Misconceptions about Process Capability Indices

M. Z. Anis
SQC & OR Unit
Indian Statistical Institute
Calcutta - 700 108
INDIA

Content

- Introduction
- Basic process capability indices
  - C_p index
  - C_{pk} index
  - C_{pm} index
  - C_{pmk} index
- Effects of Non-normality, Auto-correlation and Asymmetry
- Misconceptions concerning PCIs
- Conclusion

Introduction

- Meant for assessing the ability of manufacturing process.
- Firstly introduced by Juran et al. in 1974.
- They gave the concept of C_p to evaluate the adequacy of the process to meet the specification limits.
- Later the other indices C_{pk}, C_{pm}, and C_{pmk} were developed to avoid the drawbacks of C_p.

Introduction Cont....

- Key objectives of basic process capability indices

  i. Compare process performance with specifications.
  ii. Compare among different processes.
  iii. Provide information about processes: proportion nonconforming and closeness to target.
  iv. Provide directions for quality improvement.

Introduction Cont....

- The main objectives in quality improving centered around
  i. Minimization of the deviation of the process mean from the target value
  ii. Reduction in variability.

- Generally reduction in process variability is more difficult than correction in process mean.

The Basic Process Capability Indices

- The assumptions of the study of process capability indices are
  i. Process of interest is in the state of statistical control.
  ii. Process is generating observations according to a normal distribution.
  iii. Process mean μ and standard deviation σ are known.

- We now briefly describe the most popular process capability indices.

1. C_p index

Consider

- X: value of a measured characteristic
  - USL: Upper specification limit
  - LSL: Lower specification limit
- Values of X outside these limits will termed as Nonconformance (NC).

The process capability index C_p is defined as

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

C_p index cont....

- If the data are normally distributed and process mean is centered at the middle of the specifications then
  1. $C_p = 1$ implies 99.73% of the products are under the specification limits.
  2. $C_p = 1$ means the specification (USL-LSL) is six times the square root of the process variance.

- For calculating C_p, the knowledge of process mean is not required.

2. C_{pk} index

- It gives the influence of μ on the value of process capability.
- It is defined as

$$C_{pk} = Min\left\{\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right\}$$

$$= \frac{Min\left\{(USL-\mu),(\mu-LSL)\right\}}{3\sigma}$$

3. C_{pm} index

- It was published by Chan et al. (1988a).
- Spiring (1991a) has also thrown some light on this index.
- It is defined as

$$C_{pm} = \frac{(USL - LSL)}{6\sqrt{\sigma^2 + (\mu - T)^2}}$$

where

μ: expected value

σ: Standard deviation

T: Target value

4. C_{pmk} index

i. C_{pk} is derived from C_p by modifying the numerator.

ii. C_{pm} is derived from C_p by modifying the denominator.

- If the two modifications are combined, a new index C_{pmk} is derived.

It is defined as

$$C_{pmk} = Min\left\{ \frac{(USL-\mu)}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{(\mu - LSL)}{3\sqrt{\sigma^2 + (\mu - T)^2}} \right\}$$

Effects of Non-normality, Auto-correlation and Asymmetry

- Two assumptions are required to discuss process capability.
  1. Process is in state of statistical control.
  2. Observations are normally and independently distributed.

- But in general these assumptions are not always true.
- Most of the industrial data are non-normal.
- Similarly it is very common to have autocorrelation in the data.

Non-normal data

- A lot of work has been done assuming the data are normally distributed.
- But most of the industrial data are non-normal.
- Munechika (1992) has shown several industrial example where data were not the normally distributed.
- Somerville and Montgomery (1996-97) have shown that the indices may not give the correct figure if data are non-normal.

Non-normal data cont….

- **Empirical percentile method** discussed by Clements (1989) and McCormark (2000).
- **Transformation into normal data** by Johnson's system of distribution curve method suggested by Chou et al. (1998).
- Wu and Swain (2001) also suggested a method named **weighed variance method** to deal with non-normal data.
- Ding (2004) gave the method of calculating the process capability **using the first four moments** when data is non-normal.

Skewed data

- Wright (1995) suggested an index C_s to deal with skewed data which is defined as

$$C_s = Min\left\{ \frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - \mathrm{T})^2 + |\mu_3 / \sigma|}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - \mathrm{T})^2 + |\mu_3 / \sigma|}} \right\}$$

Skewed data cont....

- The absolute value of $|\mu_3/\sigma|$ is added to remove the effect of skewness from the data.
- Nahar et al. (2001) modified above index by replacing mod sign to minus sign.
- Chang et al. (2002) proposed weighed standard deviation method to calculate process capability for the skewed data.

Auto correlated data

- The presence of auto-correlation is inherent in the process and can not be removed.
- Jagadeesh and Babu (1994) were the first to find the difficulty in obtaining the process capability in the presence of auto-correlated data.
- Shore (1997) has shown how to calculate the process capability when the auto-correlation is present in the data.
- He showed that autocorrelation affects the sample mean and also the confidence interval.

Auto correlated data cont....

- Noorossana (2002) has shown that the variance estimate obtained from the original data is not appropriate to consider for process capability when the data are auto-correlated.

- The auto-correlation can be removed from the data using multiple regression and time series modeling method.

Some misconceptions concerning PCIs

- So far discussed above is based on the assumption that the data are normally distributed.
- Violation of normality, asymmetric and autocorrelation have also been discussed.
- There are some more notions which should be highlighted for PCIs to serve as useful indices. These notions are:

Some misconceptions cont….

I. The process with higher value of C_{pk} is a better process.
II. If the two processes have the same value of C_{pk}, they are indistinguishable in context of process capability.
III. $(C_p - C_{pk})$ is a measure of scope for improvement.
IV. The optimal value of C_{pk} is desirable to get the optimal quality of process.
V. C_{pk} and C_{pm} are used to detect the target value of process.

1. Higher the C_{pk} better the process

- From the definition of C_p and C_{pk} it is desired to have a large value of these indices.
- Small value of C_p indicates that the natural range of variation does not fit with the tolerance band.
- Here is an example which shows the high C_{pk} is not always desired.

Example- 1

- Consider an orthogonal machine with a cutting tool at a particular rake angle.
- The chip thickness is done by two production lines P1 and P2.
- Upper and lower specifications are
  - USL= 34.472 mm
  - LSL= 34.398 mm
- The other parameters are given in the following table:

Example- 1 cont….

Parameters	P1	P2
μ	34.4350	34.4425
σ	0.0123	0.0100
C_p	1.0027	1.2333
C_{pk}	1.0027	0.9833

Table- 01

Example- 1 cont….

- C_{pk} value is higher for the production line P1.
- So P1 may be considered as a better process.
- Now look at the nonconformance values.
- Nonconformance can be calculated by the formula

$$NC= \varphi[-3(2C_p-C_{pk})]+\varphi[-3C_{pk}]$$

  - The NC for P1= 2629
  - The NC for P2= 1593

Example- 1 cont….

- The nonconformance for P1 is higher although its C_{pk} value is higher.

- Here we conclude that there is an inconsistency between capability index C_{pk} and nonconformance level.

2. C_{pk} are same with different non conformance levels

- C_{pk} index is introduced to account the deviation of process mean from the target value.
- We give an example showing that two processes having same C_{pk} can have the different levels of nonconformance.

Example- 2

- Mild steel is being machined at a particular rake angle. Cutting is done by two methods C1 and C2.
  - USL= 17.80 mm
  - LSL= 17.10 mm
- The comparison of other parameters are

Parameters	μ	σ	C_p	C_{pk}
C1	17.4500	0.0875	1.33	1.33
C2	17.4889	0.0778	1.50	1.33

Table- 02

Example- 2 cont....

- For both the methods, C_{pk} values are same.
- The methods are indistinguishable in terms of process capability.
- If we calculate the nonconformance for both the processes we get
  - Nonconformance for C1= 66
  - Nonconformance for C2= 33

Example- 2 cont....

- This example shows that each process has the same value of C_{pk} but they have the different values of nonconformance.
- There is an inconsistency between capability index C_{pk} and nonconformance levels.

3. (C_p-C_{pk}) is a measure of scope for quality improvement

- C_p and C_{pk} are the measure concern for quality improvement.
- Taguchi (1985-86) suggested an approach to quality improvement by reducing extent of process mean from the target value.
- Tsui (1997) described C_p and C_{pk} may be used to monitor the scope for improvement.
- However this approach is not always true as illustrated by the following example:

Example- 3

- Consider the two processes
  - R1: Machining of valve diameter
  - R2: Machining of valve thickness
- X is a quality characteristic.
- The company has a responsibility to judge which area is needed more improvement.
- Report from the two departments show the following results:

Example- 3 cont….

Parameters	R1	R2
USL	3.2741	3.3513
LSL	3.0765	3.1041
μ	3.1830	3.1973
σ	0.0123	0.0542
C_p	2.6775	0.7601
C_{pk}	2.4688	0.5732
C_p-C_{pk}	0.2087	0.1869

Table- 03

Example- 3 cont....

- Generally if we look for scope of improvement we should take process R1 as it has a higher value of (C_p-C_{pk}).
- Now consider the nonconformance values of the two processes
  - NC for R1= $6.4889*10^{-8}$
  - NC for R2= $4.5001*10^{+4}$

- Therefore from nonconformance analysis we see that process R2 needs more attention for improvement.

4. Maximum value of C_{pk} may not be most appropriate choice

- When the target is not on the midpoint of the specification limits then C_{pk} may not give the exact process capability.
- Bothe (2001) suggested an index C_{pk}* to handle this situation.
- Following example shows that optimal value of C_{pk} may not give the optimal quality of process.

Example- 4

- Consider a part feature with target value 33.
  - USL= 33+7= 40
  - LSL= 33-4= 29
  - Midpoint of the tolerance= 34.5

- Assuming the normal distribution for the process output.
- C_{pk} value is maximum when μ= 34.5
- Unfortunately optimal product performance occurs when μ is located at target 33.

5. C_{pk} may fail to distinguish between on-target and off-target

- C_{pk} index depends upon both μ and σ.
- For fixed value of σ, C_{pk} attains its maximum when target is located on the mid value of the specification limits.
- Following is the other example showing three processes with same values of C_{pk} but different values of C_{pm}.

Example- 5

- The process is turning a cylindrical work piece with three different tools.
  - Target is to get a radius of 19 mm.
  - Upper specification limit= 23 mm
  - Lower specification limit= 15 mm.

- After getting a sample of 5000 units, the result is obtained as following:

Example- 5 cont….

Operation	μ	σ	C_p	C_{pk}	C_{pm}	C_{pmk}
A	19	1.2121	1.10	1.10	1.10	1.10
B	21	0.6060	2.20	1.10	0.64	0.32
C	22	0.3030	4.40	1.10	0.44	0.11

Table- 04

Example- 5 cont….

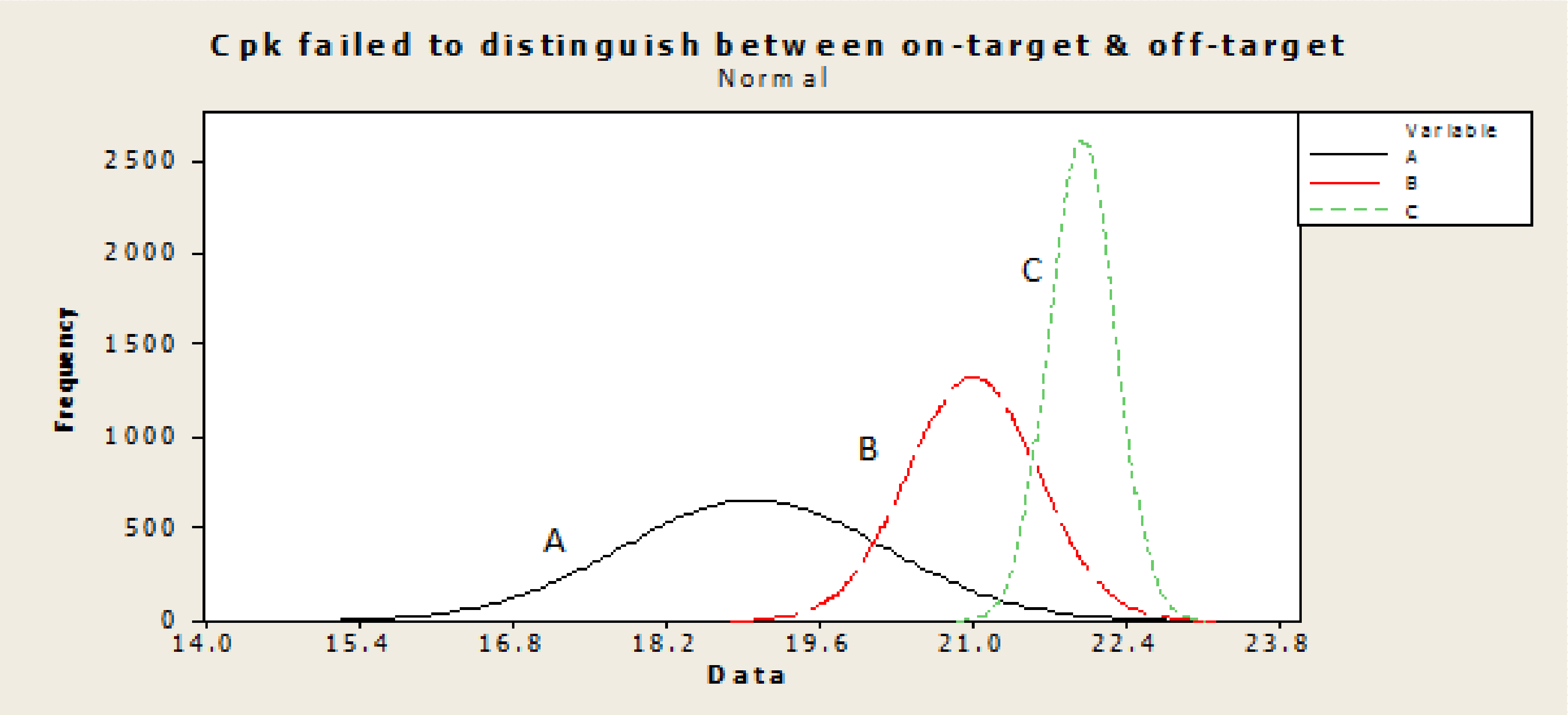

Figure- 01

Example- 5 cont….

- From the above figure we see that process A is on-target process while B & C are off-target processes.
- Although C_{pk} values are same for each of the processes.
- Therefore it is interesting to note that C_{pk} failed to distinguish between on-target and off-target processes.

6. C_{pm} may fail to distinguish between on-target and off-target

- C_{pm} index was introduced by Hsiang and Taguchi in 1985.
- Chan et al. also proposed the same index in 1988.
- It mainly concentrates on the ability of the process to cluster around the target.
- Boyles (1991) has provided a detailed analysis of C_{pm} in measuring the process centering.
- But following is an example which shows that C_{pm} may fail to detect the change in process centering.

Example- 6

- SAE- 1040 steel is being cut by three different tools.
- Target of the process is to get a chip thickness of 20 mm.
  - Upper specification limit= 23 mm
  - Lower specification limit= 17 mm
- After getting a large samples and analyzed them we get the following result.

Example- 6 cont….

Operation	μ	σ	C_p	C_{pk}	C_{pm}	C_{pmk}
A	20.00	0.8000	1.25	1.25	1.25	1.25
B	20.65	0.4664	2.14	1.68	1.25	0.98
C	20.78	0.1776	5.63	4.17	1.25	0.93

Table- 05

Example- 6 cont….

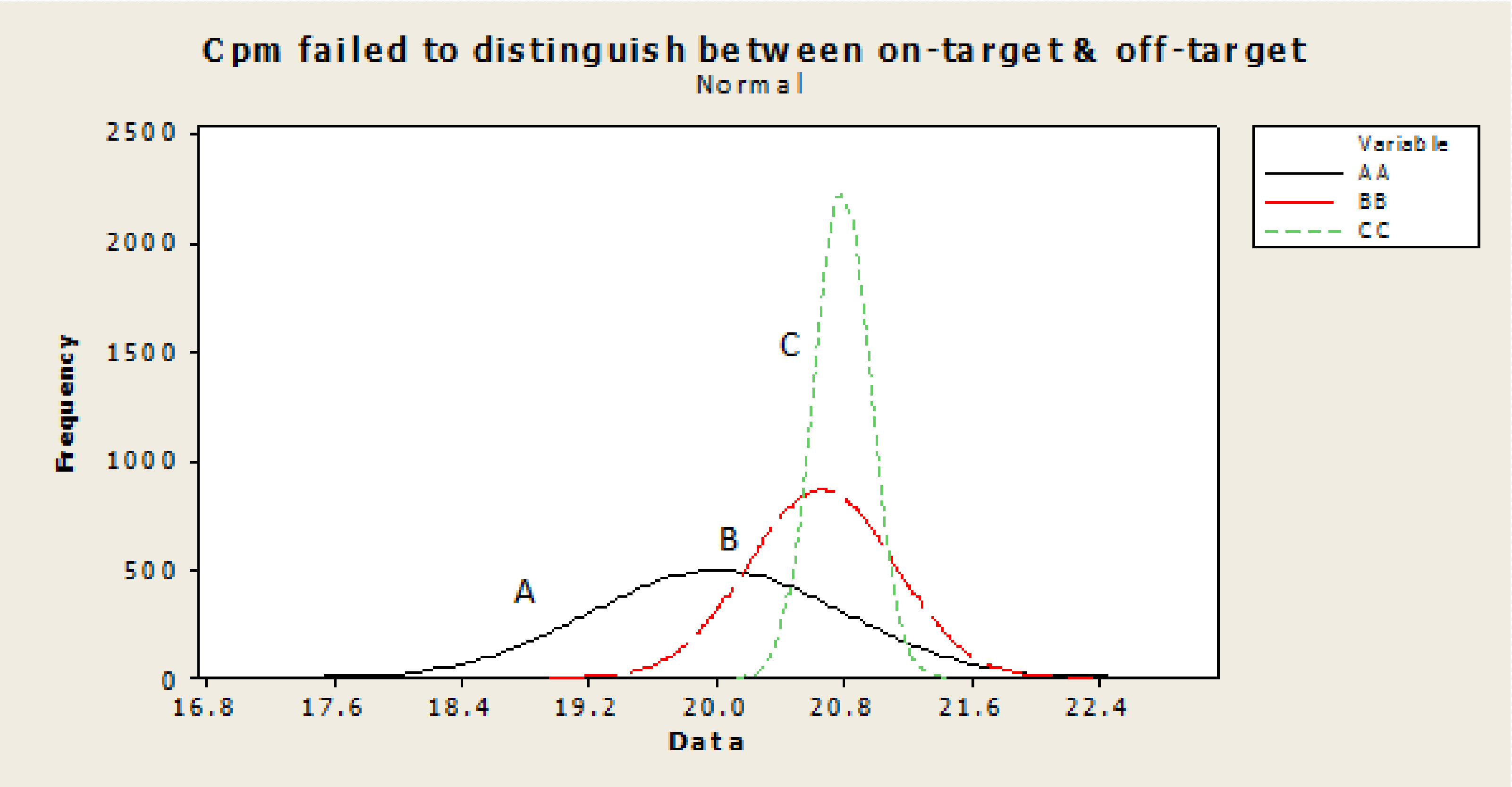

Figure- 02

Example- 6 cont….

- From the above figure we see that process A is on-target while processes B and C are off-target.
- Yet the C_{pm} value for all the three processes are same.
- Therefore it is interesting to note that C_{pm} failed to distinguish between on-target and off target processes.

C_{pk} index – inappropriate for asymmetric tolerances

Example 7
Two production lines for manufacturing steel tubes Quality Characteristic of interest: the inner diameter of the tube.

Specification : $20.23^{+0.03}_{-0.13}\ mm$
Table : Operating Details

Parameter	Line -1	Line - 2
μ	20.215	20.2375
σ	0.0075	0.0075

L = 20.10 U = 20.26 T = 20.23

Table : Calculated Indices

Index	Line -1	Line - 2
C_p	1.3333	1.3333
C_{pk}	0.6667	1.00

Which is the better production line ?

C_p for both line : Identical

C_{pk} for line - 2 > C_{pk} for line – 1

Production line – 2 is better ! ← Temptation

NC for line 1 : 0

NC for line – 2 : 1350 ppm

A better production line cannot produce so many NC. Hence, in this case, interring on the basis of the PCI is erroneous.

Even the pair C_p & C_{pk} together may not be able to discriminate between processes.

<u>Example</u> 7 : Two production lines for manufacturing steel tubes of different dimensions.

Quality Characteristic of Interest : Inner diameter (ID) of the tube.

Table : Operating Details of the two production lines.

Parameter	Line -1	Line - 2
Tube ID (mm)	$20.23^{+0.03}_{-0.13}$	$20.30^{+0.05}_{-0.06}$
L	20.10	20.24
L	20.26	20.35
T	20.23	20.30
μ (= T)	20.23	20.30
σ	0.0135	0.0225

Table : Calculated Indices

Index	Line -1	Line - 2
C_p	0.7407	0.7407
C_{pk}	0.7407	0.7407

Tempting Inference : The two production lines are in distinguishable and either one of them can be chosen for process improvement study.

NC for line 1 : 13134 ppm
NC for line – 16965 ppm

Hence line – 2 should be chosen for Quality improvement study.

Conclusion

- Every single index will suffer from some drawbacks.
- A single index value can not provide the exact figure of the true process capability.
- A process with higher value of C_{pk} can have a higher level of nonconformance.
- Two processes with the same value of C_{pk} can have the different values of nonconformance.
- Maximum value of C_{pk} may not give the optimal quality of the process.

Conclusion cont….

- The indices C_{pk} and C_{pm} may fail to detect the change in process mean.
- The users should be aware about these misconceptions while using the indices to improve the quality of process.

Thank You