How can an OC Curve be Used to Manage the Risk of Undetected Special Causes?
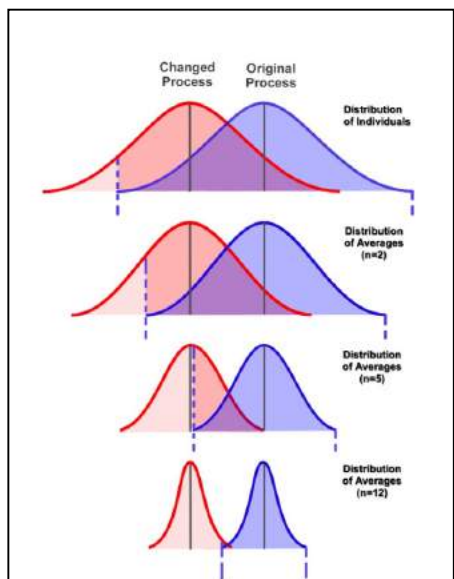
Steven Wachs
Principal Statistician
Integral Concepts, Inc.

When applied properly, SPC provides manufacturers a proven method to increase profitability and achieve a deeper understanding of their processes.. Additionally, SPC can prevent problems—saving companies money that would have been lost in scrap, rework, warranty, litigation, and market share decline. A key factor in obtaining these SPC benefits is the proper deployment of control charts. Correctly designed control charts identify significant *changes* to a process. These can be changes that are still within specification, but are statistically different than where the process was previously running. By identifying the changes, personnel can determine what caused the change and potentially improve the process or prevent the production of inferior products.

SPC practitioners frequently fail to consider how sample size and sampling frequency affect the ability to detect process changes *quickly,* a primary objective of control charts. This article describes the use of Operating Characteristic (OC) curves to understand what type of performance may be expected from a control chart. An OC curve allows practitioners to determine a sample size that will result in the detection of those process changes that are of practical significance while minimizing the occurrence of false alarms (i.e. signals for process changes of little practical significance). In this article, we limit the discussion to charts of averages (x-bar charts). OC curves for other charts will be covered in a future article.

Chart Sensitivity – Detecting Changes

The impact of sample size on chart sensitivity --the ability of a control chart to detect changes -- was explored in detail in a previous *Ask the Expert* article (see *How Should the Sample Size be Selected for an X-bar Chart* – Parts I, II).

The sequence of overlapping curves to the left illustrate the effect of increasing sample size (*n*) on the ability to detect process shifts. As *n* increases, the standard deviation of the distribution of sample averages decreases, making the normal curve narrower. A process average shift becomes

1

easier to detect since the overlap between the original process curve and the changed process curve decreases. We can increase *n* until the overlap is small enough for the next sample average (after the shift) to fall outside of the control limits set up for the original process.

When a significant process shift occurs but it's not detected, a Type II error has occurred. The control chart should have detected the change by giving us a point beyond the control limits, but failed to do so. The probability that a Type II error occurs on a given sample is denoted by β. Conversely, the probability that the shift *is* detected is known as the *Power* and this is equal to 1-β. β and Power are illustrated in the graphic below.
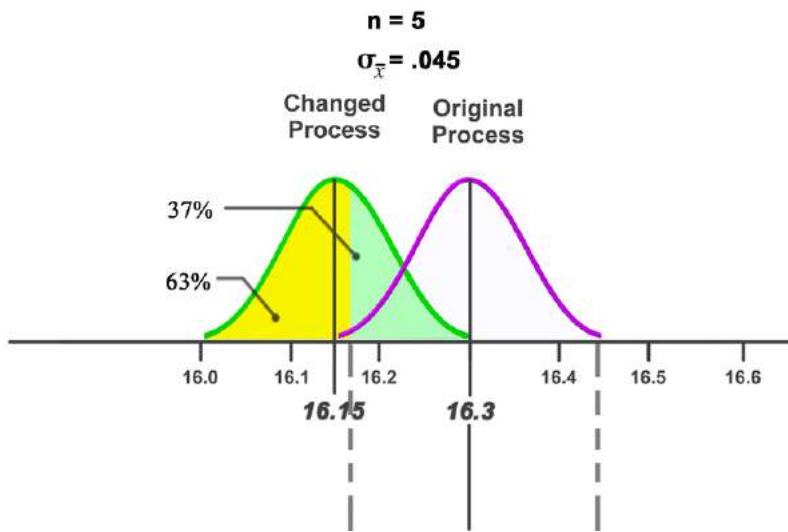
The purple curve to the right represents the original process distribution (centered at 16.3). With the sample size of *n* = 5, the lower control limit is 16.17 and is indicated by the left-most dashed vertical line. At some point, the process average shifts down to 16.15 (the green curve).

Now, on the first sample following the process shift, we are sampling from the green distribution (centered at 16.15). We can see that there is a 37% chance that the sample average will be greater than the lower control limit for the original process (the area shaded in light green). Thus, β is 0.37 since we will fail to detect the process shift if we sample from the green region. On the other hand, if we obtain a sample average less than 16.17 (in the yellow region), we *will* detect the shift. The power -- the probability of correcting detecting the shift -- is 0.63.

The Operating-Characteristic (OC) curve

The OC curve allows us to investigate how β (or power, which is 1-β) depends on the size of the process shift and the sample size. In general:

- the smaller the process shift we would like to detect, the higher β will be
- the smaller the sample size we use, the higher β will be

The size of the process shift is characterized by the number of standard deviations comprising the shift. For example, if the process standard deviation is 0.040 inches, a shift in the mean of 0.080 inches is a "2-sigma shift". Specifying the process change in this manner allows us to generalize the OC curve. Also, the OC curves shown below assume that the control limits are established at a distance of 3 standard deviations from the estimated process average (per standard control charting practice).
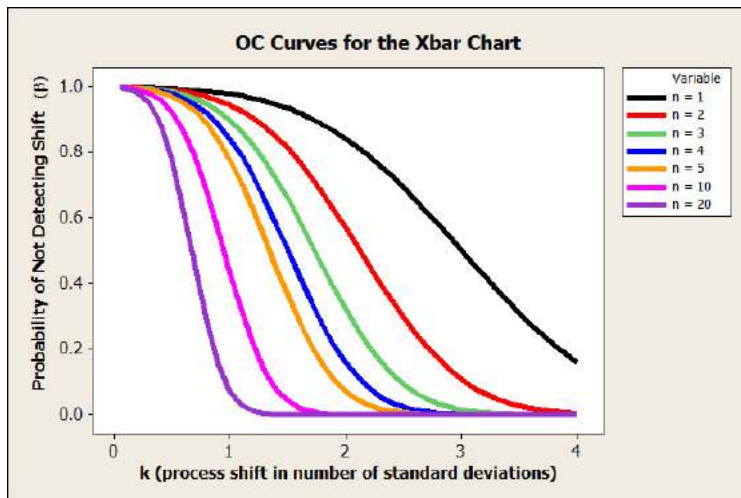
The graphic to the left shows multiple OC curves for various sample sizes. Each OC curve shows the probability of *not* detecting a specified shift in the process average on a given sample, specified in a number of standard deviations.

We can see that for a given process shift (k), as n *decreases*, the probability of not detecting the shift *increases*. Also for a given Type II error probability (β), as n increases, we are able to detect smaller shifts.
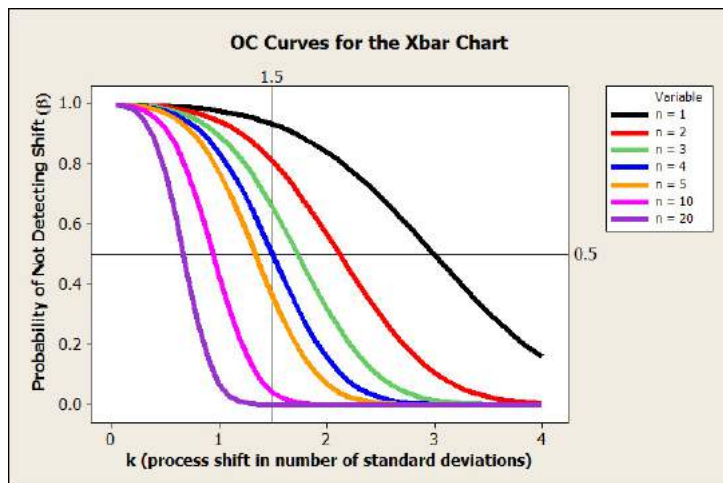
As shown on the graphic to the left, for a sample size of 4 (the blue OC curve), a mean shift of 1.5 standard deviations has a 50% chance of being detected on a given sample.

For the same size shift, a sample size of 10 (the pink OC curve) will detect the shift with high probability.

As can be seen from the above curves, for the typical sample sizes used in practice (n = 3 to 5), the x-bar chart is not very effective in detecting small shifts (e.g. shifts of 1.5σ or less) on the first sample following the shift. While this may be troubling, the probability of detection improves with each successive sample. We can compute the expected number of samples that

will need to be taken before the shift is detected. This is called the *Average Run Length* and is equal to $1/(1-\beta)$. So, if we have an 80% chance of *not* detecting the shift on a given sample (β = 0.80), then the Average Run Length = $1/(1-.8)$ = 5. That is, on average it will take 5 samples for us to observe an out-of-control point when the probability of detecting on a single sample is 0.20.

We can also determine the probability of detecting the shift after a given number (*x*) of samples, where we may actually observe the signal on any of the *x* samples. It is simply $1 - (\beta)^x$. Let's look at an example.

As we saw above, the probability that a 1.5 sigma shift is detected with a sample size of 4 is 0.50. Now, what is the probability that the shift will be detected at least by the end of the third sample following the shift? We have P(Detect) = $1 - (.5)^3$ = 0.875. So while the chance of detecting on each sample is akin to flipping a coin, the chance that we are successful after 3 attempts is nearly 90%.

Sample Size and Frequency of Sampling

The previous section illustrates that although the power of an x-bar chart to detect a specified shift on one sample may not be ideal, frequent sampling will lessen the risk of producing defective products following an undetected process change.

Where sampling effort is costly (e.g. not automatic), we must decide how to allocate our sampling effort. For example, we can opt for small, frequent samples or take large samples but less frequently. We might be considering a sample size of 5 every 2 hours, or a sample size of 10 every 4 hours.

The best sampling plan depends on several factors such as:

- The cost and ease of sampling units
- Production volumes and rates
- The cost associated with producing defective units
- The current process capability
- The impact of variation on product performance

Frequent sampling minimizes the amount of production that occurs under a changed process if a process change is undetected due to a Type II error. On the other hand, we have seen that the probability of making a Type II error can be made very small by selecting a large enough sample size.

Summary

The specification of arbitrary sample sizes and sampling frequencies imposes (frequently unknown) risks on manufacturers applying SPC. This article describes the use of OC charts for understanding and quantifying the inherent risks of undetected process changes. By considering both the designed-in probability of non-detection and sampling frequency, an acceptable level of risk can be managed effectively.