Engineering Statistics

(Entry for the 2nd Edition of *The Encyclopedia of Statistical Science*)

by

Stephen B. Vardeman*
Statistics and Industrial and Manufacturing Systems Engineering Departments
Iowa State University

Abstract

In this entry we seek to put into perspective some of the ways in which statistical methods contribute to modern engineering practice.

Engineers design and oversee the production, operation, and maintenance of the products and systems that under-gird modern technological society. Their work is built on the foundation of physical (and increasingly biological) science. However, it is of necessity often highly empirical, because there simply isn't scientific theory complete and simple enough to effectively describe all of the myriad circumstances that arise even in engineering design, let alone those encountered in production, operation, and maintenance. As a consequence, engineering is an inherently statistical enterprise. Engineers must routinely collect, summarize, and draw inferences based on data, and it is hard to think of a statistical method that has no potential use in modern engineering.

The above said, it *is* possible to identify classes of statistical methods that have traditionally been associated with engineering applications and some that are increasingly important to the field. This encyclopedia entry will identify some of those and indicate their place in modern engineering practice, with no attempt to provide technical details of their implementation.

Statistics and Measurement

It is nearly self-evident that if one is to design, build, and run technological systems and devices, one must be able to measure. And particularly when new systems are on the "leading edge" of technology, how to measure can be a serious issue. While statistics offers no direct help in suggesting physical mechanisms to exploit, it does offer important methodologies for quantifying and improving the quality of measurements. (The long-standing presence of a statistical group in the US National Institute of Standards and Technology testifies to this importance. And in passing we remark that this group's

* The financial support of the Deutsche Forschungsgemeinschaft (SFB 475, "Reduction of Complexity in Multivariate Data Structures") through the University of Dortmund is gratefully acknowledged. The author thanks Max Morris, Bill Meeker and Huaiqing Wu for their comments on an earlier draft of the entry.

online *NIST/SEMATECH e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook/ [8] provides widely accessible current information on statistical methods useful to engineers, in measurement problems and beyond.)

One fundamental class of statistical problems in engineering measurement concerns the quantification of measurement precision (variability) and identification of important contributors to *random* measurement error. Random effects models and corresponding estimation of variance components are useful statistical tools in these endeavors. The particular context where several different technicians will use a measurement device and there is interest in quantifying respectively both a baseline "repeat measurement of the same item by a single technician" variance component and a "between technicians" variance component, is known as the "gauge repeatability and reproducibility" (gauge R&R) problem in engineering and quality control circles. (See, for example, Vardeman and Van Valkenburg [15].)

A second fundamental type of statistical problem in engineering measurement is that of adjusting the output of a measurement device to agree (on average) with that of a state-of-the-art or "gold standard" device (or some fixed standard value). This is the calibration problem, and calibration is aimed at the reduction of *systematic* measurement error or bias, i.e. the improvement of measurement accuracy. (Osborne [9] provides a nice review of statistical methodology appropriate in calibration problems and available through the early 1990's.) Various forms of regression analysis are common tools in this enterprise and it is worth noting that since most often one regresses "new" measurements on gold-standard measurements or standard values, transformation of measurements to standard values involves an "inverse prediction." Accordingly, typical confidence limits for a standard value corresponding to a given new measurement come from inversion of families of prediction limits for a new measurement not contained in a calibration data set.

As measurements themselves become more complicated (for example moving from single real numbers, to approximate chemical spectra produced by mass spectrometers or to probe paths and approximate coordinates of "touch points" in space produced by coordinate measuring machines) the potential for application of methods of multivariate analysis and functional data analysis becomes clear. The recognition of other real characteristics of measurements like their digital or rounded nature (their imperfect resolution) point to the need for increasingly sophisticated statistical modeling and inference methods. And the need for efficient and effective data collection in measurement studies suggests the relevance of methods of statistical experimental design in this area.

Statistics and Empirical Optimization

Engineering practice is subject to tremendous economic pressure. Engineering designs must be produced quickly and cheaply, and the products designed must be both highly effective and cheap to make, while the systems that produce them must be made to run at high efficiency. All of this (and the lack of comprehensive scientific knowledge adequate to describe and evaluate the implications of every possible engineering alternative) implies the engineering need for methods of empirical optimization.

This need has long been recognized and addressed in the traditional engineering statistics teaching emphasis on experimental design and analysis. Methods of factorial and fractional factorial design and analysis, and so-called “response surface methodology” (empirical optimization strategies based on statistical experimental design and low order multivariate polynomial regression) have long had their place. (See, for example, Box and Draper [2].) Until fairly recently, the bulk of applications of these methods has probably been to the improvement of existing physical production processes. But statistical tools are increasingly finding application “upstream” in engineering research and design, even in contexts where “data” are not measurements on real physical systems, but rather outputs of sometimes expensive-to-run computer codes for mathematical models of potential systems. This last possibility goes in the statistical literature under the name of design and analysis of “computer experiments” and its methodology has connections to both classical experimental design theory and modern spatial statistics. (See, for example, Santner, Williams, and Notz [11], Sacks et. al [10], and Currin et. al [3].)

Statistics and Empirical Product and Process “Robustification”

Related to, but not equivalent to, the notion of optimization is that of making a product or process “robust”/able to function appropriately across a wide variety of environments and over time. The engineering need for methods of statistical experimental design and analysis to support the empirical search for robust product and process configurations was first effectively emphasized in the west in the mid 1980’s by Genichi Taguchi. Since that time, a sizeable statistical literature has grown up in “Taguchi methods.” This includes advances in both special forms of highly fractional experimental designs (purposely chosen to vary rather than control “noise”/environmental factors) and in modeling and inference for contexts where both mean and variance of response change with levels of factors whose levels are to be set in choosing a product or process design. (The panel discussion of Nair, et. al [6] is a basic early reference in this area.)

Statistics and Process Monitoring

One of the main subject areas traditionally clearly identified as part of “engineering statistics” is “statistical process control.” The traditional tacit assumptions have been that the main application of the methodology was to production, the most common tools were Shewhart control charts, and the fundamental idea (dating at least to Shewhart and the 1920’s) was that production equipment should minimally behave as if it were “stable” (consistent up to iid random variation). The traditional techniques of statistical process control have thus been aimed at detection of process change for iid processes.

In the past decade or two, standard simple tools of statistical process monitoring have found application in many business contexts beyond the engineering domain (finding prominent places in Total Quality Management and Six Sigma programs for business process improvement), and engineering applications have broadened considerably (for example including regular use in the ongoing monitoring of the stability of measurement processes and the condition of mechanical equipment in preventative maintenance). In statistically sophisticated circles, theoretically superior alternatives to Shewhart charts (particularly CUSUM schemes and their variants) have been developed and promoted, though evidence of widespread implementation of these is lacking. And there has been some recent work in engineering process monitoring taking a broader (than iid/white noise) view of what is acceptable null process behavior (that could perhaps be better informed by closer ties to the economic time series literature and its work on change detection).

The usual engineering meaning of the phrase "process control" is something different from the monitoring/detection-of-fundamental-change technology of statistical process monitoring. Most engineers (particularly mechanical, chemical and electrical engineers) understand the terminology to refer to methods (often based on quite sophisticated mathematical modeling) of ongoing adjustment of inherently dynamical systems. There have been some efforts on the part of statisticians to provide integrations of methods of "engineering control" and "statistical control" (see for example Tucker, Faltin and Vander Wiel [12]). These have had limited impact in engineering practice, due in no small part to difficulty statisticians face in acquiring the very specialized and case-by-case subject-matter process knowledge and background in control theory needed to first understand real engineering control systems.

Statistics and Process Characterization

Much of modern engineering is done in contexts where multiple devices or systems of a given design will be made. (While one-of-a kind engineering applications exist, they do not predominate.) As such, various forms of data-based process characterization are important to engineers. In some situations simple estimation of process parameters or functions of those (often called "capability indices") suffices. But it is also common to want data-based limits for likely values of either single new process outcomes or the bulk of all future process outcomes. So there is a long tradition of the use of prediction and tolerance intervals in engineering statistics (that, curiously enough, is largely unparalleled in other application areas).

Statistics and Reliability/Life Data Analysis

The issue of engineering reliability is that of how long a device or system can be expected to function before some kind of partial or complete failure. Where reliability is to be measured based on observed lifetime data, statistical methodology for single lifetime distributions like the Weibull, lognormal, and log-logistic models has been standard in engineering applications. Where systems are "repairable," inference methods for point processes (for example, renewal processes, and where there is the possibility of reliability growth or degradation, nonhomogenous Poisson processes) have found applications. There is some commonality of statistical methodology between this area and the area of medical survival analysis, and methods recognizing the presence of various kinds of censoring in data collection are essential. A comprehensive reference in the general area of life data analysis is Meeker and Escobar [4] and Meeker and Escobar [5] provide a very broad discussion of ways in which statistical thinking and tools can contribute to reliability engineering efforts, from the early design stage through the analysis of field warranty data.

Two emphases that are increasingly important in engineering life data analysis are the use of degradation data and the planning and analysis of accelerated life tests. That is, where the failure of a device or system can be characterized in terms of the value(s) of one or more measurements and it is possible to model and collect informatio n on the evolution of these over time, there is the possibility of making inferences superior to those based only on simple times to failure. (See, for example, Chapter 13 of Meeker and Escobar [4]). And in contexts where engineers aim to develop highly reliable products whose typical lifetimes must exceed the length of any sensible product development cycle, the only means of empirical testing of prototypes is to subject them to environments more severe than a normal-use environment and try to extrapolate normal-use life characteristics from "accelerated stress" life characteristics. Methods of statistical inference (lifetime model regression techniques) and study planning (experimental design optimization tools for lifetime regression models) have proved helpful in making the engineering work more systematic and efficient, particularly in applications in the electronics industry where good simple models exist for the effects on lifetime of typical stress factors, and per-unit test costs are relatively low. (Nelson [7] and Chapters 17 through 20 of Meeker and Escobar [4] are standard references here.)

Statistics and (Sampling) Inspection and Acceptance Sampling

In production contexts, there is typically a need to verify that a particular item or a product stream or lot of items meets performance/conformance goals of the producer and/or a consumer. Where one admits that individual conformance assessments are subject to uncertainty (possibly, as in Albers, Arts, and Kallenberg [1], because only indirect measurement of primary performance characteristics is possible or desirable) or only some of all items of interest will be inspected, statistical methods become useful. Traditionally, this was evident in the prominent place of methods of acceptance sampling in the engineering statistics literature. While this prominence has (appropriately) waned (see Vardeman [14] and Vander Wiel and Vardeman [13] in this regard), there remains an important role for statistics in the general area of the collection and interpretation of product inspection data.

Probabilistic Analyses

While most standard engineering analysis is deterministic, there are some areas where stochastic models are used and even fundamental. To the extent that many engineering statisticians know a fair amount of probability, they have the potential to contribute to stochastic analysis in engineering. Some of the engineering contexts in which the usefulness of stochastic modeling is well-established include: tolerancing problems, system reliability prediction and retrospective "fault-tree" analysis, project planning and analysis, production process modeling and queuing, inspection efficacy in "nondestructive evaluation," and signal processing. In some of these contexts, analytical methods are well developed and common. In others, Monte Carlo methods provide the primary path forward to improved engineering insight.

References

[1] Albers, W., Arts, G.R.J., and Kallenberg, W.C.M. (1999). "Test limits using two or more correlated product characteristics," *Technometrics*, 41, 153-165.

[2] Box, G.E.P., and Draper, N.R. (1987). *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York.

[3] Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D. (1991). "Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments," *Journal of the American Statistical Association*, 86, 953-963.

[4] Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, John Wiley & Sons, New York.

[5] Meeker, W.Q. and Escobar, L.A. (2004). "Reliability, the other dimension of quality," *Quality Technology\& Quality Management Journal*, 1, - .

[6] Nair, V.N., Abraham, B., MacKay, J., Box, G. , Kacker, R. , Lorenzen, T. , Lucas, J. , Myers, R. , Vining, G. , Nelder, J. , Phadke, M. , Sacks, J. , Welch, W. , Shoemaker, A. , Tsui, K. , Taguchi, S. , and Wu, C.F.J. (1992). "Taguchi's parameter design: A panel discussion," *Technometrics*, 34, 127-161.

[7] Nelson, W.A. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, John Wiley & Sons, New York.

[8] NIST/SEMATECH (No Date). *NIST/SEMATECH e-Handbook of Statistical Methods* [Online]. NIST. http://www.itl.nist.gov/div898/handbook/ [2003, October 20].

[9] Osborne, C. (1991). "Statistical calibration: A review," *International Statistical Review*, 59, 309-336.

[10] Sacks, J., Welch, W.J., Mitchell, T.J., and Wynn, H.P. (1989). "Design and analysis of computer experiments," *Statistical Science*, 4, 409-423.

[11] Santner, T.J., Williams, B.J., and Notz W.I. (2003). *The Design and Analysis of Computer Experiments*, Springer, New York.

[12] Tucker, W.T., Faltin, F.W., and Vander Wiel, S.A. (1993). "Algorithmic statistical process control: An elaboration," *Technometrics*, 35, 363-375.

[13] Vander Wiel, S.A. and Vardeman, S.B. (1994). "A discussion of 'all or none' inspection policies," *Technometrics*, 36, 102-109.

[14] Vardeman, S.B. (1986). "The legitimate role of inspection in modern SQC," *The American Statistician*, 40, 325-328.

[15] Vardeman, S.B. and Van Valkenburg, E. (1999). "Two-way random-effects analyses and gauge R&R studies," *Technometrics*, 41, 202-211.