Cpk: What is its "Capability?"

By: Rick Haynes, Master Black Belt – Smarter Solutions, Inc.

Cpk is one of many capability metrics that are available. When capability metrics are used, organizations typically provide a Cp and a Cpk. In this paper we will discuss the mechanics of these two capability metrics, along with the pros and cons.

In summary, the Cpk can provide insight on performance to a requirement if the process data used in the calculation comes from a normal distribution. If the process data is non-normal or it is the result of a combination of processes (a mixture of processes) then it provides an underestimation of the true non-conformance capability.

Cpk Theory

Cp and Cpk are related by algebra. The k actually represents a correction factor that adjusts the Cp to a Cpk where $C_{pk} = C_p(1-k)$. The *k* is a correction factor for an off-centered process.

$$m = \frac{USL + LSL}{2}$$

So K is defined as:

$$k = \frac{|\, m - \mu \,|}{(USL - LSL)/2}$$

The absolute value of (m-□) provides the same value when the process is above or below the middle of the specification range.

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

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

or

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

Now we have the relationship of Cp and Cpk. A company could then ask for the Cp and the mean of the process in order to find the Cpk value. This is because they should already know the specification range.

You can find this in most text books, but what is missing? The assumptions behind the calculations are:

1) The process that produced the values is stable
2) The values are normally distributed

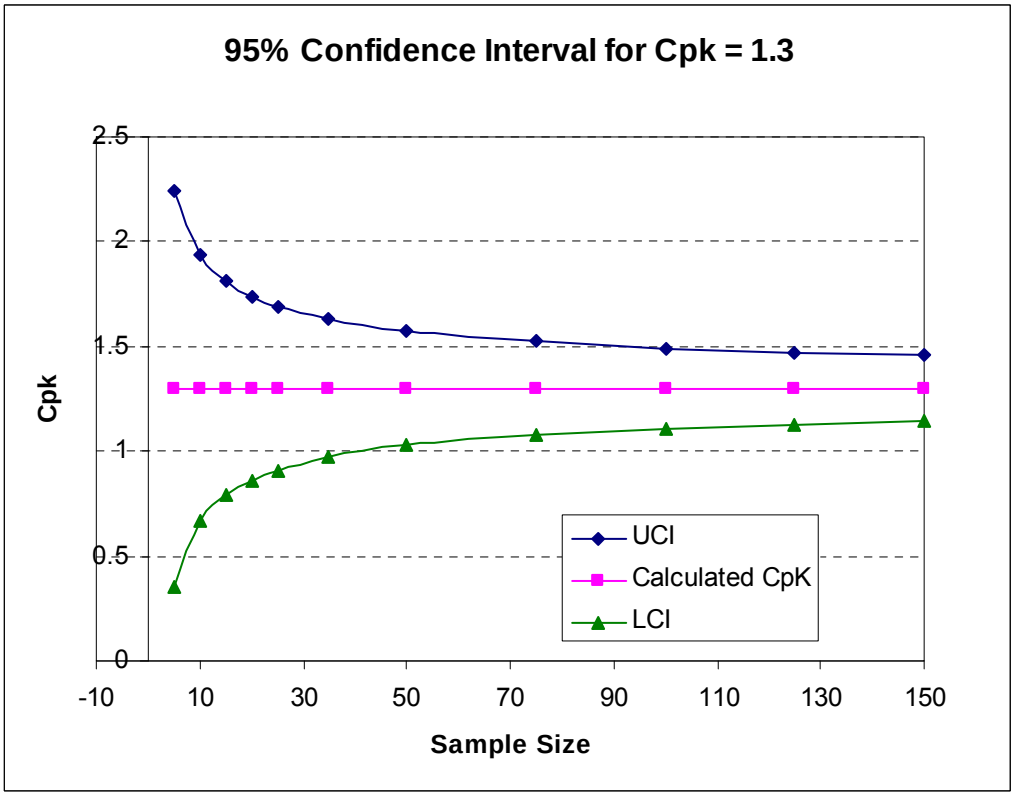
Sounds simple, but most applications of Cp and Cpk do not include a check to see if the assumptions are met. Realize that if you do not meet the assumptions, you are still able to calculate the Cp and Cpk, but they do not provide the expected insight into the process performance.

We will take a look at the impact of the assumptions on the Cpk usage; let us consider the uncertainty of the estimate with the assumptions met. Since Cpk is calculated from averages and standard deviations, you can calculate a confidence interval on every computed Cpk. This estimate is a function of the sample size and the calculated Cpk.

Confidence interval calculation for Cpk assuming a normal population.

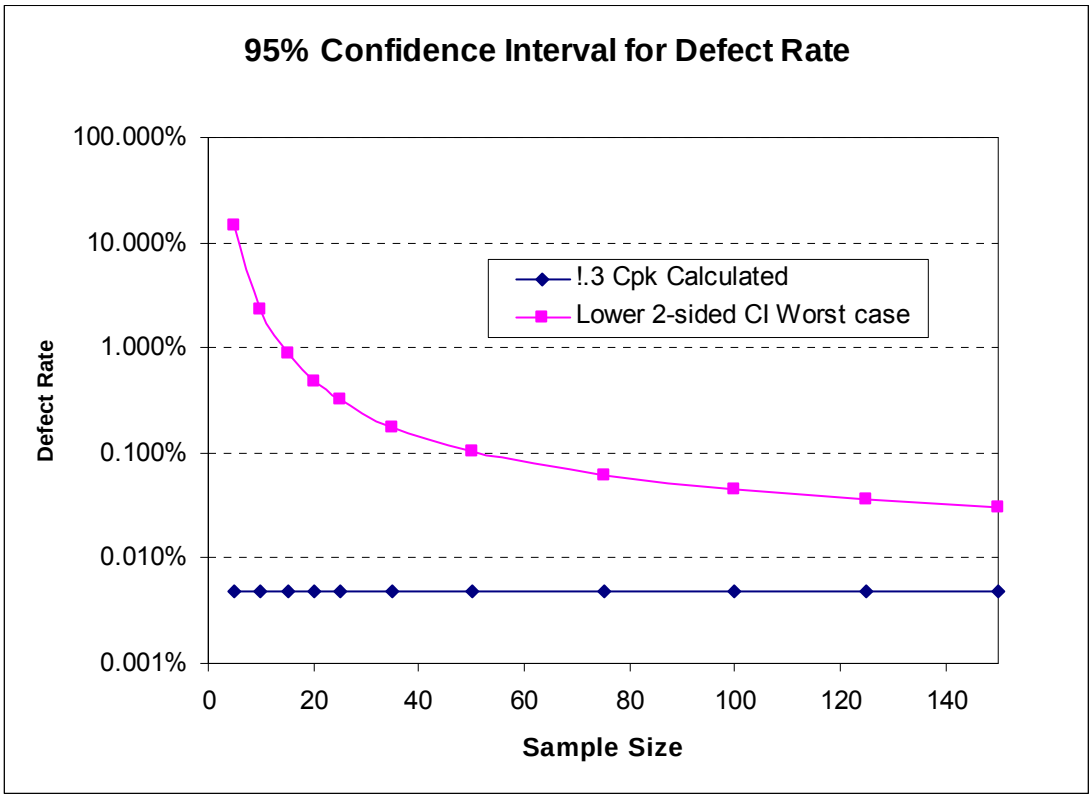
$$C_{pk} = \hat{C}_{pk} \pm z_{1-\alpha/2}\sqrt{\frac{1}{9n} + \frac{\hat{C}_{pk}^2}{2(n-1)}}$$

The derivation of this would be quite difficult, but we can calculate the simple intervals for a given Cpk.



Look how a long-term Cpk of 1.0 is possible when a 1.3 is calculated even up to a sample size of 40!

Now to convert this Cpk uncertainty to an estimated true defect rate of the evaluated material, we converted the lower confidence interval Cpk into a percent defective that would be estimated if the data is normally distributed along with an assumption that all of the defectives are found on only one side of the specification.



We do not always consider the uncertainty of calculated statistics, but it always exists.

Now that we recognize that all of the following discussions have uncertainty around the conclusions, we are ready to look at the Cpk capability to provide insight with real data sets. I am going to consider that Real Data sets do not always have normally distributed data. Real data may have a mixture of sub-processes contributing to the total performance.

Cpk Impact due to Mixture Distributions

The process produces data that derive from a mixture of distribution, such as parts off of three different manufacturing lines, or claims processed at three different claim centers, or seasonal sales volumes. In each of these cases, the customer sees only a single output no matter what the source may be, but in the organizational view, each manufacturing line provides slightly different products, the claim centers have different processing times, or there is a strong seasonal change in lead time. Most real Cpk applications are assessing process data derived from a mixture, and in these cases the Cpk will underestimate the non-conformance rate.

We will evaluate the impact of the non-normality on Cpk by using a Monte-Carlo simulation software, @Risk. We will define a sample of 25 and 100 that will be a 50:50 combination of two normal distributions that will have increasing differences in the mean.

The simulation considers two base distributions with a mean of 0 and a standard deviation of 1.0. Then a shift is introduced between the distributions. Five shifts are evaluated, {0.0,0.25, 0.5, 1.0, 1.5,2.0}, The shift is applied to each distribution but in opposite directions. For example, when the shift is 0.25, the two distribution means are -0.25 and 0.25 which lead to a 0.5 difference in the means. A two-sided specification is applied to the data, -0.2 +/- 3 or -3.2 to 2.8. The process data is not perfectly centered in the specification in order to model a typical process.

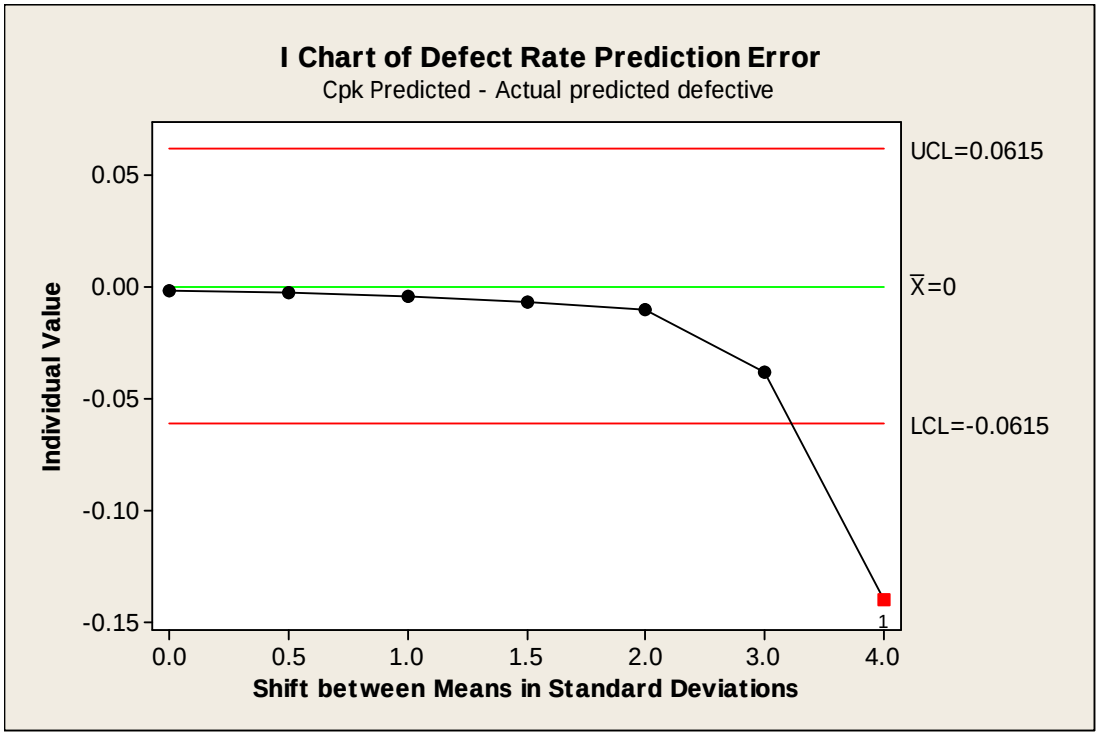
With each mean shift, a simulation using 1000 iterations is run. The output from each iteration has the following values calculated:

- Mean
- Standard deviation
- Predicted out-of-specification %
- Actual % out-of-specification (by counting data points)
- Cp
- Cpk
- Estimated out-of-specification based on Cpk Value

The last value, “Estimated out-of-specification based on Cpk Value,” is made assuming a normal distribution with the calculated mean and standard deviation from that sample. The Cpk Value, is the value time 3.0, is the number of standard deviations that the specification is separated from the mean of the distribution. This is used to calculate the percent out of specification for that specification limit. We will compare estimated defect percentage to the actual percent defective.

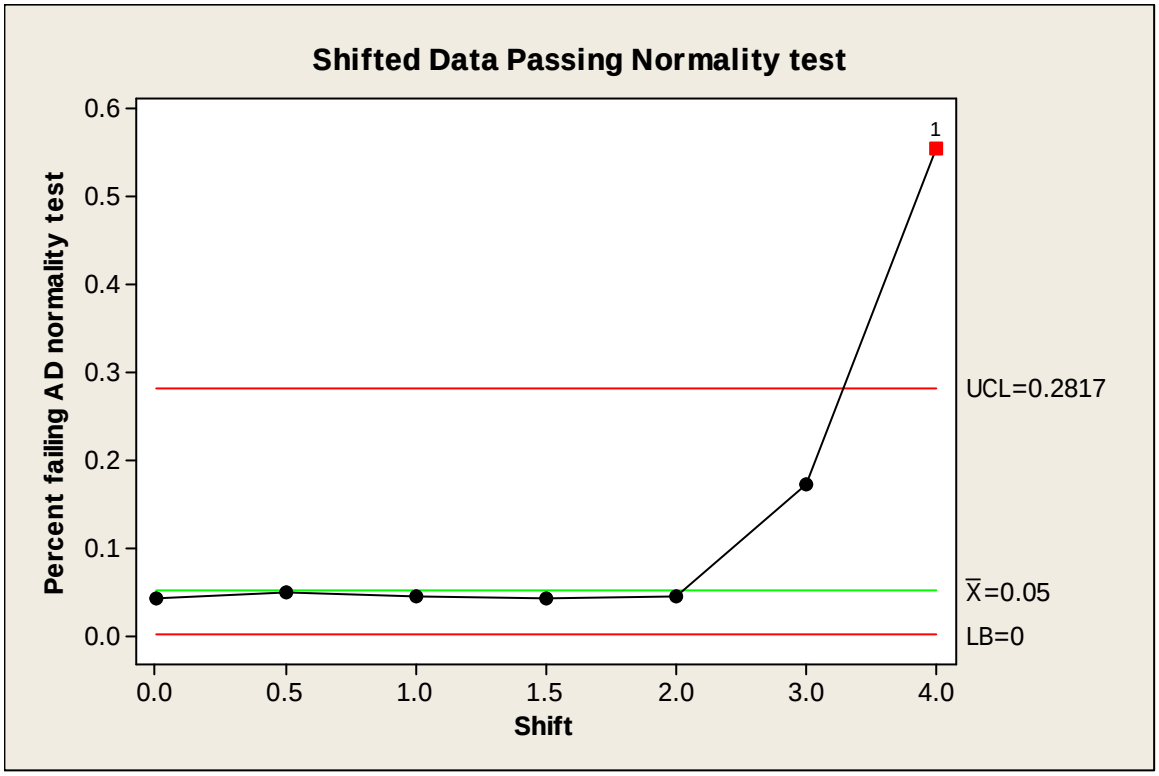
There are two ways to look at the results: how did the average perform, and how did the variation in the results impact the decision?

Starting with the average performance, we will only consider the predicted out-of-specification values. We will look at the result by comparing the estimate that the Cpk would predict, by assuming the population was normally distributed and using the actual percent defective would be, based on the true mixture distribution. In this analysis, the Cpk interpretation stayed close until the shift in the means exceeded two standard deviations. After that point, the Cpk dramatically underestimates the true non-conformance situation.



These values are based on the average response from the simulation. Earlier in this article we discussed the uncertainty of the Cpk when using small samples, so that this is just a typical result.

The data in the simulation used a mixture distribution with a mean of zero with the data points equally distributed between two normal distributions with standard deviations of 1.0. These two distributions are shifted apart by different multiples of the standard deviation. The deviation shown in the prior chart is primarily driven by the lack of normality in the combined simulation data. An analysis of each simulation run using the Anderson Darling Goodness of Fit test for normality of the combined data step also showed that the data began to fail the test at a higher rate as the shift went from 2 standard deviations to 3 standard deviations. This is demonstrated by accumulating the p-value for each random set of data in the simulation. If the data are truly normally distributed, we would expect to have 5% of the p-values to be rejected if we consider a 95% confidence in the assessment. To evaluate this hypothesis, the AD p-value is collected with each run of the simulation, and then the % rejecting the normality test should be about 5% , if I am considering a 95% confidence. The following chart shows that there is a significant increase in the rejection of normality as the shift exceeds 2.0 sigma.



This was an interesting result, in that when the samples pass the normality test, the Cpk continues to be a reasonable measure of capability. When normality begins to break down the Cpk begins to underestimate the percent defective. If you notice, there were no absolutes in these statements. Consider the 3.0 sigma shift, only around 20% of the samples failed normality but the early chart still showed an underestimation of the percent defective. Therefore, passing a normality test is a good thought, but it does not fully protect your Cpk Estimated capability.

Cpk Impact due to Homogeneous Non-Normal Distributions.

How would the Cpk respond to non-normal distributions? Non-normal data exist in many areas of business. Lognormal data are found when there is a natural limit, such as time measurements. This type of data also exists whenever quality is inspected into a product or when it leads to a removal of the long tails of the distribution. Lognormal distributed data can pass a normality test if the values are far from the exiting natural limit. The Cpk responds identically to the random normal distribution as the means between the two distributions were increased. The underestimation of the defect rate becomes severe as the shift exceeded 2.0 standard deviations.

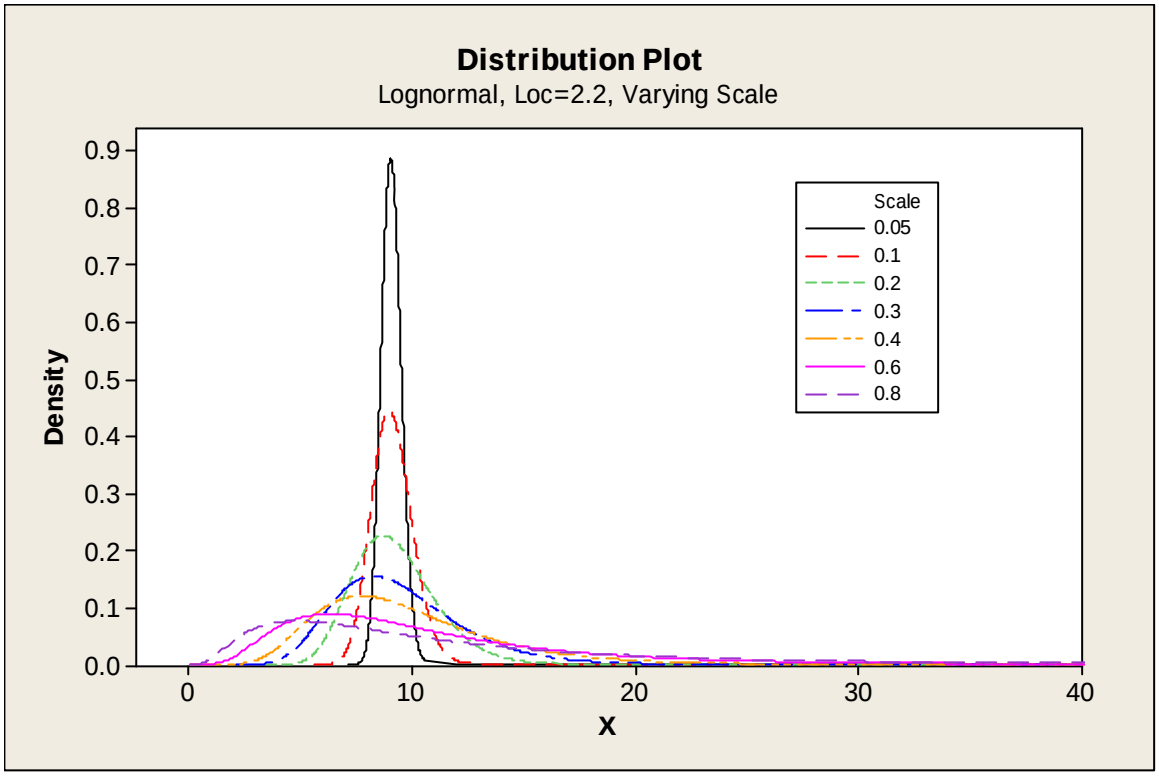
With the assessment of normal and lognormal mixture distributions, it can be generalized to say that the Cpk is a reasonable capability assessment tool as long as the data set passes an normal probability goodness of fit test.

The use of Cpk with non-normal process data occurs when you are working to comply with lead time requirements or on-time delivery requirements. In these cases, the Cpk will underestimate the non-conforming rate as the process data diverge from normality. The risk in this case is also

increased with smaller sample sizes because small samples may pass a normality test at a relatively high rate even when the true process data is not normally distributed.

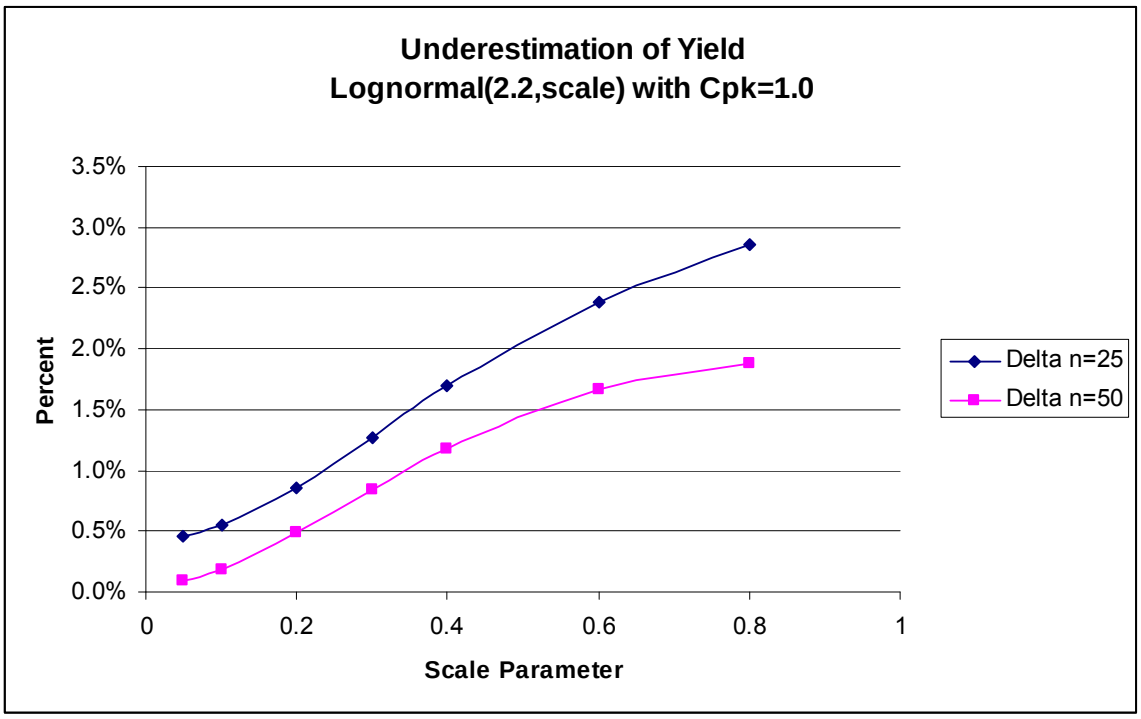
For this evaluation we will use a lognormal distribution, since it is the most common non-normal distribution found in business. Almost any data that derives from a duration of time measurement, such as a lead time or cycle time data will have a lognormal distribution.

The examination of non-normality will use a lognormal distribution with a constant location parameter with an increasing scale parameter which will drive the skewness.

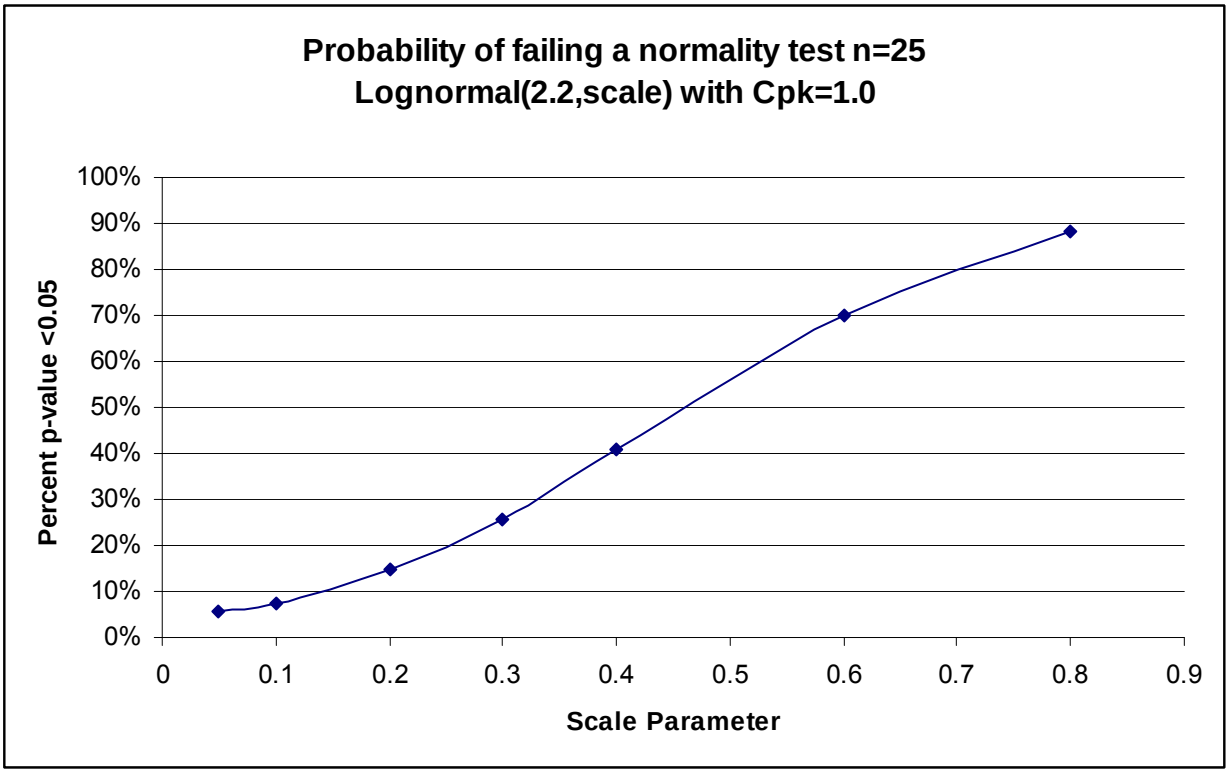


By inspection, you may see that the two smallest scale values produced curves that appear symmetric and possibly normally shaped.

The impact of the non-normality will be assessed from each of the distributions, assuming the specifications were set to have a Cp=1.0 and a Cpk=1.0. In this case the expected out-of-specification rate should be .0027 or 0.27% , the probability of being outside of the +/- 3 sigma limits for a normal distribution.



This result shows that the Cpk underestimates the yield loss when it computes to a 1.0 value when the data is increasingly skewed. Of course this is based on a very large sample using a simulation, but most applications of Cpk only use small samples. Examining small samples (n-25) for the same lognormal distributions show us that they may pass the normality test quite frequently, meaning that there is risk in underestimating yields even if the sample passes a normality test.



Conclusion:

The Cpk statistic is used as a single reference to describe the capability to meet a requirement. What is not taught at the same time is the assumptions behind the statistic that allow you to gain process insight from this statistic. To be a valid predictor of non-conformance, a large sample size is required, and the process data need to be normally distributed. Without both of these there will be an underestimation of the true non-conforming rate.

Now if Cpk is only used as a relative measure, where the goal is to move the value higher to have a better process, there may be some benefit to its use. However, if the goal is to reach a given Cpk to ensure a high quality process, it may provide a false sense of success.