Capability Analysis with Non-Normal Data

by Keith M. Bower

To assess the ability of a process to meet specifications, Six Sigma practitioners may be required to perform a capability study. In many analyses the assumption of sampling from a Normal (Gaussian) distribution may be invalid. This article addresses strategies to obtain valid capability estimates using *continuous* data under such a scenario.[1]

Process Capability Analyses

To assess the ability of a process to meet specifications for future production, capability estimates such as $\hat{C}_p$ and $\hat{C}_{pk}$ are widely employed. Associated confidence intervals may be useful as they take into consideration the amount of data used in the study.[2] Invalidly assuming a Normal distribution as an adequate model for the process can result in seriously misleading capability estimates.[3]

Evidence of Stability

A key criterion for obtaining valid estimates of future productive capability is the assumption of sampling from a stable (constant-cause) system. A probability distribution (Normal or non-Normal) may then be usefully employed.

The requirement for stability is essential; we cannot validly estimate future productive capability if there is no evidence of current stability. That is why control charts are an integral element of capability analyses.

At the start of a Six Sigma project the assessed process may be unstable and may not adequately meet specifications. Capability estimates then computed have little or no theoretical grounding – though would merely indicate the "obvious," that the process requires improvement. This part of the Six Sigma project would occur in Phase I of a control charting scheme, when the analyst is attempting to understand the system, possibly for the first time.[4]

This article considers the time frame during Phase II of control charting. The process has provided evidence of stability, and a mathematical model for the process may be meaningfully sought. The statistical tools employed in a capability study, for example confidence intervals, which are based on *enumerative* principles, may have legitimacy for this type of *analytic* study when the process is stable.[5]

The Modeling of Processes

Generally speaking, control charting procedures are designed to allow for the identification of uneconomic situations. The underlying probability distribution of the process is typically not addressed.[6] One reason is that, as discussed by Walter Shewhart, Tchebycheff's Theorem suggests that the probability of falling beyond three standard deviations from the mean will be "small" for many distributions, including, of course, the Normal distribution.[7]

A major concern with capability analyses, however, is that a good approximation of the underlying distribution is required for valid estimates of future nonconforming product. The quality practitioner must consider both the recorded measurements and the amount of data collected to assess process behavior and consider an appropriate distribution.

In "Guiding the Use of Capability Estimates – Keep it Simple!" Terry Weight suggests that a practitioner will never be in a position to state a mathematical model that would generate "exact" proportions nonconforming.[8]

It is of course true that the underlying model will never be exactly known in practice. However, this should not prevent practitioners from attempting to understand the witnessed distribution and make intelligent estimates of the distributional form (at least, in Phase II). Indeed, a useful model for a given process may be suggested by scientific theory.

The Use of "Raw" Nonconforming Proportions

In his article, Mr. Weight suggests adopting the approach of reporting solely the "raw" proportion nonconforming from some indeterminate length of study. This argument seems more appropriate for Phase I analyses as opposed to the Phase II analyses here under discussion.

In particular, a key argument against the "raw" proportion methodology concerns the amount of data used for assessment. Reporting the "raw" proportion comes into question when one acknowledges the considerable widths of confidence intervals for capability indices, even when relatively large amounts of data are used. Employing confidence intervals in capability analysis reports thus becomes all the more appropriate, rather than assuming the point estimate of the proportion nonconforming is the "correct" answer.

Non-Normal Distributions

This discussion proposes two key assumptions:

(a) The assessed process exhibits stability

(b) A non-Normal distribution is a sensible fit

If criteria (a) and (b) are met, several options may be available, including:

(1) Transforming the dataset to become approximately Normal using approaches such as the Box-Cox[9] and Johnson[10] procedures, widely available in statistical software packages

(2) Identifying a non-Normal distribution that is a useful fit

For an illustration of approach (1) consider the following example.

Example

A toothpaste manufacturer is required to fill its tubes to a set mass (115g). At the start of the cycle the tubes are consistently under-filled and then approximately meet 115g after a short while. Assume that this non-Normal distribution is both acknowledged and well understood, and the distributional form witnessed can be expected on a consistent basis; that is, this process has exhibited "stability."

As shown in Figures 1 and 2, the assumption of Normality would be unreasonable owing to the non-linear Normal probability plot, in conjunction with the Anderson-Darling test for Normality exhibiting a very low P-value. Clearly, the distribution is negatively skewed.

Figure 1

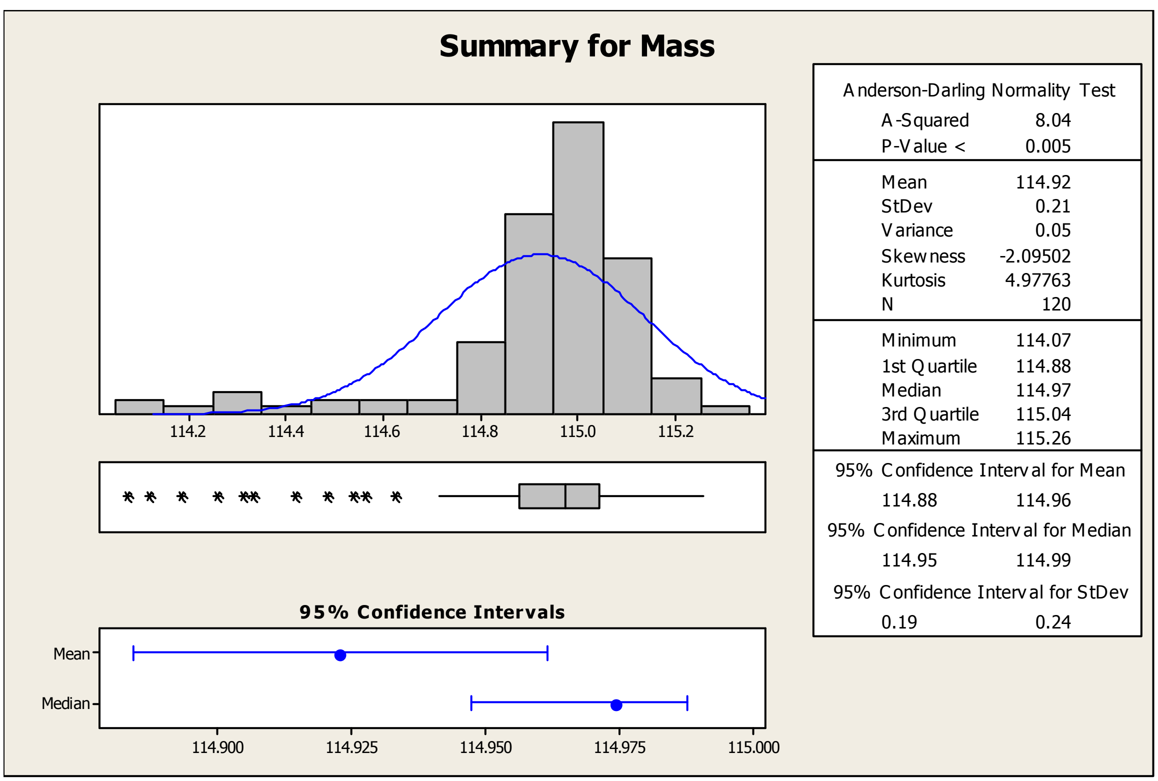
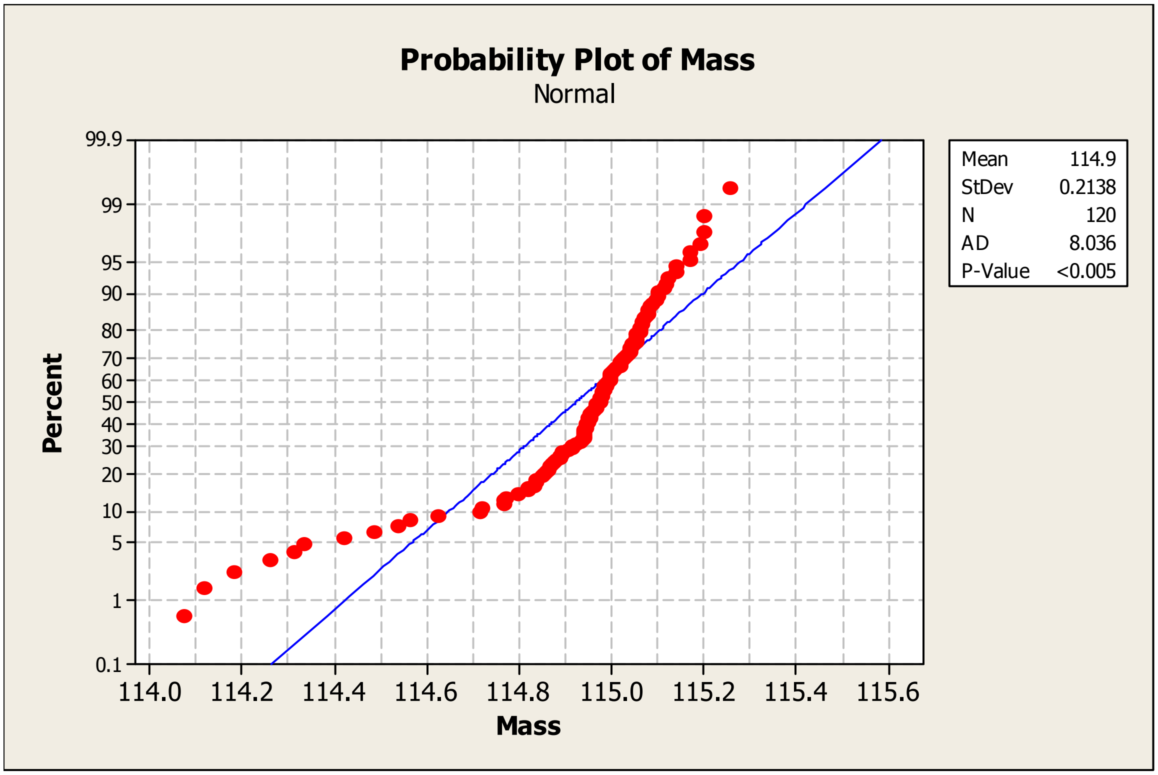


Figure 2



For these data the Box-Cox procedure does not find an adequate transformation to Normality. In addition, none of the frequently used models, such as Weibull, provides an adequate fit to the data. However, the Johnson procedure does provide an adequate transformation, as indicated in the results of Figure 3.

As shown in Figures 3 and 4, the actual model used to transform the toothpaste data into an approximate Normal distribution is rather complicated; note that it includes the hyperbolic arcsine function. Crucially, however, it appears to be a useful transformation to obtain approximately Normal results.

Using 114.9g and 115.1g as the lower and upper specification limits, respectively, we obtain:

(i) The estimate of Pp is 0.31, with lower and upper 95% confidence limits of 0.27 and 0.35, respectively.

(ii) The estimate of Ppk is 0.17 with lower and upper confidence limits of 0.11 and 0.24, respectively.[11]

Figure 3

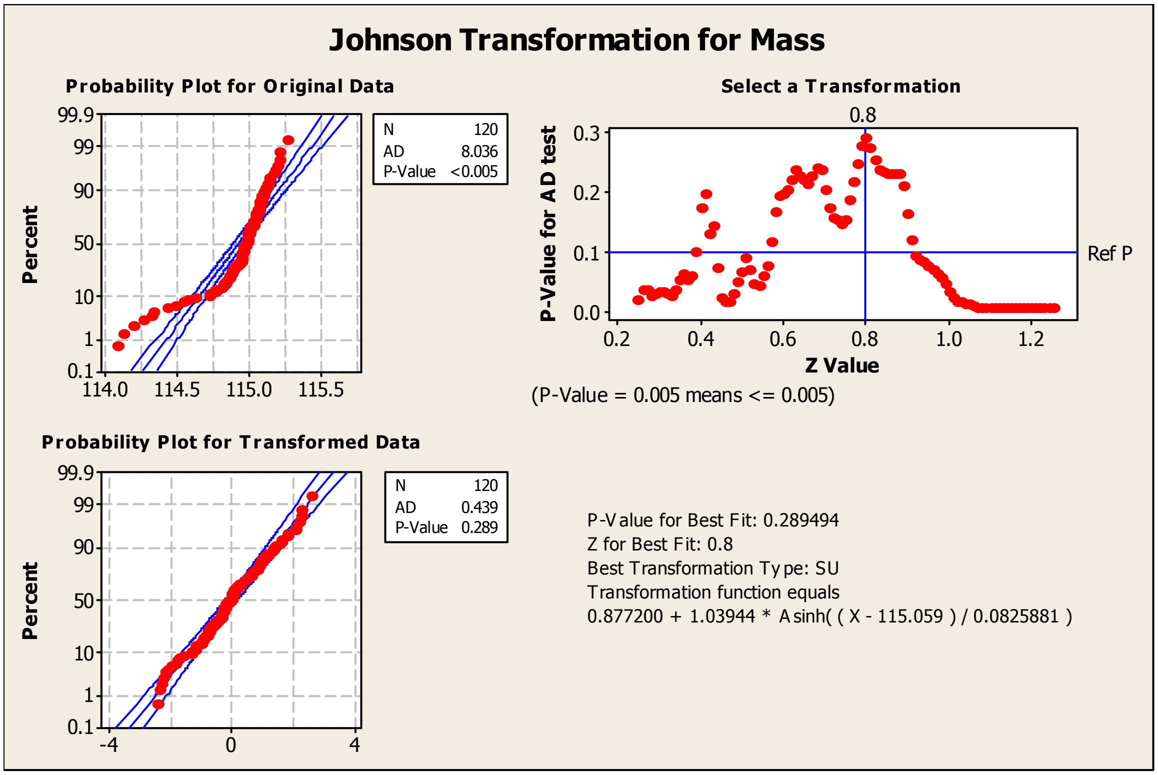
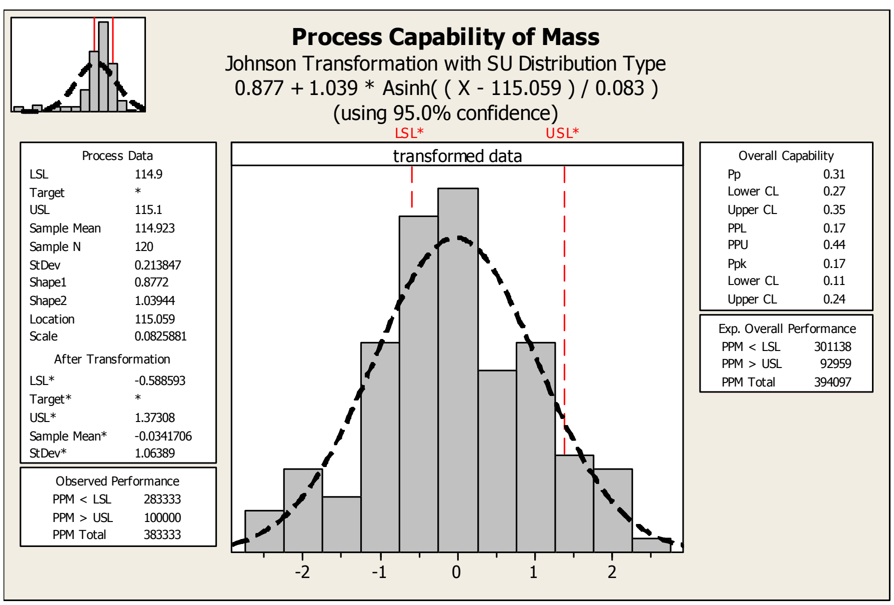


Figure 4



Summary

Performing capability analyses with Normality assumed incorrectly may result in highly misleading capability estimates. To obtain meaningful estimates when a process exhibits stability but the distribution may not be adequately modeled by a Normal distribution, alternative strategies require consideration. This article has provided some general guidelines for subsequent analyses.

References

1. For a description of continuous data see Keith M. Bower, "Non-Traditional MSA with Continuous Data," *ASQ Six Sigma Forum*, Dec. 2003, http://www.asq.org/forums/sixsigma/articles/mbb/mb_non_trad_msa1.html.
2. For information on confidence intervals for capability indices, see Keith M. Bower, "Confidence Intervals for Capability Indices," *International Society of Six Sigma Professionals: EXTRAOrdinary Sense* 2, no. 4 (2001): 6-7.
3. For information regarding the effect on widely used capability indices when Normality is incorrectly assumed, see Steven E. Somerville and Douglas C. Montgomery, "Process Capability Estimates and Non-Normal Distributions," *Quality Engineering* 9, no. 2 (1996): 305-316.
4. For more information on Phase I and Phase II studies, see William H. Woodall, "Controversies and Contradictions in Statistical Process Control," *Journal of Quality Technology* 32, no. 4 (2000): 341-350.
5. For a discussion of enumerative and analytic studies, see Keith M. Bower, "Why Divide By n-1?" *ASQ Six Sigma Forum*, Feb. 2005, http://www.asq.org/forums/sixsigma/articles/bb/bb_divide_by_n-1.html.
6. For more information on distributional assumptions and control charting, especially in Phase II, see Douglas C. Montgomery, *An Introduction to Statistical Quality Control*, 5th ed. (New Jersey: John Wiley & Sons, Inc., 2004): 237.
7. Walter A. Shewhart, *Economic Control of Quality of Manufactured Product* (New York: D. Van Nostrand Company, Inc., 1931): 176.
8. The *ASQ Six Sigma Forum* published Terry Weight's article on June 5, 2003. See the section "Failure to Achieve the Ideal Situation."
9. George E. P. Box and David R. Cox, "An Analysis of Transformations," *Journal of the Royal Statistical Society*, B, 26 (1964): 211-243.
10. Norman L. Johnson, "Systems of Frequency Curves Generated by Methods of Translation," *Biometrika*, 36 (1949): 149-176.
11. For more information on Pp and Ppk and the relation to Cp and Cpk, see Douglas C. Montgomery, *An Introduction to Statistical Quality Control*, 5th ed. (New Jersey: John Wiley & Sons, Inc., 2004): 349. Note in this example that Pp and Ppk do not have the usual interpretation (i.e., using the "overall" standard deviation). This is because the Johnson transformation is transforming each observation and in this case computing "equivalent" capability values.

About the Author

Keith M. Bower is a statistician and webmaster for www.KeithBower.com, a site devoted to providing access to online learning materials for quality improvement using statistical methods. He received a bachelor's degree in mathematics with economics from Strathclyde University in Great Britain and a master's degree in quality management and productivity from the University of Iowa in Iowa City, USA. He is a member of ASQ and the Six Sigma Forum.