A Practical Guide to Utilizing C_p and C_{pk}

Omair A. Khan[1]

[1] West Pharmaceutical Services, R&D Statistical Engineering Intern omair.adil.khan@gmail.com

June 19, 2015

This document explains the use of process capability indices as a way to understand and improve manufacturing processes. It is intended to be an empirical and pragmatic approach to capability analysis without developing the underlying statistical theory. After reading this document, the reader will have a strong grasp on the motivation for process capability indices and what is needed to calculate them, focusing on C_p and C_{pk}.

Why quantify capability?

The ability to manufacture a product within a customer's specifications or tolerances is known as **capability**. **Statistical process control (SPC)** is a methodology for achieving process stability and improving capability through the reduction of variability. In any production process, a certain amount of natural variability will always exist (**chance or common causes of variation**). Occasionally, **assignable or special causes of variation** can be present in the output of a process, arising from improperly adjusted machines, operators, or defective raw material. This type of variability is usually large compared to natural variability and tends to suggest an unacceptable level of process performance. A process that is operating with only chance causes of variation is said to be **in statistical control**. Conversely, a process that is operating under the presence of assignable causes is said to be an **out-of-control process**.

Since process variation can never be totally eliminated, the control of this variation is the key to product quality. Maintaining a stable process average and systematically reducing process variation are the keys to achieving superior quality. If process variation is controlled, then a process becomes predictable. If predictability and consistency are achieved, then a description of the capability of the process to produce acceptable products is possible. A **process capability index** is a statistical measure of process capability. These indices can be used in the following ways:[2]

[2] Deleryd M (1996)

1. As a basis in the improvement process.
2. As an alarm clock.
3. As specifications for investments. By giving specifications for levels of process capability indices, expected to be reached by new machines, the purchasing process is facilitated.
4. As a certificate for customers. The supplier is able to attach the result from the process capability studies conducted when the

actual products were produced, with the delivery.

5. As a basis for new constructions. By knowing the capability of the production processes, the designer knows how to set reasonable specifications in order to make the product manufacturable.
6. For control of maintenance efforts. By continuously conducting process capability studies it is possible to see if some machines are gradually deteriorating.
7. As specifications 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 process.
11. As a base for inspection activities.
12. As a receipt for improvements.
13. For formulating quality improvement programs.

Process capability indices in practice

SPC is primarily a method for monitoring process performance. Many engineers believe that C_{pk} can be used to quantify product quality. This is simply untrue. While C_{pk} can be used to calculate process fallout (Table 1), the decision to accept or reject a production lot of items must be made by **acceptance sampling**. Sampling plans can be derived using a variety of statistical techniques but are commonly chosen by consulting tables outlined in ANSI/ASQ Z1.4 (for attribute data) or ANSI/ASQ Z1.9 (for variables data).

As mentioned above, the main goal of capability analysis is to help reduce variability in the manufacturing process. Higher capability indices generally correspond to higher profits as they imply fewer non-conforming parts and better customer satisfaction. Table 2 contains commonly used minimum values for a variety of processes.

Finally, it is important to understand that C_{pk} does not give us the whole picture. One of the disadvantages of C_{pk} is that it does not take into consideration the target or nominal specification. Figure 1 illustrates how the same C_{pk} value can describe two very different processes. For this reason, it is good practice not to base decisions solely on the numerical value of a statistic, but also to graphically visualize the data. Another way to address this difficulty is to use a process capability index that is a better indicator of centering, such as C_{pm} or C_{pkm}.

C_{pk}	Sigma level	Yield	Fallout
0.33	1	68.27%	317311
0.67	2	95.45%	45500
1.00	3	99.73%	2700
1.33	4	99.99%	63
1.67	5	99.9999%	1
2.00	6	99.9999998%	0.002

Table 1: Relationship between C_{pk} and non-conforming items (measured in PPM).

Situation	Minimum Capabilty
Existing Process	
Regular	1.33
Critical	1.50
New Process	
Regular	1.50
Critical	1.67
Six Sigma Process	2.00

Table 2: Recommended capability values for two-sided specifications.

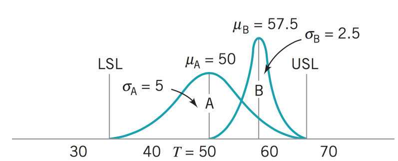

Figure 1: Two processes with $C_{pk} = 1.0$ (Montgomery, 2009).

How to calculate capability indices

While there are numerous process capability indices, the two that are most commonly used in industry are C_p and C_{pk}. These random variables are estimated with the following equations (note the use of the hat to denote the estimate):

$$\hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}}$$
$$\hat{C}_{pk} = \min\left[\frac{USL - \hat{\mu}}{3\hat{\sigma}}, \frac{\hat{\mu} - LSL}{3\hat{\sigma}}\right]$$

where USL and LSL are the upper and lower specification limits given by the customer, $\hat{\sigma}$ is the sample standard deviation (s), and $\hat{\mu}$ is the sample mean ($\bar{X}$). C_p estimates what the process is capable of producing if the process mean were to be centered between the specification limits. Many times, the mean is not exactly centered and C_p overestimates the process capability. C_{pk} more accurately quantifies capability in these cases and is generally used in place of C_p regardless of the location of the process mean. Note that $C_p = C_{pk}$ when the mean is actually centered between the specification limits.

It is recommended that standard deviation be estimated as

$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

rather than $\hat{\sigma} = \bar{R}/d_2$. This second equation is commonly used by Six Sigma practitioners but is less statistically tractable than the first equation.

Verifying assumptions

In order to use C_p and C_{pk} properly, three main assumptions must be verified:

1. The individual data must be normally distributed. Normality can be verified by visually inspecting a Q-Q plot or by using the Anderson-Darling or Shapiro-Wilk tests.
2. The individual data must be independent (a particular observation X_t cannot depend on a previous observation X_{t-1}). Independence can be assumed if a plot of the data against the order it was collected displays no obvious pattern. One can also use the Durbin-Watson test for autocorrelation.
3. The process must be under statistical control, which is verified using Shewart control charts. All data points (or subgroup averages) must fall in between the calculated control limits (not to be confused with customer determined specification limits).

Data that deviates from normality can sometimes be transformed to behave better. In practice, one should instead determine the cause for non-normality if the data is expected to be normal (e.g. dimensional data).

If any of these assumptions is not true, then the process capability indices have absolutely no interpretive value!

Determining how much to sample

Simulations have shown that one must have at least 30 samples in order to estimate C_p or C_{pk}. The exact number needed is dependent on the desired power and the type I error one is willing to tolerate. One must also take into account the length of an operator's shift and the type of manufacturing process to determine the frequency of sampling. **OC (operating characteristic) curves** are typically used for these types of calculations. However, it is usually easier to use the sampling plans tabulated in the aforementioned ANSI/ASQ standards.

Using confidence intervals

Because in practice we must estimate C_{pk} with $\hat{C}_{pk}$, the point estimate is subject to a certain degree of error. If we would like to ensure that our process has a C_{pk} of $c_k = 1.33$, for example, our measured $\hat{C}_{pk}$ must be higher. This value (the lower confidence bound) is a function of the desired C_{pk} (c_k), the sample size (n), and the probability of type I error one is willing to tolerate (usually $\alpha = 0.05$). Table 3 shows a few of these values for a variety of sample sizes.

c_k	**10**	**20**	**30**	**40**	**50**	**75**	**100**	**125**	**150**
1.30	2.29	1.87	1.73	1.66	1.61	1.55	1.51	1.48	1.47
1.40	2.45	2.01	1.86	1.78	1.73	1.66	1.62	1.59	1.58
1.50	2.62	2.14	1.99	1.90	1.85	1.78	1.73	1.71	1.69

Table 3: The minimum value of $\hat{C}_{pk}$ for which the process is considered capable (i.e. $C_{pk} \geq c_k$) 95% of the time. (Adapted from Chou *et al.*, 1990)

These values assume that the data were collected individually. When rational subgrouping is employed, the required minimum value of $\hat{C}_{pk}$ will be less than what is tabulated. The exact calculation is beyond the scope of this document, and the reader is referred to Scholz and Vangel (1998) for more details.

Final words: beware of statistical terrorism

This section is largely adapted from Kotz and Lovelace (1998).

C_p and C_{pk} can be extremely useful when used as part of a more comprehensive capability plan. However, these process capability indices have a high potential to be misused. The result is often an atmosphere of "statistical terrorism" within an organization. Burke *et al.* (1991) define statistical terrorism as "the use (or misuse) of valid statistical techniques along with threats and intimidation to achieve a business objective, even if the objective may be reasonable." Below are some examples of statistical terrorism that Burke *et al.* have outlined:

- *"Bandwagon" terrorism.* Customers require suppliers to commit to implementing SPC aggressively and may even demand a commitment to a deadline date for SPC implementation, after which proof of quality via control charts will be required with each shipment. The result? Vendors ignore the statistical methodology and focus on making attractive charts. The vendors simply won't send out-of-control charts to the customer, since they fear the material will be rejected. In this case, statistical terrorism causes the vendor to lie.

- *"Russian roulette" terrorism.* Vendors are contractually bound to a specific quality criteria, measured statistically with C_p or C_{pk} during a special qualification run. Because of the "random variability" of random variables, which include C_p or C_{pk}, sampling variability may result in a calculated value of C_p or C_{pk} below the specified minimum value, even if the process is truly capable. If only a single estimate of C_p or C_{pk} is required, and the process is exactly capable (say 1.33), there is a 50% chance that the estimate will be below the minimum value. Without including confidence limits, you are playing Russian roulette in terms of meeting their requirements.
- *"Tax audit" terrorism.* The use of standards by large customer companies forces vendors to estimate capability based on their guidelines, which may not be appropriate, for example, with non-normal data. The rigid standards deny the vendors the opportunity to understand their own processes and adjust estimation techniques to match them. The consequences of not meeting the standard may not be made clear, and these standards keep the improvement focus on the products, not the processes. The processes have to be improved in order to improve the products.
- *Other forms of terrorism.* These include "self-inflicted wound" terrorism, which results from extreme pressure that managers place upon their own employees to achieve some statistical goal. There is also "academy award" terrorism, which is the requirement that an organization compete for some renowned quality award, internal or external. Finally, there is the "one true statistician" terrorism, where an organization succumbs to the teachings of a specific individual to the exclusion of any other perspective.

Burke *et al.* (1991) suggest that statistical terrorism may be countered in the same way as physical terrorism: by intelligence, speed, and strength. The vendor should be intimately familiar with what the customer needs in their products (intelligence). Statistical expertise should be developed in-house or be readily available from a qualified outside source, so that quick statistical analysis requests by the customer can be met accurately (speed). Finally, the strength of the quality program comes from knowledge, knowledge of your own processes and how to statistically analyze them in an accurate manner.

References and Recommended Reading

ANSI/ASQ Z1.4–2003 (R2013). Sampling Procedures and Tables for Inspection by Attributes.

ANSI/ASQ Z1.9–2003 (R2013). Sampling Procedures and Tables for Inspection by Variables for Percent Nonconforming.

Chou Y-M, Owen DB, and Borrego SA (1990). Lower confidence limits on process capability indices. *Journal of Quality Technology*, **22**(3), 223-229.

Deleryd M (1996). Process capability studies in theory and practice. *Licentiate thesis*, Lulea University of Technology, Luleå, Sweden.

Kotz S and Lovelace CR (1998). *Process Capability Indices in Theory and Practice.* Arnold, London.

Sholz F and Vangel M (1998). Tolerance Bounds and Cpk Confidence Bounds Under Batch Effects. In Kahle W, *et al.* (Eds.), *Advances in Stochastic Models for Reliability, Quality and Safety* (pp. 361-379). Birkhäuser, Boston.

Montgomery DC (2009). *Introduction to Statistical Quality Control*, 6th edn. John Wiley and Sons, New York.