Don't We Need Good Measurements?

Process behavior charts work with imperfect data

Donald J. Wheeler

Good measurements are like apple pie and motherhood. Who could ever be against having good measurements? Since we all want good measurements, it sounds reasonable when people are told to check out the quality of their measurement system before putting their data on a process behavior chart. Fortunately, this is simply one more bit of advice that is completely unnecessary.

To consider why process behavior charts do not require high quality data we need to define some concepts. First of all we will need a model for an observation, then we will need a way to characterize relative utility for a measurement, and finally we will need to see how process behavior charts work with less than perfect data.

A MODEL FOR AN OBSERVATION

For our model, it should not be too hard to imagine that a single observation consists of two parts: the product value plus the contribution of measurement error.

Observation = Product Value + Measurement Error

When this measurement error is negligible, repeated measurements of the same thing will tend to produce the same observed value. However, as the measurement error gets larger repeated measurements of the same thing will start to differ. As these repeated observations diverge, the relative utility of the measurement system will suffer.

INTRACLASS CORRELATION

We characterize the relative utility of a measurement system for a given application by using the Intraclass Correlation Coefficient, *rho*.   This simple ratio is both theoretically sound and easy to interpret.   It is defined as the ratio of two variances:

$$rho \ = \ \frac{\text{variance of production process}}{\text{variance of product observations}}$$

This ratio is an honest proportion. It describes how much of the variation in the observations can be attributed to the production process itself. When the Intraclass Correlation Coefficient is, say, 0.70, we would know that 70% of the variation in the observations would be due to variation in the production stream.

Since the complement of the Intraclass Correlation Coefficient, (1 – *rho*), describes that proportion of variation in the observations that is attributable to the measurement system, we could also say that when *rho* = 0.70, exactly 30% of the variation in the observations came from the measurement system. Hence, the value of (1 – *rho*) directly characterizes the extent to which

measurement error intrudes into our observations. (For more about the Intraclass Correlation see *QDD* Dec. 2, 2010.)

THREE - SIGMA LIMITS

Now how does the process behavior chart work with less than perfect data? As we have seen in other articles, three-sigma limits will filter out virtually all of the routine variation in our data (*QDD* Aug. 5, 2009, *QDD* Nov. 1, 2010). They do this reliably, and they do this without regard to where that routine variation may have originated. If you think about the fact that measurement error is always present it should be apparent that measurement error will always be part of the routine variation in any set of data.

As the intraclass correlation gets smaller, measurement error will comprise more and more of the routine variation. But regardless of how large (1 – *rho*) becomes, measurement error will still be part of the routine variation. Since routine variation defines the limits, the real impact of increased measurement error upon a process behavior chart is to inflate the limits. Specifically, measurement error will inflate the limits of an *X* chart or Average chart by an amount equal to:

$$\text{Inflation of limits by measurement error } = \frac{1}{\sqrt{\text{rho}}} - 1$$

Eventually this inflation of the limits will desensitize the chart. Therefore, the question is how much measurement error is too much? To rigorously consider this question, we will need to consider how measurement error affects the ability of a process behavior chart to detect a shift in location. When we do this in the next section we will obtain a surprising answer:

Measurements may be used on a process behavior chart until the Intraclass Correlation Coefficient drops below 20 percent. This means that measurements will work on a process behavior chart as long as measurement error contributes less than 80 percent of the total variation.

HOW THE CHARTS WORK WITH IMPERFECT DATA

In order to incorporate the effects of all four Western Electric Zone Tests we consider the ability of an *X* chart to detect a three-sigma shift within ten time periods of when that shift actually occurred. The graph in Figure 1 is derived from my tables of the power function for the *X* chart (*Journal of Quality Technology, Vol. 15, No. 4*, October 1983).

This graph has the values of the Intraclass Correlation Coefficient along the horizontal axis. As the Intraclass Correlation goes from 1.00 to 0.00 the proportion of the total variation that is due to measurement error goes from 0 percent to 100 percent. The vertical axis has the probability of detecting a three-sigma shift in the production process within ten observations of when that shift occurred. The two curves show the ability of the *X* chart to detect this shift. The upper curve is for Detection Rules One, Two, Three, and Four used together (the Western Electric Zone Tests). The lower curve is for Detection Rule One alone (a single point outside the three-sigma limits). Both curves show that the *X* chart maintains a high probability of detecting this shift as the measurement error increases.
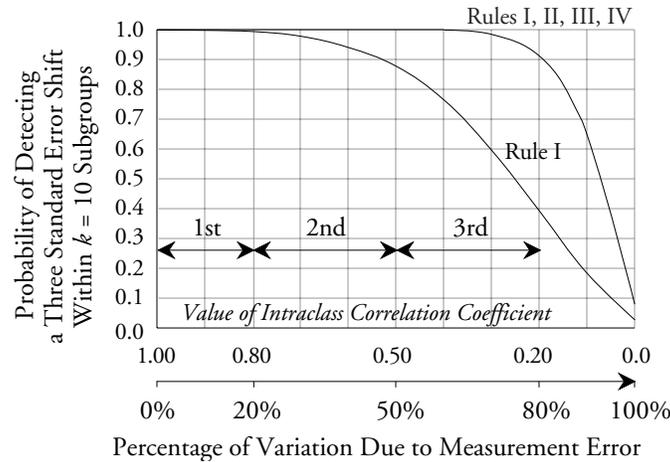
Figure 1: How Measurement Error Affects the *X* Chart

First Class Monitors: When the Intraclass Correlation Coefficient is between 1.00 and 0.80 our measurement system is said to provide a First Class Monitor. Here measurement error comprises less than 20 percent of the total variation, and there is better than a 99 percent chance of detecting a three sigma shift within ten time periods with Rule One alone.

Second Class Monitors: When the Intraclass Correlation Coefficient is between 0.80 and 0.50 our measurement system is said to provide a Second Class Monitor. Here the measurement error will amount to between 20 percent and 50 percent of the total variation, and there is better than a 88% chance of detecting a three sigma shift within ten time periods with Rule One alone. When all four detection rules are used with a second Class Monitor we can be virtually certain that we will detect a three-sigma shift.

Third Class Monitors: When the Intraclass Correlation Coefficient is between 0.50 and 0.20 our measurement system is said to provide a Third Class Monitor. Here the measurement error will amount to between 50 percent and 80 percent of the total variation, yet there is still better than a 91 percent chance of detecting a three sigma shift within ten time periods when using all four detection rules together.

Fourth Class Monitors: When the Intraclass Correlation Coefficient is below 0.20 our measurement system is said to provide a Fourth Class Monitor. Here measurement error becomes so dominant that there is little product information remaining in the observations. The probabilities of detecting a three-sigma shift rapidly evaporate. Any use of a Fourth Class Monitor is an act of desperation.

Thus, measurements will be effective on a process behavior chart up to the point where measurement error amounts to 80 percent of the total variation.

As an illustration of the truth represented by Figure 1 an average and range chart (with *n* = 8) from a client is shown in Figure 2. This measurement process was right at the borderline between a Second and Third Class Monitor, with an estimated Intraclass Correlation Coefficient of 0.52.
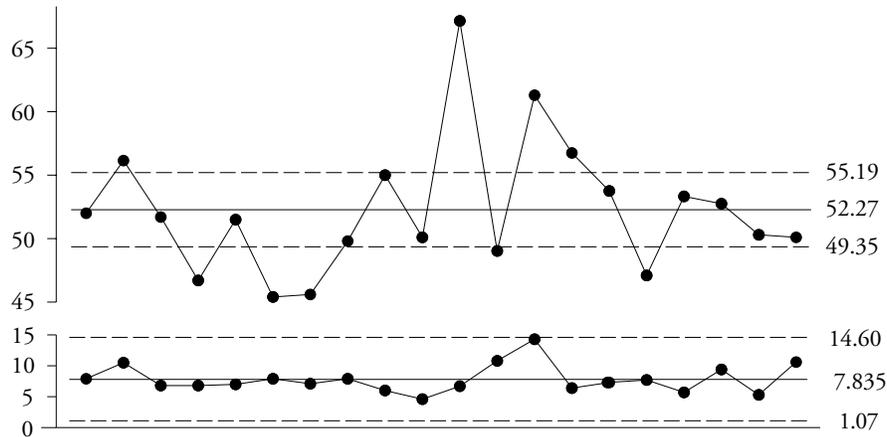
**Figure 2: Average and Range Chart for Product 2131
Created Using a Measurement System having
an Intraclass Correlation Coefficient of 0.52**

With 48% of the routine variation in the observations attributable to the measurement system, and with only 52% coming from the production process, the limits on the Average Chart have been inflated by the measurement error. However, in spite of these inflated limits, there is still no question but that this process is being operated unpredictably. Here it would be a mistake to argue about the quality of the measurements. It would also be a mistake to argue about the quality of the limits, or to estimate the inflation due to measurement error. Our imperfect measurements, used with inflated limits, have detected signals of unplanned changes in the process. Since our imperfect measurement system tends to hide signals rather than giving us false alarms, the only important question raised by Figure 2 is "What are they going to do about the signals of process change?"

The owner of this process needs to discover the assignable causes that are affecting this process. Nothing else matters. If he does not take advantage of the opportunities shown on this chart, then he will continue to suffer the consequences of excess variation, excess scrap, and unnecessary waste.

There is no way to determine the Intraclass Correlation Coefficient from Figure 2, but then we do not need to know this value in order to interpret this chart. Any signals we find are signals that are large enough to show up in spite of the masking effects of measurement error.

SUMMARY

Measurement error will inflate the limits of any process behavior chart. Yet, as shown in Figure 1, we will still be able to detect signals of exceptional variation whenever we have an Intraclass Correlation Coefficient larger than 0.20. While we always like to use the best measurements possible, the process behavior chart works with imperfect data.

This means that you do not have to worry unduly about the quality of your measurements. You do not have to qualify your measurement systems before you can do business with them. As long as you have signals of exceptional variation showing up on your process behavior chart your measurements are good enough. And if you attend to the signals on your charts you will be

able to improve your production process in spite of your imperfect measurements. As you improve your production process you will often find that you can meet specifications without having to resort to sorting and inspection.

Thus, until you get to the point of no longer finding any signals of exceptional variation on your process behavior charts, you do not need to check on the quality of your measurement system. Since money, time, and effort spent on analyzing and upgrading the measurement system is overhead that drains resources from your primary business, this should be good news.