Golf Scores and Six Sigma Metrics: Confusion and Resolution

Forrest W. Breyfogle III

Measurements affect behavior. The wrong behavior occurs when metrics do not represent what is happening. The leaders of many respected companies are now paying the price for creating an environment in which measurements do not accurately reflect what is going on in their organizations. They created an environment in which it is common practice to play games with the numbers to make short-term numbers look good at the expense of long-term performance for stockholders and employees. They did not have a strategy with which high-level business measurements could be tracked as a process. Tracking these metrics leads to the creation of a meaningful strategic plan, which then aligns process improvement/re-engineering projects to this plan.

Similarly, some measurements from Six Sigma can deceive. Care needs to be given to the metrics that will be used and how they are interpreted when creating a Six Sigma infrastructure; otherwise there can be a lot of pressure for practitioners to also play games with the numbers. The measurements used in a Six Sigma infrastructure need to drive the right behavior for all processes, not just for manufacturing.

In this article I will use golf to illustrate Six Sigma measurement alternatives, and the confusion that can be created.

Consider that we had two 18-hole golf rounds with the following scores. Think of these scores as a sequence of products produced or transactions executed over two days.

Glossary

Birdie: A hole completed in one shot under par.

DPMO: When using the non-conformance rate calculation of defects per million opportunities(DPMO), one needs first to describe what the opportunities for defects are in the process; e.g., the number of components and solder joints when manufacturing printed circuit boards. Next the number of defects is periodically divided by the number of opportunities to determine the DPMO rate.

Par: The regulation number of strokes set for a hole played perfectly, determined by yardage and design of the hole.

Process capability indices (C_p and C_{pk}): C_p is a measurement of the allowable tolerance spread divided by the actual 6s data spread (6 times standard deviation). C_{pk} has a similar ratio to that of C_p except that this ratio considers the shift of the mean relative to the central specification target.

Process capability: AIAG (5) definition for the variables data case is 6s range of a process's inherent variation; for statistically stable processes, where s is usually estimated by [is the average within sub-group range from an R control chart and d_2 is a constant]. For the attribute data case it is usually defined as the average proportion or rate of defects or defectives; e.g., center of an attribute control chart.

Process performance: The AIAG (5) definition is the 6s range of a process's total variation, where s is usually estimated by s, the sample standard deviation.

Sigma level or sigma quality level: A quality that is calculated by some to describe the capability of a process to meet specification. A six sigma quality level is said to have a 3.4 ppm rate.

Day 1

Hole number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Totals
Par	4	4	3	4	5	4	4	4	4	4	5	4	4	3	4	4	4	4	72
Score	5	6	4	5	7	4	6	5	4	7	7	5	5	5	4	4	5	4	92
Difference	1	2	1	1	2	0	2	1	0	3	2	1	1	2	0	0	1	0	20

Day 2

Hole number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Totals
Par	4	4	4	3	4	5	4	4	4	4	4	4	5	4	4	3	4	4	72
Score	5	6	4	3	4	6	5	5	6	6	5	3	6	5	4	4	5	4	86
Difference	1	2	0	0	0	1	1	1	2	2	1	-1	1	1	0	1	1	0	14

The question is - how should these metrics be reported? Within a Six Sigma infrastructure we encounter process capability/performance indices (See Glossary) such as C_p, C_{pk}, P_p, and P_{pk}. We also encounter defects per million opportunities (DPMO) and sigma quality level metrics. The general purpose of these metrics is to quantify how well a process or product is doing relative to a criterion.

Consider a Six Sigma Black Belt whose employer requires the reporting of these units for all projects. The question here is how should he or she make these calculations? For the purpose of this article I will focus on the sigma quality level.

One could define an opportunity for defect at the golf-stroke level. That is, one would say a stroke has a defect when the ball was not hit properly. However, this approach can lead to playing games with the numbers, since this definition for a defect really has various levels of "goodness" and does not necessarily relate to our overall score.

Another approach would be to define a defective event as when we exceeded par (see Glossary) for a hole. Using this definition for the two days noted above, we could count 11 defective holes out of 36 holes, noting that a birdie counted the same as a par hole. This translates to a defective rate of 0.305555 (11/36) or a parts per million (ppm) defective rate of 305555. Using a table such as Table S from reference 1, this ppm rate approximately equates to a sigma quality level of 2.0.

Now let's consider how well the sigma quality level metric describes our golf game. If we reported a 2.0 sigma quality metric to a friend, would it be clear to him how well we are doing in our golf game? Also, do we understand how our golf game is progressing over time? Do we have a strategy that naturally leads us to quantifying the source of variability within our golf game? Do we have a measurement strategy that will ask us questions about what we might do differently to improve? The answer is no to all these questions. This same confusion can occur from the metrics of C_p, C_{pk}, P_p, and P_{pk}, as described in reference 4, which gives more depth calculations for these Six Sigma metrics.

The problems described in this golf illustration with regards to the understanding of someone's golf score performance is not unlike the reporting of the "*Y*" output or Key process output variable (KPOV) of many Six Sigma projects and other metrics within an organization. I contend that often using these types of metrics leads to playing games with the numbers. This is similar to what many companies have recently done in their accounting practices in an attempt to meet short-term goals. In addition, the implication of sample size and of how samples are taken is often not addressed with these metrics.

Alternative Scoring Approach

Let's consider an alternative to reporting our golf metrics. This strategy starts with a high-level view of our golfing process. In this approach, a *30,000 Foot-Level*[SM] metric is created using what I will call infrequent sub-grouping/sampling. With this strategy, we create a sampling plan that considers operator-to-operator, machine-to-machine, raw material lot-to-lot, shift-to-shift, and other similar factors as noise to the overall output.

This strategy is in alignment with Deming's position that if performance is poor, 94% of the time it is the system that must be modified for improvements to occur. If you agree with this categorization of these variables as common cause noise variables, many standard procedures for creating control charts can lead to the wrong activity. That is, many current control charting procedures lead to fire-fighting common cause variability (such as day-to-day raw material differences) as though this were special cause (1).

Let's now consider that we collected data on 20 rounds of golf and we plotted the average number of shots above par in an *XmR* control chart; i.e., a *30,000 Foot-Level* view of our golf game, which is shown in Figure 1. I chose this procedure of reporting, as opposed to plotting the number of shots above par for a golf course, since I wanted to have the flexibility of describing process/capability at the golf-hole level, in addition to the overall course level.

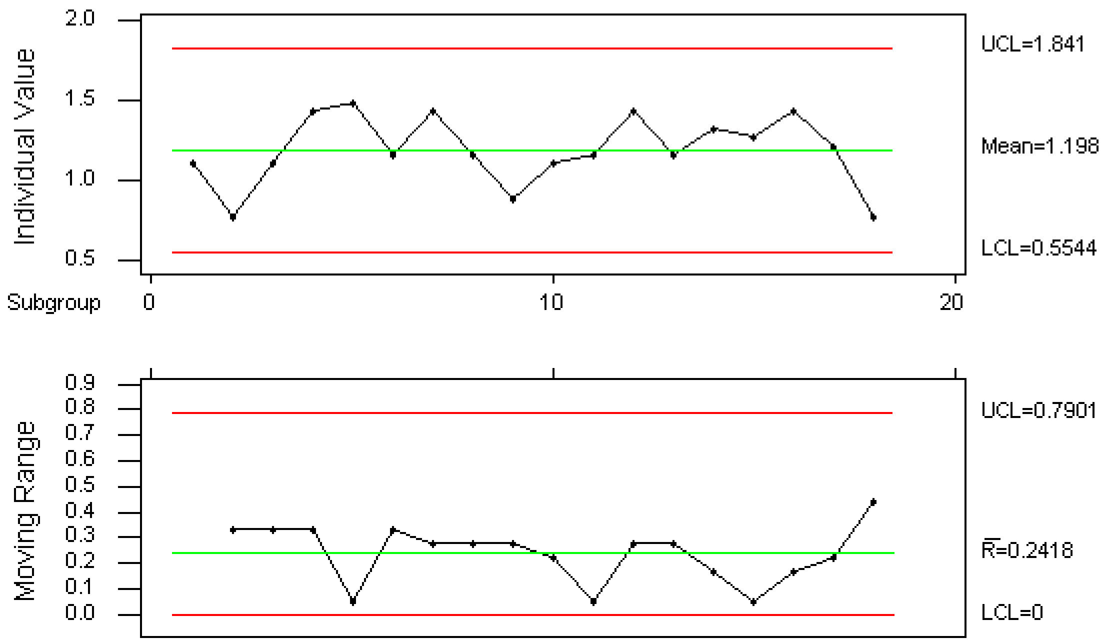
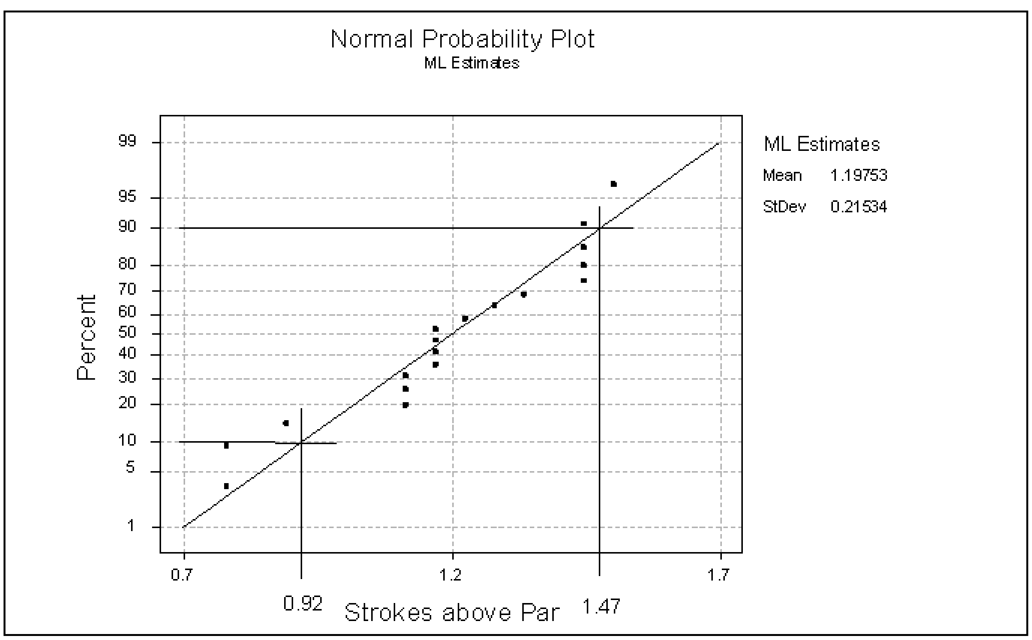


Figure 1

The center line of this plot is a best estimate for the mean response of the golfing process. Since the process is in control, it is predictable. That is, we can expect to shoot approximately 1.198 strokes per hole above par for 18 holes of golf, unless something different occurs. The normal probability plot shown in Figure 2 describes what we can expect 80% of the time when we play 18-holes of golf. This probability plot is a best-estimate that considers the variability between courses, course conditions, and our golf swinging consistency.



We could also describe an estimated process capability/performance of our golf game as an 80% (or some other percentage) occurrence rate. An 80% interval could be expressed as either:

- 0.92 – 1.47 average per hole strokes above par for an 18-hole course
- 89 (0.92x18 + 72) strokes to 98 (1.47x18 + 72) strokes for an 18-hole, 72-par course

I believe that these metrics provide a more understandable description of our golfing capability/performance than the sigma quality level and other Six Sigma metrics described above.

This *30,000 Foot-Level* view of our golf game provides a good baseline from which we can build strategies on how we should improve our game. For example, we can build a Pareto chart that shows the distribution of long-shots, short-shots, and putts. This Pareto chart can indicate where practice efforts should focus.

Other Applications

When companies are implementing Six Sigma in their organizations, it is important to align measurements and improvement activities to the overall needs of the business. Organizations need to create a set of balanced cascading measurements leading to meaningful activities. This gets organizations out of the fire-fighting mode. At Smarter Solutions, Inc. we call this approach *Smarter Six Sigma Solutions*^{SM} (S^{4} ^{SM}) or *Integrated Enterprise Excellence (IEE*). With this approach, *30,000 Foot-Level* operational metrics are aligned with *Satellite-Level*^{SM} metrics, which track the business as a process.

With this S^4/*IEE* approach, organizations can create strategic plans that are aligned with the improvement needs of the *Satellite-Level* metrics. Organizations can then create specific S^4/*IEE* projects that are targeted to these plans. These improvements can then be tracked and quantified at the *30,000 Foot-Level*. Successful projects will experience a statistically significant change in the process capability/performance at the operational level, which would have a positive impact on the *Satellite-Level* metrics.

It sounds great to create and then be able to compare sigma quality level metrics for projects. However, I believe that this approach can lead to the wrong set of activities and to playing games with the numbers. I wonder how many of the companies that are now having financial problems would still be profitable if they had followed the basic strategy described above instead of playing games with the numbers.

References

1. *Implementing Six Sigma*, Forrest W. Breyfogle III, Wiley, 1999.

2. *Managing Six Sigma*, Forrest W. Breyfogle III, James M. Cupello, and Becki Meadows, 2001.

3. *Wisdom on the Green: Smarter Six Sigma Business Solutions*, Smarter Solutions, Inc. 2001.

4. "Golf Scores and Six Sigma Metrics: Confusion and Resolution, part 2", Forrest W. Breyfogle III

5. *Statistical Process Control (SPC) Reference Manual*, Chrysler Corporation, Ford Motor Company, General Motors Corporation, AIAG, 1995.

Author

Forrest W. Breyfogle III is CEO of Smarter Solutions, Inc., which he founded in 1992. Smarter Solutions® is an international company that *wisely* integrates Six Sigma, Lean, and Theory of Constraints, helping customers improve both their bottom-line and customer satisfaction. Breyfogle was the Subject Matter Expert (SME) for a Six Sigma benchmarking study conducted by APQC. He earned a BS degree in mechanical engineering at University of Missouri at Rolla and a MS degree in mechanical engineering at the University of Texas. He is an ASQ Fellow. He can be reached at forrest@smartersolutions.com and 512-996-8288.