

The Six-Sigma Zone

Donald J. Wheeler

The objective of having a process operate with a capability index of 1.5 to 2.0 is a reasonable and economic objective. However, the arguments commonly used to explain the economics of this objective can only be described as statistical snake oil—a blend of tortured computations and incompatible, highly questionable assumptions, having a hypnotic effect and often resulting in a suspension of critical thinking.

Therefore, since the conclusion is right, but the argument used to support the conclusion is complete nonsense, this paper begins with a theoretically sound and rigorous justification for the economics of operating with capabilities in the 1.5 to 2.0 range. Furthermore, this justification allows you to convert capabilities into the one metric that all managers understand: the excess cost of production. This justification is more rigorous, more comprehensible, and more useful than the common arguments involving parts-per-million nonconforming.

Finally, in the interests of exorcism, this paper will conclude with an explanation of the fallacies contained in the common parts-per-million argument.

1. A Model for Excess Costs

The first step in an economic justification for operating with capabilities in excess of 1.5 is to develop a model for excess costs of production. So we begin with the three points with known costs.

If everything were always exactly on-target, and each part was exactly like the next part, then we should find that everything would fit together exactly, and assemblies would work exactly as intended. Measurements would not be needed, inspection would be a thing of the past, shutdowns would be unheard of, efficiencies would all be 100%, and life would be lovely. *And there would be no excess costs of production.* Thus, the first point on our excess cost curve will be a zero at the target.

Next consider what happens when we scrap an item. We have paid to produce the item, we have consumed resources, capacity, time, and effort, and now we have to throw all of this away, losing not only the item, but also the opportunity to profit from the item. Many of these real costs are easy to determine, and so we usually have a cost that we can attach to scrap. Thus, for all product values beyond the scrap point, there will be a fixed cost of scrap.

And what about rework? When an item is reworked we attempt to salvage the raw materials, resources, and effort contained in that item and to recover the opportunity to profit from that item by adding more effort and resources. Since this will only make sense when it is cheaper to rework the item than to scrap it, the cost of scrap will be an upper bound on the cost of rework. Thus, for all product values beyond the rework point, there will be a fixed cost of rework.

This means that we know the excess costs of production at three points, scrap, on-target, and rework.

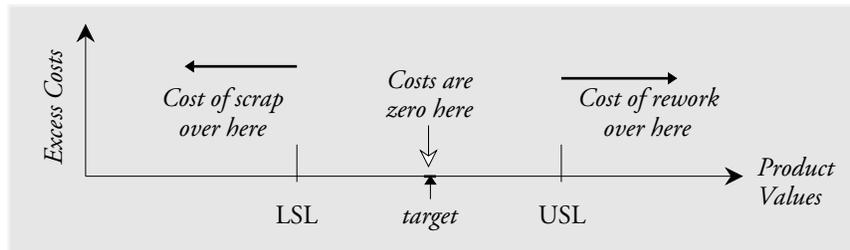


Figure 1: Excess Costs are Known for Three Conditions

The next step in developing a model for excess costs is to connect the three known excess costs. A traditional model has been a step-function where the zero cost was used for all values in between the scrap point and the rework point. As long as the part is inside the specifications the producer does not see any excess costs, but as soon as the part goes outside the specifications the producer has to face the excess costs of scrap or rework.

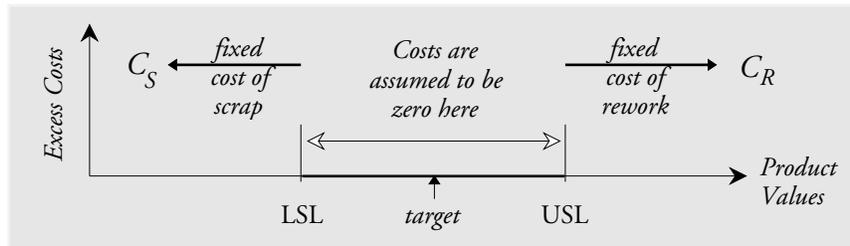


Figure 2: The Traditional Model for Excess Costs

Unfortunately, the step-function model for excess costs in Figure 2 is incomplete. While it captures the producer’s perspective it completely ignores the consumer’s excess costs—the costs associated with using the conforming product. Think about the actions you have to take in order to use the stuff that your supplier sends to you—the adjustments that have to be made, the scrap and rework that you suffer, the testing, sorting, and complex control strategies that you have to use because all of your incoming material is not exactly on-target. These costs are costs of using conforming product, and they are usually substantial.

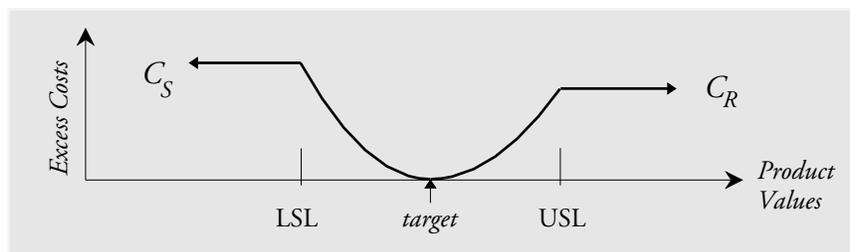


Figure 3: A More Realistic Model for Excess Costs

If you add to these costs of using conforming product the fact that your supplier’s scrap and rework

costs will always get factored into their overhead costs, and that they have to recover their overhead costs from their customers in order to stay in business, then perhaps you can see that the traditional step-function model for excess costs does not begin to tell the whole story. From the perspective of the customer, excess costs need a continuous model, and the simplest way to connect three points with a single continuous function is to use a quadratic curve.

Hence, a more realistic, macroeconomic model for the excess costs will look like the curve in Figure 3. Beyond the scrap point there is a fixed cost for scrap. Between the scrap point and the target there is a quadratic excess cost curve. Between the target and the rework point there is a second quadratic excess cost curve, and beyond the rework point there is a fixed cost of rework. This simple model provides a reasonable first order approximation to any realistic excess cost function. As the parts deviate from the target there are excess costs. The greater the deviation the greater the excess costs, until finally it is cheaper for the customer to have the part scrapped or reworked than it is to try to use it. This simple model incorporates the cost of scrap, the cost of rework, and the costs of using conforming product that is off-target.

While we usually think of scrap and rework costs as showing up at production, and the costs of using conforming product as showing up when the customer tries to use the product, the simple curve in Figure 3 provides the macroeconomic perspective needed for a realistic understanding of the consequences of our actions.

2. The Effective Cost of Production

The excess cost function defines an excess cost for each possible product value. A probability model, $f(x)$, can be used to define the frequency with which each possible product value will occur. And so the excess costs can be written as the integral of the product of the excess cost function and the probability model, $f(x)$.

$$\begin{aligned} \text{Excess Costs} &= \int \text{excess cost function} \times \text{probability model} \\ &= \int \mathcal{L}(x) f(x) dx \end{aligned}$$

Since our excess cost function in Figure 3 comes in four parts the expression above separates out into four separate integrals as shown below. There C_S will represent the cost of scrap and C_R will be the cost of rework.

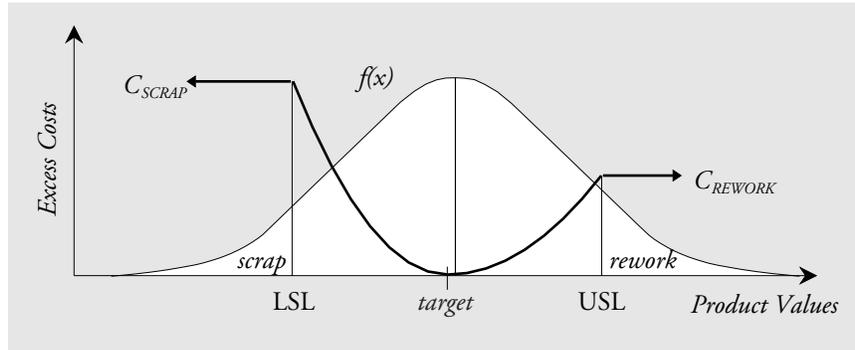


Figure 4: The Excess Cost Function and a Probability Model for Outcomes

$$\begin{aligned}
 \text{Excess Costs} &= C_S \int_{-\infty}^{\text{LSL}} f(x) dx && \text{excess costs due to scrap} \\
 &+ C_S \int_{\text{LSL}}^{\text{target}} \frac{(x - \text{target})^2}{(\text{LSL} - \text{target})^2} f(x) dx && \text{due to deviations below target} \\
 &+ C_R \int_{\text{target}}^{\text{USL}} \frac{(x - \text{target})^2}{(\text{USL} - \text{target})^2} f(x) dx && \text{due to deviations above target} \\
 &+ C_R \int_{\text{USL}}^{\infty} f(x) dx && \text{excess costs due to rework} \\
 &= C_S \text{ ISP} + C_S \text{ IBT} + C_R \text{ IAT} + C_R \text{ IRP} \\
 &= C_S \left\{ \text{ISP} + \text{IBT} + \frac{C_R}{C_S} [\text{IAT} + \text{IRP}] \right\} \\
 &= C_S \times \text{Curly Brackets}
 \end{aligned}$$

where ISP represents the integral for scrap portion, IBT is the integral below the target, IAT is the integral above the target, and IRP is the integral for the rework portion.

If we assume that the Cost of Scrap is approximately the same as the nominal cost per unit, we end up with:

$$\text{Excess Cost} = \text{Nominal Cost per Unit} \times \text{Curly Brackets}$$

The Excess Cost defines the amount by which the actual cost will exceed the nominal cost. In fact, if we add the [cost of the units shipped] to the Excess Cost, and divide by the number of units shipped, we get the Actual Cost of Production (ACP):

$$\text{Actual Cost of Production} = \frac{\text{nominal cost of units shipped} + \text{excess costs of units produced}}{\text{units shipped}}$$

With some rearranging this becomes:

$$\begin{aligned} \text{Actual Cost of Production} &= \text{nominal cost per unit} \times \frac{1.0 - \text{proportion scrapped} + \text{Curly Brackets}}{1.0 - \text{proportion scrapped}} \\ &= \text{nominal cost per unit} \times \text{Effective Cost of Production} \end{aligned}$$

The Effective Cost of Production combines the excess costs of fabrication and the excess costs of using conforming product into a simple multiplier which can be used to characterize the excess costs associated with a particular characteristic.

Inspection of the formulas above will show that the Effective Cost of Production will depend upon three things:

1. The relationship between the distribution of process outcomes $f(x)$ and the specifications, which is fully characterized by the capability indexes, C_p and C_{pk} .
2. The ratio of the cost of rework to the cost of scrap, which will generally be known for any specific situation.
3. Whether the process mean is on the rework side or the scrap side of the target value.

Extensive tables of the Effective Cost of Production are available in *The Process Evaluation Handbook* by this author. In the next section we shall use one of these tables to look at the relationship between the Effective Cost of Production and the Capability Indexes in one situation.

3. The Effective Cost of Production Curves

In the case where all nonconforming product is reworked, the excess cost function becomes symmetric and the Effective Cost of Production depends solely upon the capability indexes and the probability model. Using a normal probability model, Figure 5 shows the relation between the Effective Cost of Production and the Capability Index for the case where the process is perfectly centered (so that $C_{pk} = C_p$). Inspection of the way this curve flattens out as we move to the right will suggest that a reasonable definition of economic production would be an Effective Cost of Production of 1.10 or less. When a process is perfectly centered you will enter this zone of economic production only when you have a C_p value greater than 1.05.

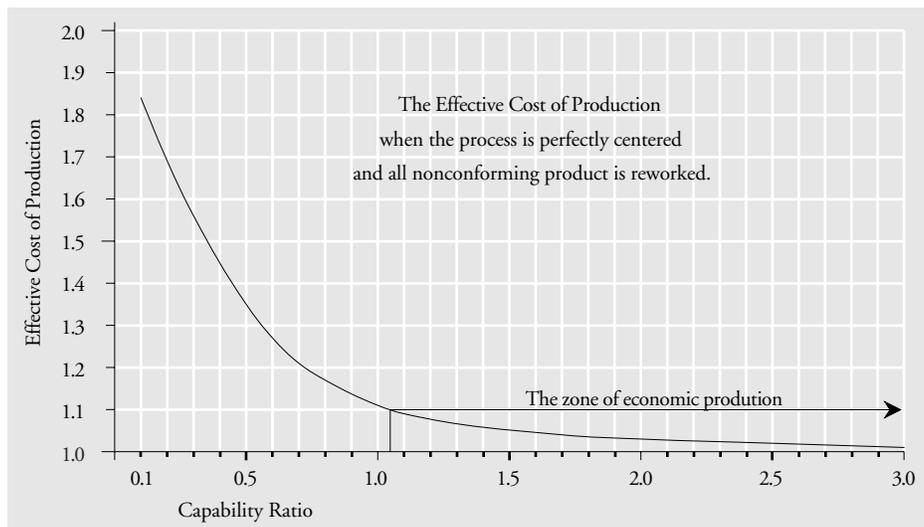


Figure 5: The Effective Cost of Production Curve When $C_{pk} = C_p$

The curve in Figure 5 is the best you can do. Values below this curve are impossible. However, if the process drifts off-target then this curve will no longer apply.

When a process is not perfectly centered the centered capability index C_{pk} will be less than the capability index C_p . This will be represented on the graph in Figure 6 by a vertical shift. A C_{pk} that is 0.3 units smaller than C_p will represent a process average is 0.9 standard deviations off target. The Effective Cost of Production Curve for this situation is shown in Figure 6.

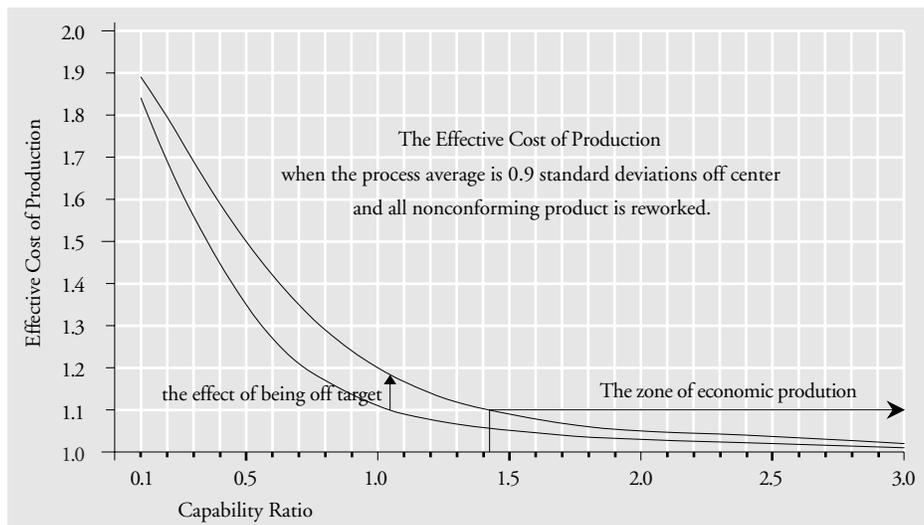


Figure 6: The Effective Cost of Production Curve When $C_{pk} = [C_p - 0.3]$

The effect of being 0.9 standard deviations off target is represented by the vertical displacement between the two curves in Figure 6, and the zone of economic production has shifted to those situations where C_p exceeds 1.42. Since no technique for analyzing small amounts of data will be very sensitive to

shifts in the process average that amount to 0.9 standard deviations or less, the curves in Figure 6 have serious implications for production. In particular, for you and your supplier to achieve the desirable goal of economic production, your supplier will need to have a capability index in the neighborhood of 1.5 or larger, and he will also need to have a way to monitor his process to keep it operating near the target value.

When the centered capability index C_{pk} is 0.5 units smaller than C_p the process average will be 1.5 standard deviations off center. The Effective Cost of Production Curve for this situation is shown in Figure 7. There we see that the effect of being up to 1.5 standard deviations off target will move the zone of economic production out to C_p values in excess of 1.87.

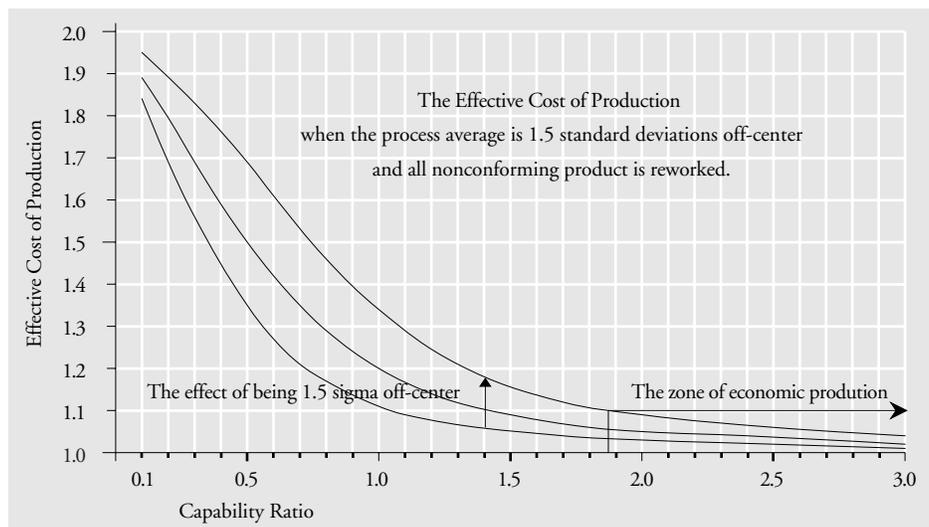


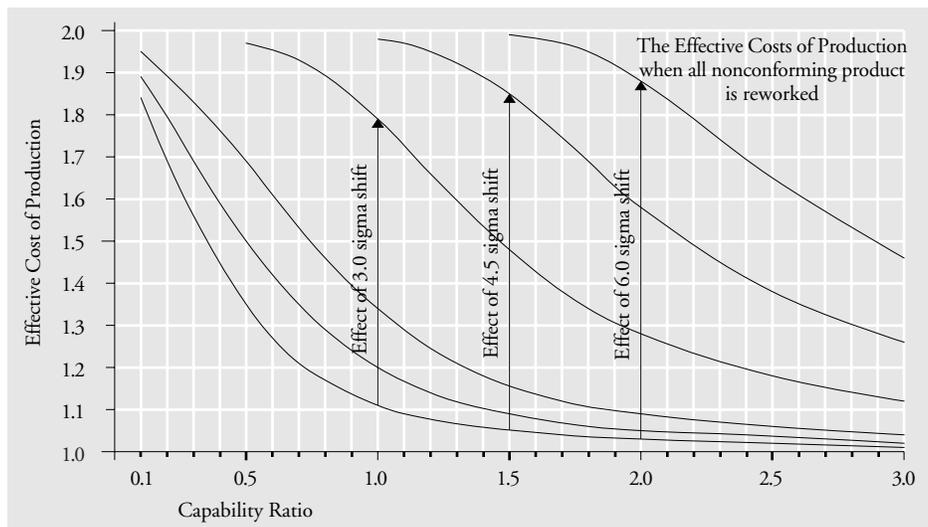
Figure 7: The Effective Cost of Production Curve When $C_{pk} = [C_p - 0.5]$

Techniques for analyzing small amounts of data will tend to be slow to detect shifts in the process average that amount to less than 1.5 standard deviations. As a result, even when your supplier is monitoring his process to keep it centered, there may be periods of time when you operate in the zone between the top two curves in Figure 7.

Continuing in the manner above, we could look at the curves for processes that are off-center by 3.0, 4.5, and 6.0 standard deviations. These curves are shown in Figure 8. Fortunately it is easy to detect when a process is this far off target. Any reasonable monitoring procedure will quickly let your supplier know when the process average has drifted by this amount. Therefore, while the zone of economic production for these last three curves is completely off the graph, it is of little interest to those who have an effective mechanism for maintaining a process on target.

Of course, if your supplier does not have an effective mechanism for maintaining a process on target, then you will suffer the consequences of an Effective Cost of Production that is much larger than it needs to be.

Shifts of 3.0, 4.5, and 6.0 standard deviations were used in Figure 8 simply because shifts of this size are quite common in virtually all production processes. Moreover, as will be illustrated in the next section, even larger shifts are not uncommon. Thus, the consequences shown in Figure 8 do not exaggerate or overstate the situation in the least.



**Figure 8: The Effective Cost of Production Curves
When C_{pk} is 1.0, 1.5, and 2.0 Units Less Than C_p**

Now that we have constructed the Effective Cost of Production Curves, what do they tell us about economic production? By taking into account the fact that small shifts (one sigma or less) are hard to detect, and shifts of up to 1.5 sigma will not be detected quickly, we can use Figure 9 to make some statements about the relationship between capability and the Effective Costs of Production we will actually experience in practice. In Figure 9 the vertical arrows represent the range of values for the Effective Cost of Production that are likely to occur when your supplier operates at various capabilities.

The first vertical line in Figure 9 shows that if we have a C_p of 1.05, and if we also have an effective mechanism for keeping the process centered within the specifications, then most of the time we will have an Effective Cost of Production between 1.10 and 1.18, with occasional periods where the Effective Cost of Production may fall between 1.18 and 1.30.

The second vertical line in Figure 9 shows that if we have a C_p of 1.42, and if we also have an effective mechanism for keeping the process centered within the specifications, then most of the time we will have an Effective Cost of Production between 1.05 and 1.10, with occasional periods where the Effective Cost of Production may fall between 1.10 and 1.17.

The third vertical line in Figure 9 shows that if we have a C_p of 1.87, and if we also have an effective mechanism for keeping the process centered within the specifications, then most of the time we will have an Effective Cost of Production between 1.03 and 1.05, with occasional periods where the Effective Cost of Production may fall between 1.05 and 1.10.

Thus, in combination with an effective mechanism for keeping the process on target, a process will operate in or near the economic zone most of the time only when it has a capability index in the neighborhood of 1.5 to 2.0 or larger.

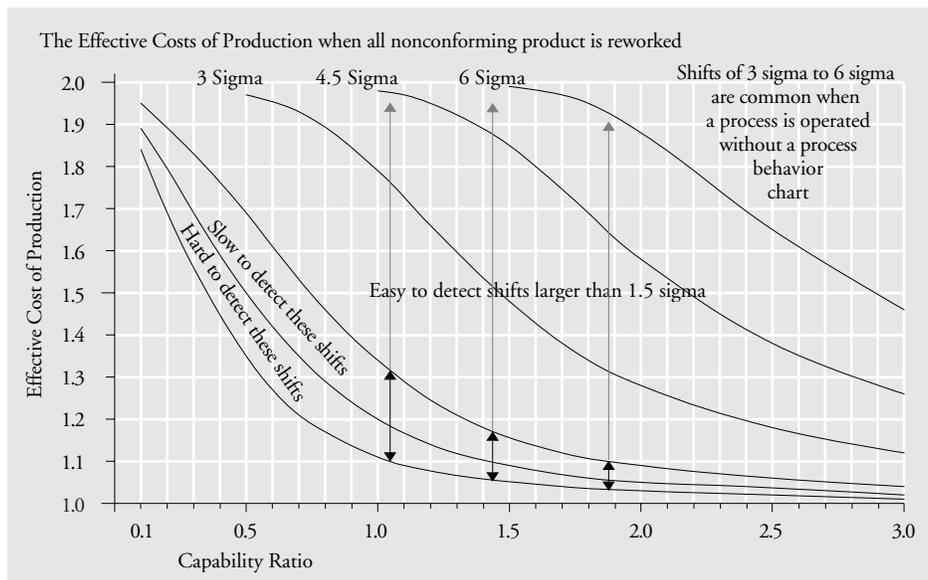


Figure 9: The Range of Effective Costs of Production to Expect at Different Capabilities

All three lines above show that in the absence of an effective mechanism for detecting shifts in the process average, no capability index value can guarantee that you will operate in the economic zone.

In the simple model used above, where all rework was assumed to be completely effective, the limit on the Effective Cost of Production was 2.0. However, in those situations where some product is scrapped there will be no upper bound on the Effective Cost of Production. The only way to realize the benefits of having a comfortably large capability index is to have an effective mechanism for detecting shifts in the process average. If you do not have a mechanism for detecting process changes there is no capability index value large enough to guarantee that you will always operate in the economic zone.

Thus, while it is desirable to have a capability index of 1.5 to 2.0, there is more to economic operation than achieving a good capability number. You will have to be able to detect changes in the process, and keep the process centered within the specifications, even when the capability is 2.0 or larger!

4. Problems with Defects Per Million

In an attempt to explain the economics of operating with capabilities of 1.5 to 2.0 many have converted such capabilities into parts-per-million nonconforming, sometimes referred to as defects per million (DPM).

The first problem with this approach is that it is incomplete. Unlike the Effective Cost of Production, the parts-per-million nonconforming argument focuses only upon the scrap and rework rate and ignores the excess costs experienced by the customer. But the conversion of capability indexes into the typical six-sigma parts-per-million nonconforming values has other problems besides being incomplete.

The conversion of capabilities into fractions nonconforming is fairly straightforward: you choose some probability model to use, determine the relationship between the model and the specification limits using the capability indexes, and compute the area under the model outside the specifications. The result

is your fraction nonconforming. While there is nothing inherently wrong with this conversion, there are situations where the results lack credibility. In particular, when the capability indexes get to be substantially greater than 1.0 the computed values will depend upon the infinitesimal areas under the extreme tails of the probability model. Since the correspondence between our model and our data will always break down in the extreme tails, the conversion of large capability indexes into parts-per-million nonconforming values has always been questionable. Using the infinitesimal areas under the extreme tails of a probability model that you have assumed will characterize your process outcomes is equivalent to raising an assumption to the third or fourth power—the result simply has no credibility. Yet it is exactly these parts-per-million values that are used to obtain the six-sigma DPM values. So the second problem with the six-sigma parts-per-million nonconforming values is that they take a straightforward conversion that is appropriate for parts-per-hundred computations and use it in an inappropriate manner.

Then we have the problem of unpredictable processes. The use of a probability model in the conversion outlined in the previous paragraph implicitly assumes that your underlying process is being operated predictably. So how can the parts-per-million values be adjusted to reflect the demonstrable reality that most processes are not operated predictably?

When a process is operated unpredictably there is not one universe, and one model, but many universes and many potential models. While we might select a generic normal probability model for a predictable process, the problem of selecting a collection of probability models to characterize an unpredictable process quickly becomes intractable. At this point the six-sigma sales pitch engages in a bit of pure rationalization—it *assumes that an unpredictable process will not shift more than 1.5 standard deviations in either direction*. As justification the older documents, those dating from the early 1990s, would usually make a vague allusion to a “research study” supporting this result, however no citation was ever given. More recently even this vague allusion is missing and the assumption is merely taken as an axiom. Given this marvelous, but unfounded, result, it was then possible to adjust the capability indexes to reflect a shift in location of 1.5 standard deviations in one direction and to recompute the incredibly small parts-per-million values. The common practice of listing these recomputed values along with the original, unshifted capability indexes has generated many questions about the origin of the “goofy” DPM numbers found in the six-sigma literature. Thus the third problem with the six-sigma parts-per-million values is their dependence upon the assumption that a 1.5 standard deviation shift in location is a *worst-case* scenario. Not only is this incorrect, it is actually the opposite of the truth.

In the previous section we saw that a 1.5 standard deviation shift was actually a *best case* number. A process operated with the benefit of an efficient mechanism for monitoring the process location will occasionally drift off center. Small shifts will be hard to detect and slightly larger shifts will only be detected slowly, meaning that the process will operate off-center some of the time. However, in the absence of an efficient mechanism for monitoring the process location, there is no limit on the size of the shifts that can occur. The following charts provide examples of this in practice.

The X Chart in Figure 10 shows a process that is shifting around by ± 3.0 sigma, which is twice as much shifting as is assumed to be the “worst case” by the parts per million arguments. While some points fall outside the limits, none fall very far outside those limits. Anyone who has created even a few process behavior charts will have encountered points further outside the limits than those seen in Figure 10.

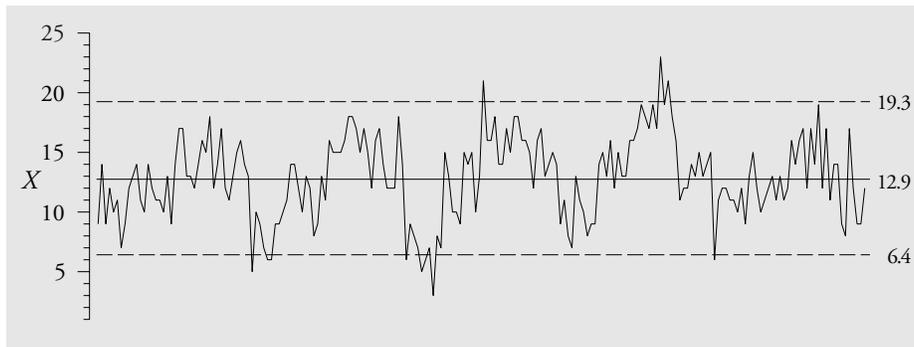


Figure 10: X Chart for Process Shifting ± 3.0 Sigma

Now consider the X Chart in Figure 11. There we see excursions with points that are up to 9 sigma above the central line and up to 8 sigma below the central line. Excursions like these are quite common in practice.

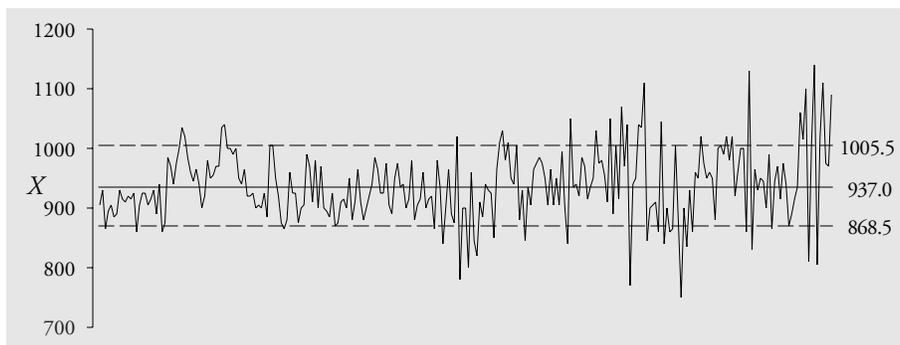


Figure 11: X Chart for Process Having Points at +9 Sigma and -8 Sigma

Or consider the X Chart shown in Figure 12, where the process location moves by an amount equal to 25 sigma during this two week period.

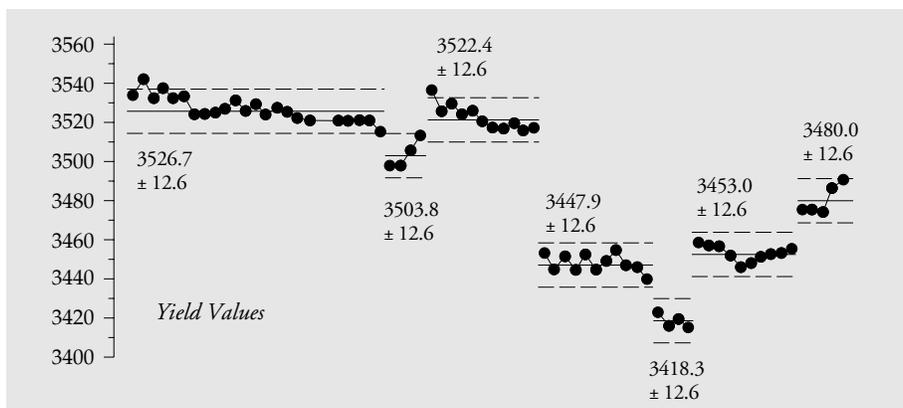


Figure 12: X Chart for Process Shifting by 25 Sigma

In the light of these examples of actual processes, the assumption that an unpredictable process will not shift location more than 1.5 sigma is completely indefensible. It is simply not true. It never was true, nor will it ever be true. Any argument built on the assumption that an unpredictable process is not going to shift by more than ± 1.5 sigma is completely undermined by examples such as those above.

Finally, the fourth problem with the six-sigma parts-per-million nonconforming values is the way they are used in reverse to define a "sigma-level" for a process. Based on the parts-per-million numbers obtained with the 1.5 sigma shifts, the actual defect rate observed during a period of production is used in reverse to create what is, in effect, a pseudo capability index for a process. There are so many problems with this tortured computation that it is difficult to know where to start. Perhaps the most fundamental flaw is that this conversion completely ignores the demonstrated fact that some processes are operated predictably while others are not operated predictably. While this conversion claims to deal with this issue based on the 1.5 sigma shift, the mythological nature of the 1.5 sigma shift undermines this claim. Finally, the pseudo capability index obtained from this conversion (usually expressed as a sigma-level for the process) is grossly incorrect for a predictable process and hopelessly optimistic for an unpredictable one.

Thus, there are four problems with the defects per million numbers commonly used in the six-sigma program. They are incomplete, ignoring the excess costs experienced by the customer. They apply a standard conversion in a non-standard way to obtain values that have no credibility. They treat a best-case adjustment as a worst-case bound. And they then are used in reverse to convert an observed level of nonconforming product into a pseudo-capability index for a process without any consideration of process predictability or the lack thereof. As a result of these four problems any use of the DPM values can only be characterized as a triumph of computation over common sense.

A second, flawed parts-per-million number commonly found in the six-sigma literature is defects per million opportunities (DPMO). As soon as DPM was used in reverse to define a "sigma level" for a process it was inevitable that someone would want to do the same with data based on the counts of blemishes or defects. Of course the problem with counts of blemishes or defects is that they have an area of opportunity that is defined as a finite portion of some underlying continuum. This means that defect rates will inevitably have units attached (e.g. 2.5 blemishes per hundred yards in a bolt of cloth). To work the magic conversion in reverse using the DPM numbers so laboriously computed earlier, these units got in the way (parts-per-million values are dimensionless). So the blemish rate of 2.5 blemishes per hundred yards was simply converted into a Defects Per Million Opportunities value by dividing by the "number of opportunities per hundred yards." If we had "one opportunity per foot" we would get:

$$\text{DPMO} = \frac{2.5 \text{ blemishes per hundred yards}}{300 \text{ opportunities per hundred yards}} \times \text{one million} = 8333 \text{ DPMO}$$

However, if we thought we had "one opportunity per inch" we would get:

$$\text{DPMO} = \frac{2.5 \text{ blemishes per hundred yards}}{3600 \text{ opportunities per hundred yards}} \times \text{one million} = 694 \text{ DPMO}$$

Thus, the DPMO is a totally subjective value that depends upon how you subdivide the continuum into potential opportunities. It is nothing more than data divided by an assumption.

Ultimately, these illogical attempts to define the sigma-level for a process are merely trying to do that which is already done by capability indexes or by direct computations of losses. Since we can now

convert capability indexes directly into Effective Costs of Production, these faulty parts per million values should be scrapped.

5. Summary

With regard to operating in the Economic Zone:

- It can be economical to operate with a capability index of 1.5 to 2.0.
- To realize the full benefits of having a capability index of 1.5 or larger you will need to have an effective mechanism for detecting changes in the process average.
- With an effective mechanism for detecting process shifts you can expect to operate with an Effective Cost of Production of 1.10 or less only when your capability index is 1.5 or greater.
- Without a mechanism for detecting process changes there is no capability index value that is large enough to guarantee that you will operate in the economic zone.

With regard to the defects per million numbers commonly used in the six-sigma program:

- They are incomplete, being based only on the costs of scrap and rework and ignoring the costs of using conforming product.
- They apply a standard conversion in a non-standard way to obtain values that have no basis in reality and no mathematical credibility.
- They treat the amount of drift you will occasionally experience in the best-case scenario as the worst that can happen in the absence of a mechanism for detecting process changes.
- They are used in reverse to convert an observed level of nonconforming product into a pseudo-capability index for a process without any consideration of process predictability or the lack thereof.
- DPMO values are merely data divided by an assumption, and are therefore totally subjective.
- Thus, the best that can be said is that these numbers represent a triumph of computation over common sense.

This manuscript was prepared as a preprint of material that was finally included in
The Six-Sigma Practitioner's Guide to Data Analysis.

Copyright © 2003 SPC Press.

www.spcpress.com/pdf/DJW177.pdf