865}-F_{0.135}}{6} + v(\mu-T)^2}}.
\]  

(3)

4. COMPARISONS

To compare the proposed generalizations \( C_{Np}(u,v) \) with \( C_p(u,v) \), we consider an example of three processes A, B, and C depicted in Figure 1. All three processes are distributed as \( \chi^2 \) with three degrees of freedom (a skewed distribution). The characteristics are summarized in Table I (\( \sigma_A = \sigma_B = \sigma_C = (6)^{1/2} \)). Process B is on-target (\( \mu_B = T \)), but processes A and C are severely off-target (\( \mu_A = LSL \) and \( \mu_C = USL \)).

Table II is a comparison between \( C_p(u,v) \) and \( C_{Np}(u,v) \) on the three processes depicted in Figure 1. The \( C_p \), \( C_{pk} \), \( C_{pm} \) and \( C_{pmk} \) values given to processes A and C are the same. Both processes are severely off-target. But, the proportion of non-conforming is 61% for process A, which is significantly greater than that for process C (which is 39%). Obviously, the basic indices \( C_p(u,v) \) inconsistently measure process capabilities of processes A and C in this case. On the other hand, the proposed generalizations \( C_{Np}(u,v) \) clearly differentiate processes A and C by giving smaller values to A and larger values to C (excluding \( C_{Np} \) which never considers process median and hence provides no sensitivity to process departure at all). For processes distributed as Weibull (often used in practice as a model for skewed data), the result is the same. In fact, for Weibull (\( \alpha,\beta \)) with \( \alpha = 3 \) and \( \beta = 1.1 \), the percentage comparisons, 61% versus 39%, displayed in Figure 1 will be replaced by 62% versus 38%.

Table III is a comparison between the proposed generalizations \( C_{Np}(u,v) \) and other generalizations \( C_{Np}(u,v) \) on the three processes depicted in Figure 1. The index values \( C_{Np}(u,v) \) given to processes A and C are the same (1.00, 0.00, 0.32, 0.00) for both A and C), which inconsistently measure process capability in this case.

5. CALCULATIONS OF \( C_{Np}(u,v) \)

Pearn and Chen proposed an estimator for calculating the indices \( C_p(u,v) \) assuming the underlying distributions are Pearsonian types. The estimators essentially apply Clements’ method by replacing the 6σ interval length by \( U_p-L_p \), which can be calculated based on available sample data collected from a stable process utilizing estimates of the mean, standard deviation, skewness and kurtosis. Under the assumption that these four parameters determine the type of the Pearson distribution curve, the \( F_u \) percentiles of the Pearson curves as a function of skewness and kurtosis can be calculated utilizing the tables provided by Gruska et al. Those estimators can be written as (see Pearn and Chen):

\[
C_{Np}(u,v) = \frac{d-u[M-m]}{3\sqrt{\frac{U_p-L_p}{6} + v(M-T)^2}}.
\]  

(4)

where \( U_p \) estimates the 99.865 percentile \( F_{0.99865} \), \( L_p \) estimates the 0.135 percentile \( F_{0.135} \), and \( M \) estimates the median \( M \). To obtain the values of \( U_p \), \( L_p \), and \( M \) tables from Gruska et al. along with some interpolation calculations are required.

Based on sample percentiles, Chang and Liu considered a different method for calculating \( F_{99.865} \), \( F_{0.135} \), and the median \( M \). The method is essentially based on sample percentiles which can be calculated using interpolations, and does not require the tables in Gruska et al. Applying this method we can

Table II. A comparison between \( C_p(u,v) \) and \( C_{Np}(u,v) \)

<table>
<thead>
<tr>
<th>Process</th>
<th>( C_p )</th>
<th>( C_{pk} )</th>
<th>( C_{pm} )</th>
<th>( C_{pmk} )</th>
<th>( C_{Np} )</th>
<th>( C_{Npk} )</th>
<th>( C_{Npm} )</th>
<th>( C_{Npmk} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.06</td>
<td>0.00</td>
<td>0.26</td>
<td>0.00</td>
<td>1.00</td>
<td>-0.08</td>
<td>0.29</td>
<td>-0.02</td>
</tr>
<tr>
<td>B</td>
<td>1.06</td>
<td>1.06</td>
<td>1.06</td>
<td>1.06</td>
<td>1.00</td>
<td>0.92</td>
<td>0.97</td>
<td>0.89</td>
</tr>
<tr>
<td>C</td>
<td>1.06</td>
<td>0.00</td>
<td>0.26</td>
<td>0.00</td>
<td>1.00</td>
<td>0.08</td>
<td>0.34</td>
<td>0.03</td>
</tr>
</tbody>
</table>

non-normal process capability indices

Figure 1. Distributions of processes A, B, and C.
obtain the percentile estimators for $C_{Np}(u,v)$, which may be expressed as the following:

$$C_{Np}(u,v) = \frac{d-M-m}{\sqrt{\frac{F_{99.865}}{6} + v(M-T)^2}}$$

$$F_{99.865} = X_{(R_1)} + \left( \frac{99.865 + 0.135}{100} - R_1 \right)$$

$$X_{(R_1)} = X_{(R_2)} + \left( \frac{0.135n + 99.865}{100} - R_2 \right)$$

$$M = X_{(R_3)} + \left( \frac{n+1}{2} \right) - R_3$$

where $R_1 = [(99.865n + 0.135)/100]$, $R_2 = [(0.135n + 99.865)/100]$, $R_3 = [(n+1)/2]$. In this setting, $[R]$ is defined as the greatest integer less than or equal to the number $R$, and $X_{(i)}$ is defined as the $i$th order statistic.

6. AN APPLICATION

To illustrate how to calculate process capability using $C_{Np}(u,v)$, we consider the following example taken from a company who is a manufacturer and supplier of speaker components (parts) supplying various kinds of rubber edges to speaker driver manufacturing factories for making speaker driver units. A standard (woofer) driver unit, as depicted in Figure 2, consists of the following components (parts) including edge, cone, dustcap, spider (damper), voice coil, lead wire, frame (basket), magnet, front plate, and back plate (T-york). The rubber edge is one of the key components which reflect sound quality of the driver units, such as musical image and clarity of the sound. One characteristic of the rubber edge which we were interested in is the weight. For each model of rubber edges, a unique production specification ($USL$, $T$, $LSL$) is set to the manufacturing processes. The weight of the rubber edge should not fall outside the specification intervals or the customers will not accept the products.

In the rubber-edge manufacturing factory, the raw rubber is first compounded through the kneader with some chemical powder. The compounded raw rubber is then cut into thin rubber strips with appropriate length, loaded onto the mold machines, and thermocasted into the desired shape of rubber edges. Different models of rubber edges have different designs, shapes, weights, and have different production specifications. One characteristic of the rubber edge which we studied was the weight. The upper and lower specification limits, $USL$ and $LSL$, of the weight for a particular model of rubber edge, which we studied, were set to 8.94 and 8.46 (in grams). The target value is the mid-point between the two specification limits, which is 8.70. The collected sample data (a total of 100 observations) are displayed below in Table IV.

Figure 3 displays the normal probability plot for the collected data. We also perform Shapiro–Wilk test for normality check, obtaining $W = 0.91$ with $p$-value = 0.0001. Since the $p$-value is sufficiently small, we may conclude that the data set comes from a non-normal distribution. To calculate the values of the estimators $C_{Np}(u,v)$, we first calculate the sample percentiles obtaining $F_{0.0135} = 8.53$, $F_{99.865} = 9.03$, and $M = 8.69$. Then, we substitute these values into the definition of $C_{Np}(u,v)$ obtaining $C_{Np} = 0.96$, $C_{Npk} = 0.92$, $C_{Npm} = 0.95$, and $C_{Npnm} = 0.91$. We note that $C_{Npk}$ value is less than 1.00, which indicates that the process is not adequate with respect to the given manufacturing specifications, either the process variation ($\sigma^2$) needs to be reduced or the process mean ($\mu$) needs to be shifted closer to the target value. In fact, there are four observations (8.98, 8.99, 9.00, 9.03) falling outside the specification interval ($LSL$, $USL$), and the proportion non-conforming is 4%.

The quality condition of such a process was considered to be unsatisfactory in the company. Some quality improvement activities involving Taguchi’s parameter designs, were initiated to identify the significant factors causing the process failing to meet the company’s requirement. Consequently, machine settings for cutting the rubber strips as well as other process parameters were adjusted. To check whether the adjusted process was satisfactory, a new sample of 100 from the adjusted process were collected.
359

Table IV. 100 Observations of weight

\begin{tabular}{cccccccc}
8.61 & 8.81 & 8.72 & 8.69 & 8.65 & 8.64 & 8.68 & 8.74 \\
8.64 & 8.68 & 8.98 & 8.70 & 8.74 & 8.75 & 8.66 & 9.00 \\
8.53 & 8.74 & 8.59 & 8.69 & 8.70 & 9.03 & 8.83 & 8.87 \\
8.76 & 8.71 & 8.74 & 8.70 & 8.62 & 8.64 & 8.66 & 8.71 \\
8.67 & 8.71 & 8.81 & 8.63 & 8.67 & 8.72 & 8.66 & 8.74 \\
8.76 & 8.71 & 8.63 & 8.74 & 8.67 & 8.69 & 8.69 & 8.70 \\
8.71 & 8.81 & 8.60 & 8.64 & 8.79 & 8.69 & 8.68 & 8.77 \\
8.66 & 8.72 & 8.81 & 8.63 & 8.78 & 8.69 & 8.70 & 8.59 \\
8.76 & 8.64 & 8.54 & 8.71 & 8.61 & 8.80 & 8.70 & 8.59 \\
8.71 & 8.81 & 8.60 & 8.64 & 8.53 & 8.74 & 8.69 & 8.53 \\
\end{tabular}

Note that for the new process the departure ratio \( k = |T - \mu|/d = 0.01 \) is quite small, which indicates that the new process is nearly on-target. As a result, the quality of the new process improved significantly, and was considered to be satisfactory in the company.

7. CONCLUSIONS

In this paper, we considered some generalizations of the basic indices \( C_p(u,v) \), which we referred to as \( C_{Np}(u,v) \), to cover non-normal distributions. If the underlying distribution is normal, then the proposed generalizations \( C_{Np}(u,v) \) reduce to the basic indices \( C_p(u,v) \). The proposed generalizations \( C_{Np}(u,v) \) are compared with the basic indices \( C_p(u,v) \) and other generalizations \( C_9(u,v) \). The results indicated that the proposed generalizations \( C_{Np}(u,v) \) are more accurate than \( C_p(u,v) \) and \( C_9(u,v) \) in measuring process capability.

In addition, we considered an estimation method based on sample percentiles to calculate \( C_{Np}(u,v) \). Computations for obtaining the estimators \( C_{Np}(u,v) \) do not require any statistical tables, or any assumptions on the underlying distributions. We also gave an example on speaker components manufacturing process to illustrate how we apply the proposed generalizations \( C_{Np}(u,v) \) to the actual data collected from the factory. The calculations are easy to understand, straightforward to apply, and should be encouraged for applications.

Acknowledgement

The authors thank the Chief Editor, Henry A. Malec, and the anonymous referees for their careful reading of the paper and several suggestions which improved the paper.

REFERENCES

Table V. Second 100 Observations of weight

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8.70</td>
<td>8.69</td>
<td>8.71</td>
<td>8.70</td>
<td>8.66</td>
<td>8.67</td>
<td>8.68</td>
<td>8.73</td>
<td>8.66</td>
</tr>
<tr>
<td>8.65</td>
<td>8.66</td>
<td>8.69</td>
<td>8.71</td>
<td>8.69</td>
<td>8.70</td>
<td>8.67</td>
<td>8.73</td>
<td>8.66</td>
</tr>
<tr>
<td>8.72</td>
<td>8.80</td>
<td>8.72</td>
<td>8.94</td>
<td>8.81</td>
<td>8.70</td>
<td>8.74</td>
<td>8.71</td>
<td>8.73</td>
</tr>
<tr>
<td>8.80</td>
<td>8.70</td>
<td>8.52</td>
<td>8.65</td>
<td>8.73</td>
<td>8.70</td>
<td>8.67</td>
<td>8.74</td>
<td>8.67</td>
</tr>
<tr>
<td>8.94</td>
<td>8.68</td>
<td>8.62</td>
<td>8.70</td>
<td>8.69</td>
<td>8.67</td>
<td>8.70</td>
<td>8.66</td>
<td>8.73</td>
</tr>
<tr>
<td>8.65</td>
<td>8.74</td>
<td>8.75</td>
<td>8.69</td>
<td>8.70</td>
<td>8.67</td>
<td>8.68</td>
<td>8.70</td>
<td>8.66</td>
</tr>
<tr>
<td>8.63</td>
<td>8.70</td>
<td>8.92</td>
<td>8.71</td>
<td>8.67</td>
<td>8.68</td>
<td>8.67</td>
<td>8.70</td>
<td>8.66</td>
</tr>
<tr>
<td>8.67</td>
<td>8.68</td>
<td>8.69</td>
<td>8.67</td>
<td>8.69</td>
<td>8.69</td>
<td>8.70</td>
<td>8.66</td>
<td>8.64</td>
</tr>
<tr>
<td>8.66</td>
<td>8.69</td>
<td>8.68</td>
<td>8.69</td>
<td>8.68</td>
<td>8.69</td>
<td>8.70</td>
<td>8.66</td>
<td>8.64</td>
</tr>
<tr>
<td>8.69</td>
<td>8.65</td>
<td>8.93</td>
<td>8.64</td>
<td>8.67</td>
<td>8.64</td>
<td>8.68</td>
<td>8.77</td>
<td>8.64</td>
</tr>
</tbody>
</table>


Authors’ biographies:

K. S. Chen received his M.S. in Statistics from National Cheng Kung University, and Ph.D. in Quality Control from National Chiao Tung University. Currently, he is an associate professor in Department of Industrial Engineering and Management, National Chin-Yi Institute of Technology, Taichung, Taiwan, R.O.C.

W. L. Pearn is a professor of operations research and quality management in Department of Industrial Engineering & Management, National Chiao Tung University, Taiwan, R.O.C. He received his M.S. in statistics, and Ph.D. in operations research from the University of Maryland at College Park, MD, U.S.A. He worked for AT&T Bell Laboratories at Switch Network Control Center as a System Engineer, and Process Quality Center as a Quality Engineer.