Characteristics of Special Cause Variation

An On-Line Debate with "Bob"

I belong to several groups on the Linked In web site. I find some of the discussions both fascinating and informative. Not long ago, on one of the group sites a member posted the following topic for discussion: "What is variation? Please provide an example of Common Cause Variation and Special Cause Variation."

Her post received a number of interesting responses, and I decided to pitch in. I wrote,

Common cause variation is controlled, random and mathematically predictable. It results from common causes of variation from within the process or system. Common causes produce outcomes that are different, but not significantly different.

Special Cause variation is uncontrolled, non-random and not mathematically predictable. It is due to some special cause (Walter Shewhart called it "assignable cause") of variation from outside the process or system. Special causes produce outcomes that are not different, they're significantly different.

Though some may view your question as elementary, I commend you for raising it. Dr. Deming insisted that understanding variation is essential. Failure to differentiate between these two types of variation, he wrote, leads to very costly mistakes.

For example, my second son, Jacob, was born with Down syndrome. At school, his cognitive ability would score significantly low – in need of special help. Unfortunately, the blind pursuit of "mainstreaming" in American education would place Jacob in a regular classroom.

Jacob will not get the special help he needs in the regular classroom. And the only way the regular classroom teacher could provide that special help would be to take away from all the other students. Thus, everyone loses; staggering cost; failure to understand and respond appropriately to your question.

[Another group member named "Bob" (not his real name) posted these comments:]

[Donald J.] Wheeler's book, *Understanding Variation*, is excellent... and I love his subtitle: "The Key to Managing Chaos!"

Jim: Sometimes a special cause can come from within the process. :-)

[In response to "Bob's" post I wrote,]

"Bob," your question (observation) is often raised in my seminars. Dr. Deming insisted, however, that special causes can only be identified via a non-random or out-of-control pattern on a control chart.

Wheeler and Chambers (*Understanding Statistical Process Control*) noted that special causes "are not part of the overall system, and should be considered abnormalities." Thus, they cannot come from "within" the process. The same authors referenced the work of Shewhart and Deming and noted that special causes will have a "detectable effect" because they are "not part of the system."

As an example, a Quality Manager at a paper mill started as a machine tender. He explained during one of my "Transformational Leadership" seminars that at an evening match in his local wrestling league he dislocated his right shoulder. So, he went to work the next day with his dominant arm ace-bandaged to his side. While awkwardly setting the machine with his left (non-dominant) hand, he caused "breaks" at the machine during his shift; but not a significant number of breaks. He asked, "Do you mean that an operator with an injured arm, causing breaks with his other arm, is not a special cause?"

I answered that, *inter alia*, it could not be a special cause because it had no significant effect. The level of common cause variation at the machine was so great that even an injured operator was not "special;" i.e., not significant; a part of the process. On the other hand, if through process improvement the level of common cause variation is reduced, the injured operator may indeed cause a significantly high number of breaks; i.e., special cause variation; no longer a part of the system.

That is the beauty of reduced common cause variation via process improvement. A future event will peak beyond the new control limit, whereas in the current process the same event will fall within the realm of random, common cause variation.

["Bob" was not convinced and replied,]

Jim, I don't disagree with your first paragraph of your latest post.

My counter example is simple. The operator using his non-dominant hand caused enough breaks that it was a Special Cause – and he was within the process.

Or, a group of 13 surgeons who do colorectal operations on cancer patients year-round – all in the same facilities, same equipment, same staff... same system. One surgeon's mortality rate is statistically higher than the other 12... a Special Cause within the system.

Just because something has a detectable effect does not mean, ipso facto, it's not part of the system. My concern is that some may take you literally and only look for Special Causes outside a system.

[I replied,]

"Bob," materials are a part of the process, as are vendors. A batch of material received from the vendor with a significantly high level of contamination, because it will have a significant effect, is not from within the process. As Wheeler and Chambers noted, it would show up as an "abnormality" and therefore no longer "a part of the overall system."

When we design and validate processes, we select materials and vendors. We did not select the significantly high level of contamination that may show up from the selected vendor. Hospitals hire surgeons and design their surgical process. A surgeon with a significantly high mortality rate would no longer be a part of the same system as his colleagues.

My Down syndrome son Jacob may be a student at the Woodstock Elementary School. He would not, however, be a part of the same system as his classmates.

Your view is broadly held, and I have seen it again and again among some of my clients. I also see a lot of Deming's Mistake 1 (tampering) among the same clients until I educate them on theory of variation. Let's return to my one-armed paper machine operator.

- 1. We observe a high, though not significantly high, number of breaks today.
- 2. Upon investigation, we find that one of the machine operators had an injured arm and caused breaks by awkwardly setting the machine with his non-dominant arm. We mistakenly conclude that he was a special cause from within our process.
- 3. We send the operator home on unpaid sick leave. (After all, the injury occurred in an amateur wrestling match, outside of working hours.)
- 4. We decide to change the process spec and SOP to make it clear that operators with injured arms are no longer to operate the machine.
- 5. We spend a week or two processing the change request (or ECN or ECO or whatever you call it) to make the changes to the SOP and process spec.
- 6. In the future, we suffer the same high levels of breaks and defects, at random because we never changed the process that produced the breaks to begin with!

As Dr. Deming put it, via such tampering we would have "made things worse, defeating our own purpose."

["Bob" tried again:]

Jim, I agree, "he was a special cause from within our process."

But once a Special Cause is identified and corrected, why would you change the system – as if you're attempting to reduce Common Cause? If the system was performing acceptably before the Special Cause... it's a form of waste to tamper with it after removing a deleterious Special Cause.

When the surgeon's mortality data were initially analyzed together - looking at one system that produced them - all appeared as Common Cause variation. But when individual surgeons' outcomes were compared, one was significantly higher than the rest.

We'll just have to agree to disagree, and leave it there. :-)

[I replied,]

You're right, "Bob." If the system is performing acceptably before the Special Cause, there would be no need to attempt to reduce Common Cause variation after one has found, removed and taken action to prevent the reoccurrence of the Special Cause.

I thought, however, that we were talking about a modern, highly-capitalized paper mill, schools and surgery. In those cases, is there any such thing as "performing acceptably?" Deming's Point 5 urges us to "Constantly and forever improve every process for planning, production and service," because in this New Economic Age there is and can be no such thing as "good enough."

On that point, I'm sure there's no disagreement.

[Another group member joined the discussion:]

Sorry I'm coming to the party late. "Bob," out of curiosity, when you say one surgeon's mortality rate was significantly higher than the rest, can you fill me in on how this was determined?

["Bob" replied,]

Data were segmented by surgeon, risk-adjusted, plotted on needle plot. Even then, the referenced surgeon's risk-adjusted mortality rate was within the +/- 3 standard deviation limits. But, "scanning" across the plot left-to-right, his datum point was suspiciously higher than the others. So, I removed it from the calculations and then his point was well outside the re-calculated UCL.

[I couldn't help myself. I decided to try again to reach "Bob."]

"A process may be stable, yet turn out faulty items and mistakes. To take action on the process in response to production of a faulty item or mistake is to tamper with the process. The result of tampering is only to increase in the future the production of faulty items and mistakes, and to increase costs – exactly the opposite of what we wish to accomplish" (W.E. Deming, *The New Economics for Industry, Government, Education,* p. 202). To remove an item within +/- 3 standard deviations because it's "suspiciously higher" than other readings and re-calculate control limits with the remaining values is to tamper.

["Bob" revealed that he could speak French and launched into some statistical gymnastics.]

Au contraire, Jim - not tampering but rather the appropriate in-context application of statistical thinking, statistical methods, and a clear understanding of process improvement:

1) If you review my posts regarding this real-life example, you'll see there were only 13 data points. [Now, before you go off on the typical myth of "but you need a minimum of 20-30 data to construct a chart" read Wheeler's very rational, and practical, comment on that issue at the bottom of page 45 of his Understanding Statistical Process Control (2nd ed, 1992).]

2) With so few data a single point (as in "outlier") can have a disproportionate influence on the value of the standard deviation - inflating it such that the Control Limits are so wide that they encompass the outlier itself.

3) Recall that the Center Line (process average) and Control Limits should be calculated only using data that vary due to Common Causes.

4) In this instance the suspect point was just inside the UCL while the other 12 varied closer about the Center Line - the suspect point "stood out" just looking at the data as a whole (thus my comment about "scanning the data left-to-right").

5) So, one may think in terms of a null hypothesis that all the data come from a single population (differences due to Common Cause variation only) and are thus "homogeneous" (constant) in their variance, with an accompanying alternate hypothesis that this is not so (a Special Cause exists).

6) One way to test the null hypothesis is by calculating 13 sets of averages (Center Lines) and confidence limits (Control Limits), holding out one datum each time and seeing if that one point falls outside the limits.

7) But two issues to consider, one practical and one statistical: (a) When you examine the plot of the data (always a good first step in analysis) only one point was "far" away from the other 12; (b) doing many statistical tests opens you up to the Bonferroni Effect (or "Inequality"), in which your actual, total Type I Error risk (alpha) is not the typical 0.05 but something much larger - roughly the # of tests times the 0.05 (not a good thing).

I hope this helps you understand what I did and why.

["Bonferroni Effect?... 'Inequality?'... Type I Error risk (alpha)?... Roughly the # of tests times the 0.05 (not a good thing)?..." I just had to respond.]

"Bob," I thought you had suggested that we should just agree to disagree. I've been very busy, so I didn't have a chance to respond to your latest convoluted attempt to defend your belief that "sometimes a special cause can come from within a process." Earlier you explained that even though one out of thirteen surgeons' mortality rates was within +/- 3 standard deviation limits, you determined that he was a special cause because "his datum point was suspiciously higher than the others."

You have every right to believe that one can dismiss a value because it looks suspiciously high; just as you have every right to believe that a special cause can come from within a process. You do not have the right, however, to misrepresent sources you cite in trying to defend your position.

For example, you cited page 45 of Wheeler's (and Chamber's) text, *Understanding Statistical Process Control*, 2nd ed. Nowhere on that page, nor on any other page of their text, do the authors advocate dismissing a value within control limits because it looks "suspiciously higher than the others." In fact, Dr. Deming used his funnel experiment in part to demonstrate the error and adverse effects of such a mistaken conclusion.

Nonetheless, I was so impressed by the statistical gymnastics you used in trying to defend your position that I decided to try them out for myself. I rolled a pair of dice thirteen times, resulting in the following outcomes: 5, 4, 7, 6, 2, 9, 7, 5, 12, 4, 6, 5, 7. The average was 6.08.

(Everyone knows that 7.0 would be the true average for a dice-rolling process. Clearly, my 13-roll sample was not large enough to accurately represent the process. You claimed, however, that Don Wheeler said it's okay to work with insufficient samples; so I continued my analysis.)

The standard deviation came out to 2.50, yielding +/- 3-sigma limits of 0 and 13.58. All thirteen values fell within the limits. Like your 13th surgeon's mortality rate, however, I decided that the 12 looked "suspiciously higher than the rest;" so I removed it. The average of the 12 remaining values came out to 5.58.

(Everyone knows that 7.0 would be the true average for a dice-rolling process. Clearly, my sample was not large enough to accurately represent the process. You claimed, however, that Wheeler said it's okay to work with insufficient samples; so I continued my analysis.)

The standard deviation came out to 1.83, resulting in revised +/- 3-sigma limits of 0.09 and 11.07. Lo and behold, the 12 fell outside those limits! I guess I could conclude that when I roll a pair of dice and get a 12, it could be due to a special cause from within the dice-rolling process.

"Bob," you have every right to reach such a conclusion; just as you have every right to believe that "sometimes a special cause can come from within a process." On both counts, however, you'd still be wrong.

Let's return to your original suggestion "that we'll just have to agree to disagree, and leave it there." I have neither the time nor the patience to keep trying to correct you.

["Bob" replied,]

Ah, Jim, you mis-represented my reference to Wheeler. It had to do with *how many* data are needed to construct a control chart, as he plainly explains, and as I clearly stated just before I cited him. I did not cite him in defense of my *analysis* of the 13 data.

Rolling dice is a random process (if they are fair dice), so naturally you get *random* results - hardly a surprise, and *not* representative of the situation I described. The one surgeon's result was *not* due to random variation compared to the variation among the other 12 surgeons' results. When those 12 are used to construct the +/- 3 standard deviation Confidence Interval about their average, the 13th lies well outside.

I can't explain it any more simply, nor in any more different ways... :-)

[In an attempt to wrap up our dialogue, I wrote,]

"Bob," you indeed misrepresent a source when you use it to defend the use of small sample sizes but fail to note that the same source says nothing about identifying out-of-control events within the control limits calculated from the small sample. You're right, however, to note that rolling dice (if they're fair dice) is a random process; just as any stable process is a random process. You report that your 13th surgeon's mortality rate fell outside the limits calculated from the other twelve rates; just as the 12 I rolled fell outside the limits calculated from the other 12 values.

Your surgeon and my 12, however, did fall within the limits calculated from all thirteen values. Therefore, neither showed evidence of being the result of a special cause.

Let's return to Elizabeth's original question, "What is variation?" One way to answer that question would be to suggest that there are two species of variation: controlled variation and uncontrolled variation (per Shewhart); or common cause and special cause variation (Deming). There are many, many – in some cases countless numbers of – common causes of variation at work in any process, none of which will have a significant effect on the process output. On the other hand, there will be a lot fewer special causes of variation at work on any process, any one of which will have a significant effect – so significant that it will fall outside the limits we used the same values to calculate.

Neither your 13th surgeon's mortality rate nor my 12 fell outside the original limits calculated from all of the data in our small samples. Thus, we cannot with validity conclude that they were the result of a special cause.

Conclusion

My Naval Academy friend and classmate, the late Wiley Voorhies, once told me, "Jim, you should never try to teach a pig to fly. You'll just end up wasting your time and frustrating the pig!" I guess I would have saved a lot of time with "Bob" if I'd followed Wiley's advice...

© 2013 James F. Leonard. All rights reserved.