When Did Six Sigma Stop Being a Statistical Measure?

Joe Schofield
Sandia National Laboratories

The term Six Sigma is widely used as an approach for process improvement and learning. It is a disciplined, structured, data-driven methodology to solving problems. Along the path to popularity, Six Sigma lost its meaning as a statistical measure and instead inherited the meaning of merely another measurement program. Organizations that intend to employ Six Sigma ought to consider which definition of six sigma *is their target: a process improvement approach, or a statistical measure for variation. This article explores the significance of the differences between six sigma and Six Sigma. Read on if you dare.*

When did *free* come to mean *free after rebate*? When did cost become *the China cost*? And when did *Six Sigma* become something other than a statistical notion? Or have you not noticed that the term *Six Sigma* no longer means a *statistical measure for variation*? For every organization that attempts to use six sigma as a statistical measure of process improvement, three other organizations use it merely to describe a process improvement effort. Most of these organizations have no intention of using six sigma statistically, but it likely impresses the folks higher up in the food chain.

Affixing lean to the term, as in Lean Six Sigma (LSS), is currently an institutional silver bullet. Do not feel left out if you have not been exposed to LSS; one Internet search found only about 125,000 LSS-related sites, but more than 1.7 million sites for Six Sigma. Very few of these sites advocate six sigma's statistical meaning, contributing to the miscommunication regarding Six Sigma processes.

As with most trendy initiatives, LSS has it own status symbols: green belts, black belts, and an assortment of colors and variations depending on the accrediting organization. In addition, LSS has a lexicon, words like kaizen, kaikaku, kanban (yes, there are more than just *k* words). There is one more Japanese word that the LSS industry may have forgotten: it is *muda,* or the word for *waste.* Without applying statistical measurement, organizations may be wasting their process improvement resources.

The application of LSS may bring numerous well-intended results, including defect reduction, work in progress reduction, cycle time reduction, cost savings, fewer hand-offs and queues, minimized changeover time, workload leveling, and more. Organizations in pursuit of process improvement are often well-advised to consider LSS to diagnose, improve, and measure their processes.

Motorola Corporation gets much of the credit for popularizing Six Sigma and the phrase *3.4 defects per million* – the battle cry of the Six Sigma world. Simply re-stated, Six Sigma has come to be synonymous with no more than 3.4 defects per million opportunities (DPMO). An opportunity might be defined as a keystroke or a mouse click, depending on whether the process being measured is developing software or writing an article.

Often, the value 3.4 DPMO is followed with a footnote or an asterisk; the fine print typically ignored. Six Sigma proponents claim that the 3.4 DPMO is the long-term process performance after the occurrence of a *sigma shift.* The *sigma shift* is a 1.5 sigma difference from 6 to 4.5 sigma performance. The underlying assumption is that short-term performance (of say 6 sigma) is really 4.5 sigma in the long term as entropy sets in. Sigma shift translates to more defects per million – 1,700 times more. Statistical 6 sigma is not 3.4 DPMO, it is actually 2 DPBO; that is defects per *billion* opportunities, a difference factor of 1,700.

Did you just get a sense of uneasiness? Remember that most companies claiming the use of six sigma for process improvement are not using either of these statistical values; they are merely targeting their processes for measured improvement.

What if performance improved over time, though, in contrast to being subject to entropy? A sigma shift for better would be a 7.5 sigma process. A 7.5 sigma process would have three defects per hundred trillion (3.1 DPhTO [Schofield notation]).

While a 7.5 sigma process seems an unreasonable expectation, at this rate the commercial airline industry would encounter a fatal event every 17,500 years, U.S. highways would incur 23 deaths per year instead of 40,000, and three deaths per annum would be realized from prescription defects instead of 7,000.

But a 7.5 sigma performance is not unreasonable in the computing world.

Consider for a moment a teraflop machine that operates at one trillion floating point operations per second. In a mere 100 seconds, three defects would be generated. Within one year, 1,246,080 defects would be generated.

It gets worse. Within the next year (or so) the petaflop machine will be released. A machine operating at that speed could generate over more than one billion defects per year if operating at 7.5 sigma. Do you feel more unease? Do not get prematurely paranoid – a petaflop machine is unlikely to appear on your desktop anytime soon.

Fortunately, hardware performs far more reliably than the sigma levels just described suggest, but that does not apply to software. Software defects cost the U.S. economy almost $60 billion a year [1]. Of course, software defects are not limited to software. Auto companies such as BMW, DaimlerChrysler, Mitsubishi, and Volvo have all experienced software-related product malfunctions (defects) that include engine stalls, wiping interval problems, gauge illumination defects, and transmission gear errors [2]. Software technicians in Panama were charged with murder after 21 patients died from gamma ray overdoses in just 40 months [3]. Sorry, no sigma levels released. And yet, 62 percent of polled organizations lack a software quality assurance group [4].

Practicing statistical *something sigma* is an industry best practice. The Software Engineering Institute's Capability Maturity Model® Integration recognizes the relevance of measurements and analysis by placing it prominently as a Level 2 Process Area in its staged representation. Later in the model, there is the need to identify assignable and common cause variation at maturity Levels 4 and 5, respectively.

So when did statistical notions become ambiguous with words like Lean and Six Sigma? Perhaps organizations should raise an *alert* when the term *six sigma* is used to investigate its contextual alignment with

expectations, visions, and goals. Perhaps, too, the process improvement initiative will have an increased likelihood of success – regardless of what it is called.

Conclusion

Given the abundance of quality improvement and Six Sigma tools available to organizations today, incorporating six sigma measurements might not be that difficult – if the organization chooses to do so. For instance, brainstorming techniques for current state weaknesses could be validated with statistical data (perhaps not to a six sigma threshold, but the introduction of any statistical validation on root cause analysis might provide relevant insight into weaknesses). Root causes listed on cause and effect (Fishbone) diagrams could similarly be validated with statistical data collection. Process flow maps could use the distribution of a statistical sample in assigning hands-on and queue time measurements. Each of these uses of statistics would begin to reintroduce the use of quantitative measures into the Six Sigma movement, perhaps leading to the reemergence of six sigma quality thresholds.

Mark Twain probably was not thinking about Six Sigma when he described the three types of lies as lies, darned lies (paraphrased), and statistics, but his quote seems apropos given how Six Sigma proponents use six sigma today. Six Sigma should be reserved for, well, six sigma performance – a statistical measure for variation. Maybe then quality will translate to fewer product recalls, lower costs will mean that costs are decreased, and six sigma performance will equate to two defects per billion – maybe that is asking too much. Distinguishing between statistically measured performance and measured performance can help assess the true progress of an improvement effort. When applying Six Sigma for process improvement, do not leave out the six sigma.◆

References

1. Zaino, Jennifer. "Behind the Numbers." Information Week. 29 Mar. 2004: 94.
2. Sullivan, Laurie. "Software Quality." Information Week. 15 Mar. 2004: 56.
3. Gage, Debbie, and John McCormick. "We Did Nothing Wrong." Baseline-The Project Management Center, 4 Mar. 2004.
4. Surmacz, Jon. "Why Software Quality Stinks." CIO. 1 Dec. 2003: 28.

About the Author

Joe Schofield is a distinguished member of the technical staff at Sandia National Laboratories. He is a trained Lean Six Sigma Black Belt, chairs the organization's Software Engineering Process Group, is the Software Quality Assurance Group leader, and is accountable for the introduction of the Personal Software Process℠ and Team Software Process℠. He has dozens of publications and conference presentations. Schofield chairs the Management Reporting Committee for the International Function Point Users Group, is active in the local Software Process Improvement Network, and has taught graduate level software engineering classes since 1990.

Sandia National Laboratories
MS 0661
Albuquerque, NM 87185
Phone: (505) 844-7977
Fax: (505) 844 2018
E-mail: jrschof@sandia.gov