

Michel Baudin's Blog

Ideas from manufacturing operations

Is SPC Obsolete? (Revisited)

Six years ago, one of the first posts in this blog — [Is SPC Obsolete?](#) — started a spirited discussion with 122 comments. Reflecting on it, however, I find that the participants, including myself, missed the mark in many ways:

1. My own post and comments were too long on what is wrong with SPC, as taught to this day, and too short on alternatives. Here, I am attempting to remedy this by presenting two techniques, *induction trees* and *naïve Bayes*, that I think should be taught as part of anything reasonably called statistical process control. I conclude with what I think are the cultural reasons why they are ignored.
2. The discussions were too narrowly focused on *control charts*. While the [Wikipedia article on SPC](#) is only about control charts, other authors, like [Douglas Montgomery](#) or [Jack B. Revelle](#), see it as including other tools, such [scatterplots](#), [Pareto charts](#), and histograms, topics that none of the discussion participants said anything about. Even among control charts, there was undue emphasis on just one kind, the XmR chart, that Don Wheeler thinks is all you need to [understand variation](#).
3. Many of the contributors resorted to the argument of authority, saying that an approach must be right because of *who* said so, as opposed to *what* it says. With all due respect to Shewhart, Deming, and Juran, we are not going to solve today's quality problems by parsing their words. If they were still around, perhaps they would chime in and exhort quality professionals to apply their own judgment instead.

[Example: Final Inspection and Failure Analysis](#)
[Sequencing For Fast Identification Of Defectives](#)
[Failure Analysis Of Defectives](#)
 [The Quality Professions's Body of Knowledge](#)
[Conclusions](#)
[Acknowledgements and Further Reading](#)

Example: Final Inspection and Failure Analysis

At the end of [Lean Assembly](#), I devoted a chapter to a subject that is generally taboo in the Lean literature: inspections, test and rework operations. Such operations may not be *supposed* to exist but they do, in manufacturing processes ranging from socks to car engines and integrated circuits, and they won't be eliminated by next week. Let's assume the following context:



At each operation *Op. 1*, ..., *Op. n*, the production team does everything it knows how to avoid damaging

workpieces and passing defectives onwards. Nonetheless, *Final Test* still uncovers a few defectives that are passed on to *Failure Analysis* to identify the origin of the defects in the process and issue a report to the team in charge. The tools discussed below address operational aspects of both *Final Test* and *Failure Analysis* and neither is part of the standard SPC toolkit as described in the literature and taught in certification courses.

Sequencing For Fast Identification Of Defectives

If you need a final check, you need effective tests or inspections, the design of which is specific to each product. On the other hand, there are general principles on how to run a test and inspection operation, like sequencing by a decreasing *figure of merit* like the percentage of rejects found per unit time. It is a method I was taught early on, which I later realized applies only to tests or inspections that are independent, like paint scratches and leaks. I later realized that it method did not work for tests that were *not* independent, such as the ability of a microprocessor to work at different clock speeds. If it won't work at 1 GHz, there is no need to check it at 2GHz.

For the general case, I found a method called induction, where the outcome of each test or inspection takes you down a different branch of a decision tree, chosen for maximum information gain. It was pioneered in 1972 in health care, applied to the diagnosis of kidney diseases using 22 different tests. It reduced to 7 the mean number of tests needed to arrive at an answer.



Ross Quinlan

To implement this technique in those days, researcher Monica Bad probably had to start from the math and write FORTRAN code on punch cards to run on something like an IBM 370; today, you can download various free versions of a package called “C5.0” developed in Australia by [Ross Quinlan](#) and run it on your laptop.

Induction trees are just as applicable to the sequencing of industrial tests on outgoing products as to medical tests but, in the 45 years since Monica Bad first used it, somehow it never found its way into the literature or professional courses on manufacturing quality. I would love to be wrong about this and welcome any contradictory input. If I am correct, why? Induction trees are practical, proven, understandable, and capable of improving operations that quality professionals are involved with or responsible for.

Failure Analysis Of Defectives

At *Final Test*, your objective is to prevent defectives escaping to customers. It is a filter. The units that pass are shipped, and you are left with the defectives. However you dispose of them, you want to establish the cause of their defects and the point in the process where they appeared so that you can send problem reports to the right recipients for action. That's *Failure Analysis* and its output is this information, not the defective units from which it is retrieved. The units delivered to *Failure Analysis* are a trove of problems to be mined, and this operation's productivity is measured in problems found per unit time.

Time is of the essence. If it takes you four months to issue a report in a fast-changing industry, the information in it is obsolete; a week, and the information is still relevant but more defectives will have been produced while waiting; an hour, and it is both relevant and timely.

There are many techniques you can use to organize this work — including a more detailed version of the

induction trees described above — but I would like to focus on another one called [Naive Bayes](#), that is equally absent from the literature on quality, even though it has been applied to [failure analysis for hard disk drives](#).

Assuming engineers have conducted an FMEA (Failure-Mode-Effect Analysis) on the product, you have a list of failure modes and observable effects, or symptoms. You may also have the relative frequency of each effect on units with that failure and on units without it. Now you are working your way back to probabilities of failure modes given a set of symptoms.



Thomas Bayes

Bayes is Naive Bayes refers to Thomas Bayes, an 18th-century British clergyman who came up with a formula to transform the probabilities of symptoms given a cause to the probability of a cause given the symptoms. We know that if a unit is missing a gasket, there is a 60% probability that it will leak, 45% that it will overheat, and 30% that it will make a rattling noise. Bayes formula turns this around to tell us that, if it leaks, overheats and rattles, it has, say, a 70% probability of missing a gasket.

The method is called *Naive* because it treats the symptoms as if they were independent. This means that the probability of having several symptoms simultaneously is the product of probabilities for each:

$$\text{Prob}(\textit{Leak and Overheat and Noise}) = \text{Prob}(\textit{Leak}) \times \text{Prob}(\textit{Overheat}) \times \text{Prob}(\textit{Noise})$$

We know it's not true, but it simplifies the calculations and the method has been found effective in spite of this flaw.

Given the symptoms we have from the functional tests, Naive Base will give probabilities that they are due to a missing gasket, a loose fastener, or a cracked casting, etc. You can then use these probabilities — in combination with other information — to decide which one to investigate first. The other information includes, for example, whether the method to confirm a cause is destructive and how much time it takes.

[Bayes formula](#) has been around for almost 280 years and has served as the basis for many analysis techniques collectively known as *Bayesian statistics* that are used today in areas ranging from spam filtering of email to the search for shipwreck survivors. They are not completely ignored in [Juran's Quality Control Handbook](#): the latest edition, from 2016, references them on pages 612 and 620 — that is, in 2 of its 992 pages. On the other hand, there is no mention of them at all in Douglas Montgomery's [Introduction to Statistical Quality Control](#) or in Pyzdek and Keller's [Six Sigma Handbook](#).

In this blog, simple Bayesian methods were applied in the discussion of [Acceptance Sampling In The Age Of Low PPM Defectives](#) to questions that cannot be answered otherwise, like “When all previous deliveries of an item from a supplier have been defect-free, what is the probability that the *next one* will be defect-free?” If you google “Bayesian quality,” you only encounter a few academic papers.

The Quality Professions's Body of Knowledge

It's 2017, and, in its handling of data, the quality profession still relies exclusively on a body of knowledge frozen at the level of what could practically be used and was academically recognized in 1945. Induction trees are based on information theory, which was created by Claude Shannon in 1948.

Bayesian methods existed before 1945 but were out of fashion. Shewhart, in his own words, was an enthusiastic adopter of the developments in statistics from 1900 to 1920. These developments, however, were

based on a philosophy now called “frequentist” that did not consider the Bayesian approach legitimate. The tables have turned today, but, without taking sides in this debate, one clear effect was that Shewhart had no exposure to Bayesian statistics and neither did his disciples.

These explanations are at the level of a first why: a body of knowledge frozen in 1945 cannot contain concepts invented or recognized as valid afterward. It begs a second question: why was this body of knowledge frozen in 1945?

World War II was a period of rapid innovation in manufacturing in the US, with [TWI](#) introduced to train a new workforce of highly motivated women, a moving assembly line for bombers in Willow Run, Liberty Ships built by arc-welding vertical hull slices,... and the invention of new statistical methods in quality.

The *aftermath* of World War II, on the other hand, rolled back much of this progress, with accountants taking over US manufacturing and starting it down the path of focus on financial metrics. By the late 1950s, instead of working to improve quality, some appliance companies had “reliability” departments engineering products to fail as soon as the warranty expired.

In a previous post on [Lean accounting](#), I cited the 1953 movie [Executive Suite](#), in which the R&D manager and the chief accountant of a furniture manufacturer compete for the job of CEO in a board meeting. In the movie, the R&D manager wins over the directors with a passionate speech about products, quality, and pride in workmanship; in real life, on the other hand, the accountants won.



Perhaps, an environment where quality was not valued and professionals were fighting to prevent backsliding was less conducive to keeping up with the latest development in what is now called data science than to a hardening of the contemporary body of knowledge into dogma.

It's been 60 years and, since the early 1980s, the pendulum in US manufacturing has swung back towards an emphasis on quality but the *statistical* body of knowledge of the quality profession is still stuck in 1945. In management habits, effects endure long after their causes are gone.

Conclusions

The world of data science is faddish, and, for the quality profession to follow every fad would be as questionable as remaining stuck in 1945. Induction Trees and Naive Bayes are basic workhorses of data science, and not particularly fashionable.

[Neural Nets](#), on the other hand, are fashionable but it doesn't mean they are an appropriate tool to identify the cause of quality problems. By feeding a Neural Net enough examples, you can train it to recognize a handwritten “8.” You can easily gauge its ability to do so, and you don't need to understand *how* it manages this feat. Banks run into problems, however, when they apply this technique to credit risk assessment because Neural Nets don't explain why they recommend granting or denying a loan. Likewise, Neural Nets don't have

much of a role in quality problem solving, as humans need to understand and validate the reasoning behind the identification of a root cause.

What the quality profession needs is *technology watch*. It needs to evaluate tools systematically on a continuing basis for their validity and relevance, for the competitive advantage their application may bring, and for the risk of being left behind that ignoring them carries.

Acknowledgements and Further Reading

In alphabetical order, the contributors to the [original discussion](#) were [Cornelio Abellanas](#), [RW Ainsworth](#), [Sanda Albu](#), [Jagu Barot](#), [Kicab Castaneda-Mendez](#), [Mike Clayton](#), [Bryan Coats](#), [Paul Coverstone](#), [Gagandeep S. Datta](#), [Bob Doering](#), [Stephen R Edwards](#), [Wayne Fisher](#), [Attila Fulop](#), [Mark Graban](#), [Andrew Gwie](#), [Eric Hansen](#), [Greg Hjelmeland](#), [Tim Higgins](#), [Joseph \(Joe\) E. Johnson](#), [Roland Johnson](#), [Zaheer Khan](#), [Glen Knight](#), [Bob Mallett](#), [Lasse Marjanen](#), [Sid Martin](#), [Bob Matthew](#), [Hormoz Mogarei](#), [Al Onyeagwa](#), [Drew Peregrin](#), [Randy Pound](#), [Robert Reid](#), [Steve Ruegg](#), [Mikel Sanchez](#), [David \(Dave\) Scharrer, Jr.](#), [Karen Schultz](#), [Dan Strongin](#), [James Thornton](#), [Shashank M. Vaislay](#), [John Vandenbemden](#), and [John Witcher](#).

The literature on manufacturing quality that I cited in this post includes the following:

- [Juran's Quality Handbook](#), Joseph A. Defeo
- [Introduction to Statistical Quality Control](#), Douglas Montgomery
- [The Six Sigma Handbook](#), Thomas Pyzdek & Paul Keller
- [Understanding Variation](#), Don Wheeler
- [Quality Essentials: A Reference Guide from A to Z](#), Jack B. ReVelle

The only book I found to consider that there may be more than one characteristic of interest to measure on a workpiece is Peihua Qiu's [Introduction to Statistical Process Control](#). It has several chapters on *multivariate* control charts.

A popular means of learning data science is the [course concentration](#) on the subject offered by Coursera since April, 2014. It is a series of 10 online courses taught by three biostatistics professors from Johns Hopkins University. Attendance statistics are not easily available. According to [simplystats](#), as of February 2015, 1.76 million students had signed up. 71,589 had completed one course. 976 had completed the first 9 courses, and 478 the tenth, which is a "Capstone project."

Most of the literature show to use the tools with R, in cookbook fashion, without any theoretical justification. R itself, however, is not a tool I would recommend to any manufacturing professional I know. It's a programming language and requires users to think like software engineers. The only practical way to use it in a manufacturing environment is through applications developed by specialists.

The analytical tools are described in John Mount and Nina Zumel's [Practical Data Science with R](#). For data preparation, visualization and project documentation, see Hadley Wickham and Garrett Grolemund's [R for Data Science](#).

[#SPC](#), [#StatisticalProcessControl](#), [#DataScience](#), [#Quality](#)

Share this:



Like this:

Loading...

Related

[Is SPC obsolete?](#)

December 27, 2011

In "Technology"

[Why Many Japanese Organizations Think Six Sigma Is A Joke | Ian Moore |](#)

[Nipponica](#)

October 14, 2017

In "Blog clippings"

[Why Your Lean and Six Sigma Improvement Efforts Aren't](#)

[Driving Better Results | IndustryWeek | John Dyer](#)

May 28, 2015

In "Press clippings"

October 11, 2017

 [Leave a reply](#)

[« Previous](#)

[Next »](#)

Leave a Reply

Your email address will not be published. Required fields are marked *

Comment

Name

*

Email

*

Website

Post Comment

☐ Notify me of follow-up comments by email.

☐ Notify me of new posts by email.

Lean Leadership Summit 2016

About the author



[Michel Baudin](#)

Fascinated with the art of making things, Michel is working to improve it. Trained in engineering and applied math, he got his feet wet in production in the early 1980s, and later apprenticed under master Japanese consultant Kei Abe for eight years, starting his own group in 1996. He has been consulting since 1987, teaching courses and writing technical books. He intends to keep working with like-minded partners in the Takt Times Group and contributing improvements in the management and technology of manufacturing as a consultant, trainer, and writer.

Personal Links

[The takt times group](#)

[Blog about manufacturing](#)

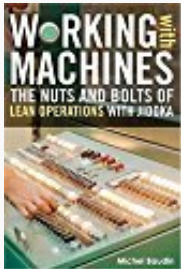
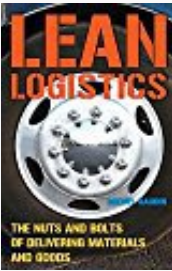
[Press and blog clippings about Lean](#)

Verified Services

My books



[Shopping Cart](#)



[My Bookstore](#) | [Shopping Cart](#)

To find out more:

Name (required)

Email (required)

Phone number

Website

Your question (required)

Submit »

Videos



"Introduction to Lean 2013"
by Michel Baudin

Follow Blog via Email

Enter your email address to follow this blog and receive notifications of new posts by email.

Join 2,494 other subscribers

Follow

Most popular posts and pages

[Is SPC Obsolete? \(Revisited\)](#)

[Deming's point 2 of 14: Adopt the new philosophy...](#)

[What is Karakuri Kaizen?](#)

[Why 5S fails](#)

[Deming's Point 3 of 14 - Cease dependence on inspection to achieve quality...](#)

Recent Posts

[Now It's Humans Assisting Robots | Sheelah Kolhatkar | The New Yorker](#)

[Why do we call it "value stream"?](#)

[How Hard Is It To Create Flow?](#)
[More Sophisticated Graphics In Today's New York Times](#)
[Objections To The History Of Quality As Told On The ASQ Website](#)

Categories

[Announcements](#)
[Answers to reader questions](#)
[Asenta selection](#)
[Blog clippings](#)
[Blog reviews](#)
[Book reviews](#)
[Data science](#)
[Deming](#)
[Events](#)
[History](#)
[Information Technology](#)
[Laws of nature](#)
[Management](#)
[Metrics](#)
[News](#)
[Organization structure](#)
[Personal communications](#)
[Policies](#)
[Polls](#)
[Press clippings](#)
[Technology](#)
[Tools](#)
[Uncategorized](#)
[Web scrapings](#)

Follow me on Twitter

[My Tweets](#)

My tags

[5S](#) [Automation](#) [Autonomation](#) [Cellular manufacturing](#) [Continuous improvement](#) [Deming](#) [ERP](#) [Ford](#)
[Government](#) [Health care](#) [industrial engineering](#) [Industry 4.0](#) [Information technology](#) [IT](#) [jidoka](#) [Kaizen](#)
[Kanban](#) [Lean](#) [Lean assembly](#) [Lean Health Care](#) [Lean implementation](#) [Lean Logistics](#)
[Lean management](#) [Lean manufacturing](#) [Logistics](#) [Management](#)
[Manufacturing](#) [Manufacturing engineering](#) [Metrics](#) [Mistake-Proofing](#) [Poka-Yoke](#) [Quality](#)
[Six Sigma](#) [SMED](#) [Standard Work](#) [Strategy](#) [Supply Chain Management](#) [Takt](#) [Takt time](#) [Toyota](#) [Toyota](#)
[Production System](#) [TPS](#) [Training](#) [VSM](#) [WCM](#)

Michel Baudin's feed

 [RSS - Posts](#)

 [RSS - Comments](#)

[View Full Site](#)

Proudly powered by [WordPress](#)

Follow

Follow Michel Baudin's Blog

Get every new post delivered to your Inbox

Join other followers:

[Sign me up!](#)