1.1 Statistical Process Control

A production process is a unique combination of four Ms, namely Machines, Material, Methods and Man; the role of each of which is to work in a synchronized manner to produce the item/service under consideration with all its intended features at its best possible extent, (Montgomery, 2005 a). Capability is typically defined in dictionaries as the ability to carry out a task or goal. In connection with a production process, this task or goal, is its ability to meet the customer requirements which are more formally referred to as specification limits. These specification limits reflect the true requirements about the product/service in terms of its form, fit and function. Process Capability, is thus concerned with the ability of the process to meet the customer specifications regarding the
product/service. So, the success of a production process engaged in mass production, is determined by the extent to which it meets the customer requirements.

To give the discussion a more precise form, let \( X \) be the characteristic of interest regarding the product/service, the formal reference to which is process variable/characteristic, the lowest value of \( X \) which the customer is willing to accommodate be LSL, the Lower Specification Limit and the corresponding highest value of \( X \) be USL, the Upper Specification Limit. Also, assume that \( \mu \) and \( \sigma^2 \) are the mean and variance of \( X \), respectively. The probability distribution of \( X \) is assumed to be normal, unless otherwise mentioned. Under this set up, it is known that 99.73\% of the observations on \( X \) will fall in the interval \([\mu - 3\sigma, \mu + 3\sigma]\).

Process capability analysis deals with how to assess the capability of a manufacturing process, where information about the process is used to improve the capability. With process capability analysis, one can determine how well the process will perform relative to product requirements or specifications. As the subject matter of process capability analysis is to judge/predict the capability of the process in meeting the customer requirements, it has to be ascertained before hand that the process operates in a stable/predictable manner. To understand what is meant by this predictable state of a production process, we have to consider that branch of Statistical Quality Control called Statistical Process Control, see Duncan (1986), Grant and Leavenworth (1996) and Montgomery (2005 a). It should be noted that a process capability analysis could be performed even when the process is unstable, but the results are of momentous importance and hence, of limited use.

Every production process is so designed that the output come out with the aimed-at / target value of \( X \), as they are best acceptable to consumers. The quality of the output will be considered as maximum if there are no variations from the target value. But no process in this universe is, hitherto, found operating at this level of perfect output. Wheeler and Chambers (1992) explain it like this: no two things are alike. The villain of perfect process output is, variation, which invades every process and system known to this world.

The first significant attempt to understand and remove process variation in a scientific way was made by Walter A. Shewhart of Bell Telephone Laboratories in 1920’s. He conducted his studies on understanding the various components of variation, and that made the birth of the science Statistical Process Control (SPC). He recognized that every process is influenced by two sources of variability common/natural/random/chance cause and special/assignable cause. Special cause variability is due to identifiable sources which can be systematically identified and eliminated using various process
control methods. Common cause variability is due to the cumulative effect of many unidentifiable sources which cannot be removed without undue expenses and/or process redesign. The postulates and the findings of Shewhart were first submitted on May 16, 1924 and the final conclusions are available in his book *The Economic Control of Quality of Manufactured Product* (1931).

A process under the current design is said to be operating in the state of statistical control, if only common cause of variability is present in the process. This is the best level at which any process can perform as there will only be minimum variation from the target/aimed-at value of $X$. This is when the process is said to be working in the predictable state, and suitable for predicting the capability of the process in meeting the requirements on $X$. SPC has developed powerful tools like control chart for checking and bringing the process into the stable state after eliminating all assignable causes. Further reduction of process variation from the state of statistical control may also be possible by overhauling the present production set-up by using Design Of Experiment (DOE) methods (Montgomery 2005 b).

Maintaining the production process in the state of predictable has quality connotations also, besides capability dimensions. Since chance variation can never be totally eliminated, keeping the production process in the state of control, is the key to product quality. Maintaining a stable process average and systematically reducing processes variation are the keys to achieving superior quality. Various ISO standards and other quality programmes like Six - sigma are being used world wide for this purpose.

### 1.2 Process Capability Analysis

After having reached the predictable/in-control state, one can describe the capability of the process in meeting the actual product requirements. A process operating in this state of control can be used to ascertain what is known as “natural tolerance limits”, the limits within which majority of $X$ will lie. The lowest value of these is called Lower Natural Tolerance Limit (LNTL), and the biggest, Upper Natural Tolerance Limit (UNTL). If the distribution of $X$ is assumed to be normal with mean $\mu$ and standard deviation $\sigma$, then these limits will be $\mu - 3\sigma$ and $\mu + 3\sigma$ respectively, and it will contain 99.73% of the values of $X$. Process capability compares the output of an in-control process with the customer requirements expressed in terms of specification limits. A production process operating in the state of statistical control, on the process variable $X$ is said to be capable of meeting the requirements if all the values of $X$ are such that LSL < $X$ < USL. This can happen when the spread of the probability distribution of $X$ is fully contained in the specification spread, USL - LSL. In ordinary circumstances, an
in-control process can meet this situation and hence capable. This situation when the probability distribution is normal, is pictorially given in figure 1.1.

Thus, a “normal” process will be capable if the “natural tolerance interval”, \([\mu - 3\sigma, \mu + 3\sigma]\) \(\subset\) [LSL, USL], the specification tolerance interval. It should be noted that the “natural tolerance” \((\mu + 3\sigma) - (\mu - 3\sigma) = 6\sigma < \text{USL-LSL}\), the “specification tolerance”, is not sufficient for the process to be capable, in general. This is because, only when the mean of the distribution of \(X\) is located at such positions between LSL and USL that the spread of the distribution is fully contained in the specification interval, the process will be capable. In general, if LNTL and UNTL are the natural tolerance limits of the distribution of \(X\) between which it varies when the process is in control, then the process will be capable of meeting the customer requirements also, if \([\text{LNTL}, \text{UNTL}]\) \(\subset\) [LSL,USL], and other wise, not. An ideal situation in which the process can be capable is when the mean of the distribution is at mid-way between the specification interval and the natural tolerance is much less than the specification tolerance. That is, when \(\mu = \frac{(\text{LSL} + \text{USL})}{2}\) and \(6\sigma < \text{USL-LSL}\). A process whose mean \(\mu = \frac{\text{LSL} + \text{USL}}{2}\) is called a centered process. If the process is not fully capable, the extent to which it is capable can be understood by comparing natural tolerance with specification tolerance. The above being the idea of process capability, it can also be used to answer other capability related problems such as

1. What happens to the capability if the mean of the distribution of \(X\) is exactly at the more interested target value of \(X\), than at the mid-point of the specification
Figure 1.2: Normal distribution and proportion outside the specification limits.

2. What happens if the range of variation of $X$ is increased or decreased from the present “manageable” magnitude?

3. What happens if the mean of $X$ is shifted to a lower or a higher value from the present “safe” position?

A complete solution to these problems can be obtained by computing what is known as the proportion of non-conforming items in the following way.

**Calculation of the proportion of non-conforming items**

If the process is in control and the entire output is not lying between LSL and USL, then non-conforming items will be produced. The proportion of non-conforming (PNC) items produced can be calculated as

$$PNC = 1 - F(USL) + F(LSL)$$

where $F(\cdot)$ is the cumulative distribution function of the process variable $X$. If the distribution of $X$ is assumed to be normal with mean $\mu$ and variance $\sigma^2$, then

$$PNC = 1 - \Phi \left( \frac{USL - \mu}{\sigma} \right) + \Phi \left( \frac{LSL - \mu}{\sigma} \right)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. PNC for a “normal” process with mean $\mu$ and standard deviation $\sigma$, and fixed LSL and USL, is shown in figure 1.2.
Process Capability and Quality of Conformance

In Process Capability Analysis, what is explicitly ensured is the production of almost all items within the specification interval. All items produced with the value of $X$ such that $LSL < X < USL$ will be acceptable to consumers. So, the quality of conformance to the requirements is ultimately achieved, though implicitly. Even though no rejection will take place as long as $LSL < X < USL$, there are usually more/most preferred values which in most of the cases is $X = (LSL+USL)/2$. In such cases production processes will be so arranged that production takes place with this value of $X$ as far as possible, for which various process control techniques such as control charts can be used. This way, an improvement in the quality of conformance is the ultimate benefit of every process capability analysis.

Deleryd (1996) made the following remarks about process capability: the objective of using process capability studies is first of all to receive information about the process, information that can form a base for improvement efforts leading to a capable process, or perhaps an even more capable process. The information received during a process capability study is also useful in the overall improvement of the process with in an organization. Based on the information, correct priorities can be made of where to initiate most of the improvement efforts. It is important to know that it is not the process capability studies themselves that bring about improvement, but rather the actions taken, based on the result from the process capability studies.

Process capability analysis has proved to be a very valuable engineering decision tool. Quality engineers and managers are often faced with decision problems such as whether to accept or scrap a batch of process output, to intervene or not to intervene in a production process, or to review or not to review a managerial decision to intervene in a production process etc. These types of decisions are based on the estimated proportion of non-conforming material in a batch or lot. Exciting new decision-making procedures for on-line control situation which are proactive with respect to both prediction and control have been introduced. The results of properly executed process capability analysis will provide accurate information for these decisions.

Deleryd (1996) had developed a list of 13 most common ways of using the results from process capability studies. They are:

1. As a basis for improvement in the process
2. As an alarm clock
3. As specifications for investments
4. As a certificate for customers
5. As a basis for getting specifications at reasonable levels.
6. For monitoring machine capability
7. As a basis for introducing new products
8. For assessing the reasonableness of customer demands
9. For motivation of co-workers
10. For deciding priorities in the improvement of the process
11. As a base of inspection activities
12. As a receipt for improvements
13. For formulating quality improvement programs

It is said that process capability analysis must be addressed and discussed while the product’s life cycle is being considered, rather than considering it at the manufacturing stage. Finley (1992) recommended the following procedure for implementing process capability analysis in the life cycle of a product:

1. The customer defined the nominal or target value.
2. The customer’s engineers propose specification limits which allow manufacturing variability, while also ensuring that the part will provide the intended functions.
3. The manufacturer analyzes the process that would make this product, to determine if it can hold the customer’s required specification.
4. The customer and the manufacturer agree on specification limits. The resulting contract is called an approved drawing or blue print.
5. After the manufacturer has begun production, he conducts process capability studies to compare his manufacturing output to the required specification range of the customer.
CHAPTER 1. PROCESS CAPABILITY INDICES AND GENERALIZED CONFIDENCE INTERVALS

1.3 Process Capability Indices (PCIs)

Though the essence of process capability is to compare natural tolerance with specification tolerance, the method of practising it in actual production sites may be difficult unless it is simplified to some other form wherein which the computation of a single quantity can tell the extent to which the process meets the requirements. A search for providing proof of quality via process capability study naturally ends up with the construction of one or the other of a set of process capability measures more commonly known as process capability indices (PCIs). The concept of PCIs was introduced by Juran et al. (1974). They realized the need for a single ratio or index to compare the specification interval with the actual process spread or variation as a measure of process capability. Construction of a process capability index has become the most frequently used practice in any process capability analysis. A process capability index is a unit-free measure that quantifies the relation between the actual performance of the process and its specified requirements. These indices are proved to be extremely useful in attaining various production targets when used in conjunction with Engineering Quality Control methods. A process is defined capable if the capability index exceeds a threshold value which can never be less than one in any case. Its value can be used to measure the extent to which the process meets the specifications. In general, the higher the value of the index, the lower the amount of products outside the specification limits. If the value is not up to the requirements, improvement efforts can be initiated based on the value of the index.

The need for such a single unit-less number like PCI, might be felt by all concerned with the production process. People from managerial to floor level can make use of the value of the index to evaluate the capability of the process in meeting the requirements simultaneously with the product is being produced. It also helps initiate suitable corrective steps on-line to make the process capable by examining the value of the index.

From what is discussed so far, it is clear that a PCI must take into account both natural tolerance interval and specification interval simultaneously. So,

\[
\text{general form of PCIs} = \frac{\text{specification interval}}{\text{process spread}}
\]

If specification interval is greater than process spread in such a way that the natural tolerance limits are within the specification limits, then the process is capable of meeting the requirements, as can be seen in Figure 1.1.

From the general form of a PCI, it is clear that no process can certainly be capable
Table 1.1: Classification of processes based on the value of the capability index

<table>
<thead>
<tr>
<th>Capability value</th>
<th>Grading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 1</td>
<td>Inadequate</td>
</tr>
<tr>
<td>1 and less than 1.33</td>
<td>Capable</td>
</tr>
<tr>
<td>1.33 and less than 1.5</td>
<td>Satisfactory</td>
</tr>
<tr>
<td>1.5 and less than 2</td>
<td>Excellent</td>
</tr>
<tr>
<td>2 and above</td>
<td>Super</td>
</tr>
</tbody>
</table>

if the value of PCI is less than one. The situation is almost the same even when the value is one. Higher the value of the PCI, greater will be the capability of the process. A PCI value of $\frac{4}{3}$ is generally accepted as the minimum value of a capable process, with mild changes associated with individual PCIs. See Kotz and Lovelace (1998). Recommended minimum values of PCIs for existing and new processes are given in Table 2.1. Ensuring a high value to the PCI while the product is being built, can help achieve international standards like six-sigma philosophy, various ISO standards etc.

There is also the practice of grading the production processes based on the values of the index, as given in Table 1.1.

### 1.4 Some Important PCIs

In this section we consider some important PCIs in frequent use. The work on PCIs had been actually started much earlier before the formal introduction of the first PCI $C_p$ by Juran et al. (1974). Since then, there has been an avalanche of PCIs to suit the needs of various requirements. For a review of the work on PCIs during the period 1992-2000, see the article by Kotz and Johnson (2002). For a review of even earlier work, see the books by Kotz and Johnson (1993) and Kotz and Lovelace (1998). See also Chapter 7 of Montgomery (2005 a). An extensive bibliography of papers on PCIs during the period 1990-2002 is available in Spiring et al. (2003). The latest book by Pearn and Kotz (2007) is an authoritative treatise on PCIs.

#### The $C_p$ Index

The history of PCIs began with the introduction of the $C_p$ index by Juran et al. in 1974, though it did not get considerable acceptance, until early 1980s. They defined the index as

$$C_p = \frac{USL - LSL}{6\sigma} = \frac{\text{allowable range of measurements}}{\text{actual range of measurements}}.$$  

Finley (1992) refers to this index as Capability Potential Index (CPI), and Montgomery (2005 a) as Process Capability Ratio (PCR). The numerator of $C_p$ gives the size of the
range over which the process measurements can vary. The denominator gives the size of the range over which the process is actually varying. Obviously, it is desirable to have a $C_p$ as large as possible to attain high capability. But even when $C_p$ is large, the process may not be capable, unless the process mean is located at safer positions between LSL and USL as can be seen in figure 1.3.

Thus the fact that neither the numerator nor the denominator of $C_p$ takes into account the actual location of the process mean, remains to be a serious drawback of the most commonly referred PCI, and hence it is regarded only as a process potential index and not as a process performance index. From the expression of PNC, the proportion of defective produced when the process is centered, ie. when $\mu = \frac{LSL+USL}{2}$ is calculated as $2\Phi(-3C_p)$ and in all other cases it will be greater than this quantity. Table 1.2 gives the proportion of defective in Parts Per Million (PPM) for various values of $C_p$, when the process is centered.

**The $C_{pk}$ Index**

As has already been noted, a change in process capability occurs due to either a change in process mean or process variation or both. A capacity index that takes into
account the change in both the process mean and process variation was felt by many, and that resulted in the proposal of the $C_{pk}$ index by Kane (1986). $C_{pk}$ is defined by a three-step procedure as follows: Let
\[
C_{pu} = \frac{USL - \mu}{3\sigma} = \frac{allowable\ upper\ spread}{actual\ upper\ spread}
\]
and
\[
C_{pl} = \frac{\mu - LSL}{3\sigma} = \frac{allowable\ lower\ spread}{actual\ lower\ spread}
\]
then
\[
C_{pk} = \min\{C_{pl}, C_{pu}\}
= \min\left\{\frac{\mu - LSL}{3\sigma}, \frac{USL - \mu}{3\sigma}\right\}
= \frac{d - |\mu - M|}{3\sigma},\ d = \frac{USL - LSL}{2},\ M = \frac{USL + LSL}{2}
\]
Figure 1.4 illustrates what is represented by the $C_{pk}$ index. Gunter (1989) picturesquely describes $C_{pk}$ as a way to measure the ratio of the amount of room needed to the amount of room available to produce product within specifications.

$C_{pu}$ compares the upper $3\sigma$ range of the process data to the distance between the upper specification limit and the process mean. $C_{pl}$ compares the lower $3\sigma$ range of the process data to the distance between the process mean and the lower specification limit. Basically, each ‘half’ of $C_{pk}$ measures how close the outer tail of the process distribution is to the specification limit. $C_{pk}$ is evaluated as the worst of the two. For different combinations of $\mu$ and $\sigma$, $C_{pk}$ can have the same value as is seen in figure 1.5.
Figure 1.4: Comparison of amount of room available to amount of room needed

Figure 1.5: Three processes with $C_{pk} = 1$. 
Table 1.3: Approximate proportion non-conforming associated with given $C_{pk}$ values.

<table>
<thead>
<tr>
<th>$C_{pk}$</th>
<th>Parts outside tolerance limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>16 out of 100</td>
</tr>
<tr>
<td>0.50</td>
<td>7 out of 100</td>
</tr>
<tr>
<td>1.00</td>
<td>13 out of 10,000</td>
</tr>
<tr>
<td>1.33</td>
<td>3 out of 1,000,000</td>
</tr>
<tr>
<td>1.67</td>
<td>1 out of 10,00,000</td>
</tr>
<tr>
<td>2.00</td>
<td>1 out of 1,000,000,000</td>
</tr>
</tbody>
</table>

$C_{pk}$ was created in Japan to offset some of the weaknesses of $C_p$, primarily the fact that $C_p$ measures capability in terms of process variation only and does not take process location into account. This is best displayed by process $C$ whose $C_p$ value is higher than $A$ and $B$ but it is likely to be in danger in the event of its location gets shifted further right. Thus location of the process at safe positions between the specification limits, is as important as acquiring the minimum possible process variation for achieving high process capability. Therefore, a higher value of $C_p$ cannot ensure process capability unless and until the process average is located at such positions between LSL and USL that small changes in the process average in the immediate future cannot make defectives. Kotz and Johnson (1993) obtained the following relationship between the percentage non-conforming and the value of $C_{pk}$

$$100\Phi(-3C_{pk}) \leq PNC \leq 200\Phi(-3C_{pk}).$$

Finley (1992) developed Table 1.3 of approximate proportion non-conforming associated with given $C_{pk}$ values. Finley’s proportion non-conforming estimates assume that there are no non-conforming pieces for the ‘better half’ of the process distribution.

The $C_{pm}$ Index

The indices $C_p$ and $C_{pk}$ do not take into account the fact that the process mean $\mu$ may differ from some other more interested target value $T$. As per modern quality improvement theories, it is important to use target values and to keep the process on target. See for example Bergman and Klefsjö (2003). The idea is to design capability indices that deem the process as non-capable if the process mean is far away from the target value even when the probability of non-conformance is small.

The case when $\mu \neq \frac{USL + LSL}{2} = M$, is referred to as asymmetric tolerances and most of the earlier PCIs were designed for symmetric tolerances. PCIs, free of target values $T$, generally assume their maximum value when $\mu = \frac{USL + LSL}{2}$, and hence in
such cases process conditions are so arranged that \( \mu \) is at \( M \), as far as possible. This is not the case of indices incorporating a target value.

The following are two cases where asymmetric tolerances can occur in practice. The first is the case where the process distribution is not normal but skewed, in which case the tolerances also will naturally be asymmetric. This will be the case even when the process distribution is symmetric but non-normal, as the assumption of normality is key to process capability inference. In such cases where \( X \) has a non-normal distribution with/without a symmetric tolerance (LSL, T, USL), a suitable transformation \( Y = g(X) \) yielding normality for some increasing non-linear function \( g \), will generally result with the asymmetric specification \( (g(\text{LSL}), g(T), g(\text{USL})) \). Such an example is considered in sections 2.4.4. where Generalized Confidence Intervals are constructed in connection with \( C_{pmk} \).

In the second case, an asymmetric tolerance can arise from the customers’ view that deviations from target are less tolerable in one direction than in the other. For example, suppose that an initial symmetric specification of say, \( 42 \pm 3 = (39, 42, 45) \) has been proposed by the customer for a new product, before knowing much about the capability of the process. But if it is now seen, that the supplies’ production process could not hold the upper specification at 45 without expensive process overhauling, the customer may in that context be willing to expand the upper specification limit to, say 51 and thereby establishing an asymmetric tolerance. See for more details Boyles (1991, 1994), Pearn et al. (2001, 2004), and Lin and Pearn (2002).

Chan, Cheng and Spiring (1988) introduced such a capability index that incorporates a target value into its formula. The index proposed by them can be taken as a measure of process centering, and it is defined by

\[
C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{d}{3\sqrt{\sigma^2 + (\mu - T)^2}}.
\]

It is clear, from the above expression, that process capability can be increased by either reducing the process variation or moving process average close to the target value or both. \( C_{pm} \) can also be expressed in terms of \( C_{pk} \) and \( C_p \) as

\[
C_{pm} = \frac{C_{pk}}{(1 - k)\sqrt{1 + \frac{(\mu - T)^2}{\sigma^2}}} = \frac{C_p}{\sqrt{1 + \frac{(\mu - T)^2}{\sigma^2}}}
\]

where \( k = \frac{|\mu - M|}{d} \) is one of the original Japanese indices. The index \( k \) describes process capability in terms of process location only and provides a quantified measure of the
amount that a process is off-centre.
Parlar and Wesolowsky (1998) noticed that \( C_p, C_{pk}, \) and \( C_{pm} \) are related by the formula

\[
C_{pm} = \frac{C_p}{\sqrt{1 + 9(C_p - C_{pk})^2}}.
\]

Kotz and Johnson (1993) found that the expected proportion of non-conforming items associated with a given value of \( C_{pm} \) is

\[
p = \Phi \left( \frac{-d - \mu}{\sqrt{\lambda^2 - \mu^2}} \right) + 1 - \Phi \left( \frac{d - \mu}{\sqrt{\lambda^2 - \mu^2}} \right)
\]

where \( \lambda = \frac{d}{3C_{pm}}. \)

The \( C_{pk}'' \) Index
A process is said to be centered or to have a symmetric tolerance if the target value \( T \) is the midpoint of the specification interval i.e. when \( T = M \). Though symmetric cases are common, asymmetric cases do occur in the manufacturing industry as just seen in the case of \( C_{pm} \). Since \( C_{pk} \) index is independent of \( T \), it fails to distinguish between on-target and off-target processes. Hence \( C_{pk} \) cannot be used to evaluate process capability when asymmetric tolerances are there.

To overcome this problem, Pearn and Chen (1998) proposed the \( C_{pk}'' \) index, defined by

\[
C_{pk}'' = \frac{d^* - A^*}{3\sigma}
\]

where

\[
d^* = \min\{USL - T, T - LSL\} \quad \text{and} \quad A^* = \max \left[ \frac{d^*(\mu - T)}{USL - T}, \frac{d^*(T - \mu)}{T - LSL} \right]
\]

The distributional properties of this index are studied in detail by Pearn et al. (2004).

The \( C_{pmk} \) Index
This third-generation capability index was introduced by Pearn et al. (1992), which is a combination of \( C_{pk} \) and \( C_{pm} \) and is defined as

\[
C_{pmk} = \min(USL - \mu, \mu - LSL) \quad \text{and} \quad C_{pmk} = \frac{d - |\mu - M|}{3\sqrt{\sigma^2 + (\mu - T)^2}}
\]
It can also be expressed in terms of $C'_{pm}$ and $C'_p$ as given below.

$$C'_{pm} = (1 - k)C'_{pm}$$

$$C'_p = \sqrt{1 + \left(\frac{\mu - M}{\sigma}\right)^2}$$

Thus $C'_{pmk}$ and $C'_{pm}$ are related the same way as $C'_{pk}$ and $C'_p$ ($C'_{pk} = (1 - k)C'_p$).

Wallgren (1996) had studied the sensitivity of $C'_{pmk}$ in detecting the deviations from target value and came to the conclusion that $C'_{pmk}$ is more sensitive than $C'_{pk}$ or $C'_{pm}$.

Vännman (1995) had compared $C_p, C'_{pk}$ and $C'_p$ with $C'_{pmk}$ and found that $C'_{pmk}$ is more restrictive with regard to process means’ deviation from the target value than the other indices. Pearn and Kotz (1994), in fact, ranked the indices $C_p, C'_{pk}, C'_p$ and $C'_{pmk}$ in their order of sensitivity to differentiate the difference between the process mean and the target value as (1) $C'_{pmk}$ (2) $C'_{pm}$ (3) $C'_{pk}$ (4) $C'_p$.

The $C_p(u, v)$ index

Instead of simply subtracting $|\mu - M|$ in the numerator and adding $(\mu - T^2)$ in the denominator of the general form of a PCI, Vännman (1995) proposed a new index as a function of suitable multiples of these two quantities, and is defined as

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

It is clear from this expression that the values of $u$ and $v$ are to be large if we want the index to be sensitive to departures of $\mu$ from $T$. Obviously,

$$C_p(0, 0) = C_p, \quad C_p(1, 0) = C'_{pk}$$
$$C_p(0, 1) = C'_{pm}, \quad C_p(1, 1) = C'_{pmk}$$

Based on detailed numerical studies involving statistical considerations, including power of a test argument, she suggests that $u = 0$ and $v = 4$ will produce a useful PCI.

1.5 Conventional Confidence Intervals Revisited

Confidence intervals quantify our knowledge, or lack thereof, about a parameter or some other characteristic of a population, based on a random sample. For example, asserting that a 95% confidence interval to contain the mean life of a particular brand of light bulb is 800 to 900 hours is considerably more informative than simply stating that the mean life is approximately 850 hours.

The classical confidence interval estimation procedure for obtaining exact confi-
CONFIDENCE INTERVALS INVOLVING NUISANCE PARAMETERS IS FOUND SUCCESSFUL ONLY IN SPECIAL CASES. FOR EXAMPLE, EXACT CONFIDENCE INTERVALS (CIs) FOR $\mu^2 + \sigma^2$, WHERE $\mu$ AND $\sigma$ ARE THE MEAN AND STANDARD DEVIATION OF A NORMAL DISTRIBUTION ARE NOT AVAILABLE. THE BEHRENS-FISHER PROBLEM IS PROBABLY THE MOST WELL KNOWN PROBLEM IN WHICH EXACT CIs BASED ON MINIMAL SUFFICIENT STATISTICS DO NOT EXIST. THE UN-AVAILABILITY OF EXACT CIs IN MANY SITUATIONS CAN BE ATTRIBUTED TO THE CONVENTIONAL APPROACH TO CONSTRUCTING CIs. TO EXPLAIN THE IDEA OF GENERALIZED CONFIDENCE INTERVALS, LET US BEGIN WITH THE CONSTRUCTION OF CONVENTIONAL CIs.

Consider a parametric statistical problem where we observe $X = (X_1, X_2, \ldots, X_n)$ whose distribution belongs to some family of distributions parameterized by $\xi \in \Xi \subset \mathbb{R}^p$. Suppose we are interested in the $100\gamma\%$ CI for the scalar parameter $\theta = \pi(\xi), \theta \in \Theta$. Let $A(X)$ and $B(X)$ be two statistics satisfying

$$P[A(X) \leq \theta \leq B(X)] = \gamma$$

(1.5.1)

where $\gamma \in [0, 1]$. Then $[A(x), B(x)]$ is a conventional CI of confidence coefficient $\gamma$, where $x$ is an observed value of $X$.

Though this method of constructing CIs is conceptually simple and easy to implement, its drawback is that in most applications especially involving nuisance parameters, it is either not easy or even impossible to find $A(X)$ and $B(X)$ satisfying equation (1.5.1) for all possible values of the nuisance parameters. The idea in GCIs is to make this possible by making probability statements relative to the observed sample, as done in Bayesian and Non-parametric methods. In other words, we allow the functions $A(\cdot)$ and $B(\cdot)$ to depend not only on the observable random vector $X$ but also on the observed data $x$.

The conventional CIs obtained by the definition of equation (1.5.1) have the following two repeated sampling properties in the context of repeated experiments:

Property 1: If the same experiment is repeated a large number of times to obtain new sets of observations, $x$, of the same sample size $n$, then the true value of $\theta$ is going to be included in say $95\%$ of the CIs constructed as per definition (1.5.1).

Property 2: After a large number of situations of setting $95\%$ CIs for certain parameters of independent statistical problems, the experimenter will have correctly included the true values of the parameters in the corresponding intervals $95\%$ of the time.

Clearly, Property 1 implies Property 2. Property 1 is useful in simulation studies where one needs to study the performance of various types of CIs. Except in this
kind of theoretical studies, Property 1 cannot be entertained as a practically useful repeated sampling property, because if one can obtain repeated samples from the same experiment, the claimed confidence level will no longer be valid. Moreover if sampling is repeated a large number of times, then the parameter will become known exactly so that the computation of a CI is not of an issue. Insisting on the repeated sampling property with the same experiment may leave the experimenter with great difficulties or no possibility of constructing CIs. Also, the various repeated sampling properties of interval estimates obtained through the conventional method are mere implications of the probability statement (1.5.1), rather than something deliberately required in the construction of the interval.

1.6 Generalized Confidence Intervals (GCIs)

A search for CIs having at least Property 2, resulted in the origin of another class of CIs known as Generalized Confidence Intervals (GCIs) by Weerahandi (1993, 1995), thus enhancing the class of CIs. This was made possible by insisting on exact probability statements which will ensue more desirable interval properties rather than on repeated sampling properties. This will enable us to construct CIs in the context of large number of problems than what is possible with conventional methods. The idea is to do the best with the observed data rather than on all possible samples that could have been observed but were not. Since GCIs are generalizations of classical CIs, the former is reduced to the latter in the special cases in which the latter exists. Also, as in the case of conventional CIs, when there exists a GCI with $100\gamma\%$ confidence, there is usually a class of $100\gamma\%$ GCIs.

In many practical contexts, including the one in our consideration namely Statistical Process Control and Process Capability Studies where the parameter of interest may itself vary around the intended value, idea of repeating the same experiment is meaningless as the interval is no longer valid when additional samples are taken, and hence it is of little practical use. So the method of GCIs is found theoretically more meaningful than aspiring solely for the repeated sampling properties, in the context of interval estimation of PCIs. Besides this, the construction of conventional CIs for PCIs is more often not possible because of the presence of nuisance parameters.

As a conventional approach for constructing CIs is based on the notion of a pivotal quantity, a parallel approach is considered in the construction of GCIs to define what are known as Generalized Pivotal Quantities (GPQs). The CIs obtained by inverting the GPQs will be GCIs.
Definition of Generalized Pivotal Quantities (GPQs)

Let $S = (S_1, S_2, \ldots, S_k) \in \mathbb{R}^k$ denote a statistic based on $X = (X_1, X_2, \ldots, X_n)$. In theory we can consider an independent copy $X_1^*, X_2^*, \ldots, X_n^*$ of $X_1, X_2, \ldots, X_n$ and denote the statistic based on $X_i^*$'s by $S^*$. Thus $S^*$ is an independent copy of the observable random vector $S$ whose distribution is indexed by $\xi \in \mathbb{R}^p$. We will use $s$ and $s^*$ to denote realized values of $S$ and $S^*$ respectively. Then a GPQ for a scalar parameter $\theta = \pi(\xi)$, denoted by $R_\theta(S, S^*, \xi)$ is a function of $(S, S^*, \xi)$ with the following properties:

(GPQ1): The conditional distribution of $R_\theta(S, S^*, \xi)$, conditional on $S = s$, is free of $\xi$.

(GPQ2): For every allowable $s \in \mathbb{R}^k$, $R_\theta(s, s, \xi) = \theta$.

It may be noted that the second condition uses $s$ in both the first and the second argument positions of $R_\theta(\ldots, \xi)$. This is an important aspect of the definition of generalized pivotal quantity and explains both its similarity to, and its difference from, an ordinary pivotal quantity. $R_\theta(S, S^*, \xi)$ will become an ordinary pivotal quantity based on $(S, S^*)$ if, in condition (GPQ 2), we require that $R_\theta(s, s^*, \xi) = \theta$. The idea is that we do not intend to observe a realization of $S^*$ because under appropriate conditions, $S$ and $S^*$ will be sufficiently close to each other. Hence, we can substitute $S$ in both the argument positions of $R_\theta(\ldots, \xi)$ and use the distribution $R_\theta(s, s, \xi)$ from (GPQ1) to make approximate confidence statements about $\theta$.

As in the case of conventional pivotal quantities, (GPQ1) is imposed to ensure that a subset of the sample space of all possible values of $R$ can be found at a given value of the confidence coefficient $\gamma$, with no knowledge of parameters. (GPQ2) is imposed to guarantee that such probability statements based on a GPQ will lead to confidence regions involving observed $s$ only.

Now let $r_1$ and $r_2$ be such that

$$P[r_1 \leq R_\theta(S, S^*, \xi) \leq r_2] = 1 - \alpha,$$

then $[\theta; r_1 \leq R_\theta(s, s, \xi) \leq r_2]$ is a 100$(1 - \alpha)$% confidence interval for $\theta$. For example, if the value of $R_\theta(S, S^*, \xi)$ at $S = s$ is $\theta$, then $[R(s, s^* / 2), R(s, 1 - s^* / 2)]$ is a 100$(1 - \alpha)$% confidence interval for $\theta$, where $R(s, s^*)$ stands for the $s^*$th percentile of the conditional distribution of $R_\theta(S, S^*, \xi)$ conditionally on $S = s$.

At the time of introduction of the idea of GCIs, it was not sure what all repeated sampling properties would be satisfied by GCIs. But the idea has been soon found applied in a wide variety of problems, and the results were astonishing. The areas of application included comparison of means, testing and estimation of parametric func-
tions of normal and related distribution by Tsui and Weerahandi (1989), Johnson and Weerahandi (1998), Chang and Huang (2000), Krishnamoorthy and Mathew (2003), Gamage, Mathew and Weerahandi (2004); in testing fixed effects and variance components in repeated measures and mixed effects ANOVA models by Zhou and Mathew (1994), Gamage and Weerahandi (1998), Chiang (2001), Krishnamoorthy and Mathew (2004), Weerahandi (2004), Mathew and Webb (2005), Arendacka (2005), Lin and Liao (2006); in laboratory testing by Iyer, Wang and Mathew (2004); in Bioequivallence by McNally, Iyer and Mathew (2003); in growth curve modelling by Weerahandi and Berger (1999), Lin and Lee (2003); in reliability and system engineering by Tiao and Cappellari (2004), Roy and Mathew (2005); in health studies by Chen and Zhou (2006); in process control by Hamada and Weerahandi (2000), Burdick, Borror and Montogomery (2005). The simulation studies carried out in these problems showed that the true coverage of the GCIs were at least as large as the intended coverage. Thus, it took no time to establish that GCIs obey the practically more useful Property 2 of CIs.

Estimation of confidence intervals in the case of PCIs has also been experiencing the same problem of the presence of nuisance parameters in the probability distribution of the estimators, and hence the failure in obtaining a confidence interval of the conventional type. Finding the success of GCIs in the above contexts, we have decided to try its prospectus in the context of PCIs, and it came out with fruitful results: See Sebastian and Kurian (2006), Mathew, Sebastian and Kurian (2007) and Kurian, Mathew and Sebastian (2008). Our findings are also in line with that of others viz. repeated sampling properties of GCIs are fairly near to the nominal levels, and are better in many cases than that of the existing approximate methods. But it was always a concern for the advocates of GCIs to get the recognition of the conventional practitioners by showing the exact frequentist coverage (Property 1) for GCIs also. Weerahandi himself attempted to give a proof for the same. But, because of in sufficient conditions, he could show only the following result:

A $100\gamma\%$ GCI is a conventional CI with confidence coefficient at least $2\gamma - 1$. That is, a 97.5% GCI is satisfying property 1 and Property 2 of a conventional CI of confidence coefficient at least 0.95.

Hanning et al. (2006) came up with a set of fairly mild conditions under which the GCIs have the correct asymptotic coverage, and this is given in the next result:
CHAPTER 1. PROCESS CAPABILITY INDICES AND GENERALIZED CONFIDENCE INTERVALS

Result

(Hanning et al. 2006). Suppose, for each fixed \( s, n, \) and \( \gamma \in (0, 1) \), there exists a real number \( C_n(s, \gamma) \) such that

\[
\lim_{n \to \infty} P_{\xi}[R_{\theta}(s, S^*, \xi) \leq C_n(s, \gamma)] = \gamma,
\]

Then \( \lim_{n \to \infty} P_{\xi}[R_{\theta}(S, S, \xi) \leq C_n(S, \gamma)] = \gamma \)

In particular, since \( R_{\theta}(s, s, \xi) = \theta \), it follows that the interval \(-\infty < \theta \leq C_n(S, \gamma)\) is a one-sided confidence interval for \( \theta \) with asymptotic coverage probability equal to \( \gamma \).

1.7 A General Recipe for Constructing GPQs

As the construction of GCIs is based on the notion of GPQ, the construction of an appropriate GPQ is crucial. Until Hanning et al. (2006) proposed a general method, all the available works on GCIs were based on GPQs constructed independently, and in a completely initiative manner. Chiang (2001) proposed the method of surrogate variables for deriving approximate CIs for functions of variance components in balanced mixed linear models. But, Iyer and Patterson (2002) identified that the CIs of Chiang were identical to GCIs. Though, Chiang had proposed a systematic method for computing GCIs for the class of problems he considered, he had not extended it to other class of problems.

Iyer and Patterson (2002) formulated a method for constructing GPQs for \( \theta = \pi(\xi) \) and proved that the method works for the class of problems where there exists a \( k \)-dimensional pivotal quantity \( \mathbb{R} \) that possesses an invertible pivotal relationship with the parameter \( \xi \). Nearly every GCI in the published statistical literature may be obtained by the use of this recipe.

To understand the general recipe, we need the following two definitions.

Definition 1.7.1 Let \( S = (S_1, \ldots, S_k) \in S \subset \mathbb{R}^k \) be a \( k \)-dimensional statistic whose distribution depends on a \( p \)-dimensional parameter \( \xi \in \Xi \). Suppose there exist mapping \( f_1, \ldots, f_k \) with \( f_j : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R} \), such that , if \( E_i = f_i(S; \xi) \), for \( i = 1, \ldots, k \), then \( \mathbb{E} = (E_1, \ldots, E_k) \) has a joint distribution that is free of \( \xi \). We say that \( f(S, \xi) \) is a pivotal quantity for \( \xi \) where \( f = (f_1, \ldots, f_k) \).

Definition 1.7.2 Let \( f(S, \xi) \) be a pivotal quantity for \( \xi \) as described in Definition 1.7.1. For each \( s \in S \), define \( \varepsilon(s) = f(s, \Xi) \). Suppose the mapping \( f(s, \cdot) : \Xi \to \varepsilon(s) \) is invertible for every \( s \in S \). We then say that \( f(S, \xi) \) is an invertible pivotal quantity for \( \xi \). In this case we write \( g(s, \cdot) = (g_1(s, \cdot), \ldots, g_k(s, \cdot)) \) for the inverse mapping so that whenever \( e = f(s, \xi) \), we have \( g(s, e) = \xi \).
The Recipe: Let $S = (S_1, \ldots, S_k) \in S \subset \mathbb{R}^k$ be a $k$-dimensional statistic whose distribution depends on a $k$-dimensional parameter $\xi \in \Xi$. Suppose there exist mappings $f_1, \ldots, f_k$, with $f_j : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$, such that, $f = (f_1, \ldots, f_k)$ is an invertible pivotal quantity with inverse mapping $g(s, \cdot)$. Define

$$R_\theta = R_\theta(S, S^*, \xi) = \pi(g_1(S, f(S^*, \xi)), \ldots, g_k(S, f(S^*, \xi))) = \pi(g_1(S, E^*), \ldots, g_k(S, E^*))$$

where $E^* = f(S^*, \xi)$ is an independent copy $E$. Then $R_\theta$ is a GPQ for $\theta = \pi(\xi)$. When $\theta$ is a scalar parameter, an equal-tailed 2-sided $100(1 - \alpha)\%$ GCI for $\theta$ is given by $R_{\theta, \alpha/2} \leq \theta \leq R_{\theta, 1-\alpha/2}$. Here $R_{\theta, \gamma} = R_{\theta, \gamma}(s)$ denotes the $100\gamma^{th}$ percentile of the distribution of $R_\theta$ conditional on $S = s$. One-sided generalized confidence bounds are obtained in an obvious manner.

Hanning et al. (2006) had used this recipe to derive GPQs in a wide variety of problems. We use this method for constructing generalized pivotal quantities in the following chapters in connection with the construction of generalized confidence intervals for various process capability indices.
Bibliography


