Key Challenges for Statisticians in Business and Industry
Author(s): Gerry Hahn and Roger Hoerl
Source: *Technometrics*, Vol. 40, No. 3 (Aug., 1998), pp. 195-200
Published by: American Statistical Association and American Society for Quality
Stable URL: http://www.jstor.org/stable/1271171
Accessed: 18/01/2010 13:55

Editor's Note: The following article and discussion are a bit out of the mainstream for *Technometrics*, but they focus on issues of critical importance to many of our readers. The authors' original intent was to raise questions and ideas in the hope that the panelists might provide partial answers and reactions. The mix that has resulted is perhaps a little more symmetric, but each participant brought an interesting and valuable point of view to the discussion. (Diane Lambert's discussion was received after the authors had completed their summary.) The authors and I hope that you will find these comments useful in your own thinking about the future of statistics in business and industrial settings.

Key Challenges for Statisticians in Business and Industry

Gerry HAHN and Roger HOERL

GE Corporate Research and Development
Schenectady, NY 12301
(hahn@crd.ge.com, hoerl@crd.ge.com)

This discussion examines the exciting and changing role of the statistician in business and industry as we enter the 21st century. We describe the new environment—statistics without statisticians, new growth areas, escalated management expectations, and so forth—and the role changes it requires. We then present some key nontechnical and technical challenges and our view on how some of these should be addressed. We conclude by raising some specific questions to stimulate the panelists' comments.

KEY WORDS: Quality; Reliability; Role changes; Six Sigma improvement.

Statisticians in business and industry are experiencing substantial changes in their work environment, facing many exciting new opportunities, and urgently needing to address some major challenges. A candid review of where we are and where we should be going is in order.

We begin with some background on the changing work environment and the key ways in which these changes are having an impact on our roles. Next, we describe some of the general nontechnical, as well as technical, challenges facing statisticians in the new environment. We conclude by asking the panelists specific questions on how statisticians should respond to these challenges.

THEN AND NOW

To set the stage, we review a few of the important ways in which today's environment differs from that of the past, both technically and strategically. Recognition of these changes is an important prerequisite to understanding today's challenges.

Statistics Without Statisticians

Perhaps the most significant change in our environment is what has been called the "democratization of statistics" [analogous to the democratization of mathematics; see Vere-Jones (1995)]. It is, to a great degree, a consequence of the accessibility of statistical software. It has tremendous ramifications both for the routine use of existing statistical methods and for the development of new ones. Moreover, there appears to be a growing awareness of the value and need of quantitative and statistical methods in many segments of U.S. business and industry. Our employer, the GE Company, has, most enthusiastically, adopted the Six Sigma approach, originally pioneered by Motorola, as *the* major initiative of the company until the year 2000. Tens of thousands of people are being trained in statistical methods as part of this initiative. Perhaps most startling is that high-level managers are asking questions about the use of statistics in their organizations and are even receiving training themselves. Moreover, the fervor in our company is not limited to manufacturing—the traditional stalking ground of industrial statistics. Equally important are product design and financial transactions—both for nonmanufacturing businesses (e.g., finance, information, and broadcasting) and for the nonmanufacturing operations, such as billing and order entry, within services and manufacturing businesses. Although our experience is specific to GE, we see evidence of similar changes in other North American businesses—although, in some cases, such as the automobile industry, the interest might have peaked several years ago.

In part because of the relatively limited number of professional statisticians, most Six Sigma practitioners (colorfully identified as master black belts, black belts, green belts, etc.) are applying the methods without the benefit of much professional statistical advice. In fact, the vast majority were trained in statistical methods by someone other than a statistician. We see this trend continuing within U.S. business and industry as user-friendly software becomes even more generally available.

With regard to new techniques, the last decade or so has seen the introduction of various important, and well-recognized, new approaches for data analysis. A surprising number of these have been developed independently of the statistical community, frequently by engineers and computer scientists. These techniques include neural networks, fuzzy logic—and, more recently, data mining and Bayesian networks.

The obvious environmental change revealed by these situations is that neither practitioners nor researchers from other disciplines are asking our permission to do statistics. Often, we may not even be aware of what they are doing. This trend has caught statisticians off guard. Our profession has been inwardly focused for so long that it is now being bypassed by others, not only in the use of existing techniques but also in the development of new ones. In summary, the increased popularity of statistical methods has not necessarily resulted in increased popularity of statisticians.

New Growth Areas

As already suggested, another key change is that traditional areas of application of statistical methods, such as manufacturing and survey design, are no longer the major growth areas. Current growth areas include pharmaceuticals (primarily R&D), health care in general, environmental, and, perhaps most significantly, financial applications.

GE, for example, has a large ($40 billion in revenue in 1997) financial services business. Traditionally, statistical methods were used sparsely in this part of the corporation. This is changing very quickly! There is now tremendous demand for such things as developing statistical models for scoring credit applications, forecasting the residual value of leased cars or equipment, or suggesting an "optimal" price to bid for a municipal bond based on variables pertaining to the bond issue. These applications are just the tip of a very large iceberg. Unfortunately, relatively few statisticians have experience in these types of problems, notwithstanding the existence of an active business statistics community. It appears that integrating the continuous improvement focus of industrial statisticians with the financial experience of business statisticians could reap great benefits.

Escalated Management Expectations

There was a time when many statisticians resided in staff functions and lived in a world of relatively "benign neglect" by management. Although expected to support immediate corporate needs, they typically spent at least some of their working hours in conducting research of general benefit to their company and to the world as a whole. Much of the published research in statistics that was actually motivated by real problems emanated from this group. In recent years, many staff functions, especially those whose impact is not highly visible or required by government regulations (such as in pharmaceuticals), have been targets for corporate downsizing. To paraphrase a reviewer, business is no longer willing to have isolated specialists on their payrolls. Those that remain need to apply their time to customer projects for specific tasks performed. A key aim is to be responsive and vital to today's business needs and to be able to prove their contributions quantitatively. Thus, even though management has a better appreciation of the value of a quantitatively based approach in general, and statistics in particular, there is greater pressure than ever before for us to deliver tangible results quickly. A subtle, but important, issue this raises is that we can expect even fewer publications in the future from applied statisticians working in business and industry. Given that there is already serious concern within the profession that our publications often lack practical usefulness, this raises a significant issue. One possible way of addressing this issue may be by more extensive teaming in publications between statisticians in industry and those in academia.

Changing Focus of Quality Improvement Efforts

For much of our careers, quality improvement tended to be synonymous with massive inspection of finished product and then later with statistical process control (SPC) of manufacturing processes. Today, there is a passion, perhaps as a consequence of competitive pressures, for *proactive* quality improvement in many businesses. A major goal is to improve the design process, so that the resulting products do not require intensive improvement efforts after market introduction. This creates issues for statisticians because we typically have little knowledge of design processes, especially in the nonmanufacturing growth areas.

Speed as a Driving Force

The pace of everything has accelerated—as evidenced by such concepts as rapid new-product introduction, agile and flexible manufacturing, and just-in-time inventory. Thus, for statisticians to provide insights is not enough; our response must also be timely. This raises an issue because our profession generally has emphasized performing "ultimate" analyses over "appropriate" analyses.

Other Changes

We mention only briefly a variety of other changes that have significantly had an impact on industrial statisticians but that require little elaboration:

- The computer revolution: Statistical researchers need now focus on how to squeeze the most important information out of immense databases, instead of methods that are simple to analyze with limited or no computer resources.
- Organizational delayering: The fact that there are generally fewer layers of management means that statisticians have more opportunities to directly communicate with, and have an impact on, top management.
- Communications: Detailed written reports are being downplayed in favor of concise management-oriented presentations—often presented remotely via video or teleconferences. Statisticians need to adapt to these new forms of communication.

- Outsourcing and globalization (terms not yet recognized by our computer spell checker): Fewer and fewer products are being built in their entirety in one location—instead parts are obtained from numerous suppliers, and often the same product is built in various locations worldwide. This has significantly broadened the role of the "plant statistician."

Some Examples

To be more specific, we briefly describe some recent activities, by ourselves or our associates, that would have been impossible, or highly unlikely, say, a decade or two ago:

- Proposal of a system that remotely tracks and continuously updates the measured performance of a locomotive in the field via a Web site that can be accessed dynamically by responsible engineers and managers; the system also provides early warning of significant trends and impending problems so that these can be averted before they result in a field failure.
- Presentation to top management of how an advanced classification and regression-tree approach can help identify the most likely, and the most desirable, customers for a new insurance policy.
- Evaluation of commercial simulation software, making a proposal for a corporate license for the selected package to senior management, and holding a "train the trainers" session to company components.
- Meeting with a diverse group of managers from throughout Asia to discuss implementation of the Six Sigma initiative, training them in the basic tools, and consulting on their projects.
- Using a database involving millions of records, constructed for billing customers, to develop an algorithm to help determine appropriate authorization limits to credit-card holders.
- Serving on the staff of a business quality leader to help guide the business's quality efforts, with a specific focus on the Six Sigma initiative.

REQUIRED ROLE CHANGES

Broader Roles for Statisticians

Not surprisingly, the roles of statisticians in business and industry have been affected dramatically by these changes in environment. Therefore, those considering a career in statistics need, as a preliminary step, to do a sober and probing self-evaluation of their qualifications. More than ever before, business and industry require statisticians who have the insight to get at the real root causes of issues, both technical and nontechnical; general problem-solving and scientific thinking skills; a broad base of statistical and subject-material knowledge; the ability to learn quickly what they do not know; the skill to adapt such knowledge to the problem at hand; the confidence to work effectively with others in teams; the willingness to balance thoroughness with getting the job done on time; the stomach to work in a dynamic, high-pressure environment and to put in long hours when needed; outstanding communications and training skills; and the enthusiasm (and immodesty) to sell themselves and their ideas. More detailed attributes were discussed by Hoerl, Hooper, Jacobs, and Lucas (1993) and in the references they provided. In short, statistical knowledge is a necessary, but far from sufficient, condition for a successful career in business or industry. The traditional education of statisticians provides few of these skills beyond the knowledge of the statistical tools themselves.

Once on board, today's statisticians must strive to get involved in the strategically most significant projects—that is, those that combine importance to the company with a high likelihood of leading to a useful contribution. Unlike in academia or the industrial environment of the past, the technical sophistication of the problem is not the major criterion. Statisticians can no longer afford to sit patiently at their computers, waiting for someone to bring them data to analyze. In addition, with the proliferation of statistical software, simply performing a statistical analysis is no longer a marketable task—anyone with a laptop can do that. Statisticians must convincingly articulate how they can add significant value, and they generally need to sell their services to secure the needed funding to proceed.

Moreover, the statistician in business or industry must become a deeply immersed and proactively committed team member—rather than a consultant whose advice is sought only on occasion. In fact, the consulting role may be "outsourced" to a professor at a local university. Being a contributing team member requires a broad approach to problem-solving—that is, getting out of one's technical "box" to learn about other disciplines and often performing tasks that do not directly relate to one's narrow specialty. By contrast, engineers and business people tend to have a relatively broad background. All too often, statisticians know only statistics, mathematics, and, more recently, computer science. They would often have been better served if they had also obtained significant education, perhaps as undergraduates, in such areas as physics, engineering, or finance. One reason business and industry tend to view statisticians narrowly is that we have tended to choose a narrow role for ourselves.

If a statistician proves successful as a team member, what next? The profession must give more thought to what the broad career progression for statisticians might look like, beyond aspiring to manage a statistical consulting group. In contrast, engineers often begin by doing the nuts and bolts of engineering. After gaining experience, and demonstrating competency, some migrate to more visible tasks, such as moving from manufacturing to design or being accountable for major equipment installations. Some go into engineering management. Eventually, a few move to a truly strategic role, often called "Chief Engineer," in which they are expected to make recommendations to senior management on important technical issues, negotiate with key vendors, or provide final approvals on new designs. This role bears some similarity to Deming's (1986) concept of a "statistical leader." Although significant parts of Deming's thinking have been accepted and adopted by industry—and are among the driving forces behind the enthusiasm for quality improvement—his concept of a statistical leader has, by and large, not materialized. Moreover, although there are some exceptions, such as striving to become a partner in an actuarial firm, there is no well-defined career path for statisticians in business and industry that possibly culminates in a broad leadership role. We feel there should be. But what must we do to help make this happen?

Being Responsive to the World of Statistics Without Statisticians

What should be our unique role given the dramatic spread of statistics without statisticians? First, statisticians should accept this trend and welcome it. After all, many of us have been advocating the more universal application of statistical tools and concepts for a long time, perhaps based on the Japanese model. We should not be jealous of abdicating some of our "rights" and, in fact, should see it as an opportunity to do more challenging work. We can look to computer science for an analogy. At one time, software was developed almost exclusively by computer professionals. With the advent of more user-friendly computer languages, this has been less and less the case. Today's computer scientists, however, are not exactly on the unemployment line—they are just working at a higher level, such as making software development by the rest of us easier and more effective, ensuring compatibility of the components of a system, or planning long-term strategies to meet their organization's future information technology (IT) needs. Statisticians need to similarly expand their horizons.

Our unique value added to the organization of the future may include bringing an "improvement" mindset—based on proactive avoidance of problems—to the team and the ability to take a holistic approach to problem definition and resolution, as well as our critical quantitative skills.

KEY NONTECHNICAL CHALLENGES

Ensuring That the Democratization of Statistics Is Successful

While welcoming this change as beneficial, we have the somber responsibility to help ensure that it is successful and does not leave our profession with a bad name. Thus, some of the key issues we need to address are the following:

- How can we help ensure that the often limited training of practitioners in statistics best meets their needs—and that there is a proper emphasis on understanding basic concepts such as variability, the proper planning of investigations, and the use of graphical analysis tools over the rote teaching of statistical software? Our point is that being able to generate a regression model using JMP, Minitab, or Excel does not make one a competent modeler.
- How can we build an atmosphere of mutual trust so that practitioners and managers are eager to involve us in the most strategically important, and most technically complex, situations?
- How can we expedite the development of tools, often directed at large datasets, that are robust to potential misuse or misunderstanding in the hands of minimally trained practitioners?
- How can we help dismantle the sometimes self-imposed boundaries between us and other professionals and ensure that we are at the forefront of the development of the next "fuzzy logic" or "neural net" methodology? How can we foster more joint research between statisticians and other professionals? Similarly, how can we develop better communications links to make statistical researchers more aware of the real problems faced by other disciplines? Wanting to do this is no longer controversial—the key question is *how* to do it.

Developing the Appropriate Skills

Given the diversity of skills required to be effective in business and industry today, how can we most effectively develop these skills in the industrial statisticians of the future? Knowledge of statistical tools alone is just not responsive to today's needs. At the same time, we must agree that it is impossible to develop all the needed nontechnical skills, in addition to gaining solid knowledge of the statistical concepts, theory, and tools themselves, through formal courses in college. In fact, our task is compounded by the fact that, as our customers become more knowledgeable, it is essential for us to be ahead of them in our technical knowledge. All of this points to the critical challenge for our profession.

Helping Build Useful Databases

The tremendous improvements in computer technology have been an essential ingredient of many of the issues raised in the preceding comments. Clearly, we must leverage the IT revolution. Thus, there is a great opportunity to work with Management Information Systems operations in building new generations of systems for broad use that incorporate statistical thinking, graphics, and models to help management sift through the volumes of data that stream in. Business and industry's ability to generate data and, for example, make it generally accessible to a wide user community via the Web continues to outstrip our ability to transform such data into meaningful information. Our capabilities for conducting intensive data analyses by various ingenious and sophisticated methods continues to advance rapidly. At the same time, the cost of running experiments—one of our key weapons—has probably not changed radically. Thus, there is a greater premium than ever before on better use of the data that is, or can readily be, obtained routinely as a precursor to experimentation.

The problems are compounded by the fact that, very frequently, the data were collected for purposes other than analysis. Thus, in-process inspections are made not to assess the impact of process variation on product life but to make on-the-spot accept/rework/reject decisions, and the results are not fully recorded, if at all. Similarly, credit-card payment records are kept mainly for collection purposes—not to develop algorithms for determining credit limits. Thus, even today, the needed data are sometimes not recorded or kept on sheets or tracing tags that are then stored in file cabinets. Even when the data are computerized, the records may often be on those variables that were easiest to measure rather than most informative or are inconsistent over time, incomplete, or stored in different databases that do not readily communicate with one another. As a result, analysts spend a vast amount of time in cleaning up inadequate data. How can we sensitize students to the fact that, unlike problems in statistical textbooks, data are generally not delivered on the proverbial silver platter? And, most importantly, how can we help ensure effective up-front data acquisition? Perhaps a clue was provided by a reviewer who said that the key issue is to "get people to collect better data in recognition that they will be analyzing it themselves!"

Ensuring That Computers Are our Servants—Not our Masters

We have some concerns about our profession's current response to the IT revolution. We see too many similarities to the way in which mathematics began to dominate statistics in the 1970s. Although mathematics is the foundation of all the "hard" sciences, our profession has, in many authors' opinions, taken respect for mathematics to an extreme. Our journals have tended to emphasize exact mathematical solutions to unrealistic problems, as opposed to approximate solutions to real problems. In addition, as noted by Moore (1997), we have downplayed many of the fundamental but nonmathematical aspects of statistical science, such as data quality versus data quantity, the difference between sampling from a dynamic process versus a static population (analytic versus enumerative studies), or the value of reducing versus just modeling variation.

In a similar vein, George Box has often noted (see Box 1990) that "mathematics makes a good servant, but a poor master." Our concern is that IT will become the "master" rather than the "servant" of statistics in the next century. Although we certainly agree that statistical computing is critically important (as is mathematics), the challenge we see is effectively using the IT revolution to our advantage, without letting it dominate or overshadow the fundamental concepts of statistical science, which are neither mathematics nor computer science.

KEY TECHNICAL CHALLENGES

Building Quality Into Product Design

A major technical challenge for today's industrial statistician is that of helping ensure high quality and reliability in new products and in the redesign of existing products. Traditionally, quality and reliability improvement have been associated with statistical quality control and manufacturing operations. Thus, statisticians have focused on ensuring consistent manufacture of products. The quality of a product, however, is often principally determined once its design has been frozen and it has been transitioned to manufacturing. Thus, the biggest opportunity in driving toward high quality is at the product-design stage. This is an important element of the message that Taguchi, himself an engineer, tried to convey when he talked about "off-line quality control." We also feel that this is part of what Deming had in mind when he claimed that the greatest proportion of variability is "within the system" itself. The specific challenge is to effectively integrate a statistically disciplined approach and tools, such as modeling and accelerated testing, into the design process. This may also require the development of new tools, specifically targeted for design improvement. Similar comments apply for the design of systems for financial or health-care services or for environmental protection.

Making Reliability More Proactive

Reliability has been defined as "quality over time." A more formal definition is the "probability that a product performs its required function for a specified time under conditions encountered in the field." "Unreliability" causes customers to experience quality defects in a major way, and it is a significant part of the long-term cost of poor quality. Concern with reliability has become more acute as a consequence of the focus on rapid new-product introduction, the refusal of customers to accept a "ship and fix" mentality, and global competitive pressures. Assurance of high reliability presents a special challenge because, unlike scrap and rework, the costs of unreliability are often delayed. Statisticians have, in the past, been involved in estimating average life and in reliability demonstration programs. A typical question that we have been asked is: "How much testing is needed to demonstrate statistically that a product is achieving a specified level of reliability?" This question is frequently premature, however—it is hard to demonstrate something that does not exist! In fact, it is obvious that the ultimate objective should be to improve reliability, not just to estimate it! Often, the first step to improvement of reliability needs to be identification of reliability problems as speedily as possible so that the root causes can be determined and appropriate corrective action taken. This might involve both statistical approaches, such as accelerated testing, and nonstatistical ones, such as Failure Mode and Effects Analysis. Thus, a key challenge here is changing the focus from reliability estimation to reliability improvement. Reliability has also become a major issue in software systems, as well as in health-care diagnostic equipment, and therefore is important beyond the traditional "smokestack" industries.

Returning SPC to its Proactive Roots

As noted by a reviewer, Shewhart's original intent was for SPC to be a proactive tool that would enable those working with a process to actively improve it. The unique role of the control chart was to help decide the improvement approach—identifying and fixing root causes of variation if "special causes" existed, or studying the whole system to make fundamental improvements if only "common causes" existed. This intent appears to have lasted as long as engineers and quality professionals were the primary proponents of SPC. Mathematically oriented statisticians have tended to view SPC as a simplistic time series modeling approach, an equally simplistic on-line control algorithm, or a very inefficient hypothesis test. This tends to obscure the fact that the real intent of SPC is to guide human intervention to improve the process.

A key technical challenge is to "rediscover" SPC as a proactive tool, while at the same time taking advantage of the IT advances that have occurred since Shewhart's time. One example is the integration of automatic control theory with SPC (see Box and Kramer 1992; Vander Wiel, Tucker, Faltin, and Doganaksoy 1992). This was not feasible on a broad scale in Shewhart's time. Another issue is how to better adapt SPC procedures to provide a starting point for identifying the root causes of out-of-control situations. Again, with financial and health-care organizations being more data based, use of SPC has expanded well beyond the factory floor. It is equally important that financial analysts, as well as engineers, understand how to use SPC in a proactive manner.

OUR UNIQUE CONTRIBUTION—A SUMMARY

So what, in our opinion, is the industrial statistician's significant added value in the new environment? It is to ensure that in addressing key company problems and in making decisions under uncertainty one obtains the best possible information, retains such information in a readily accessible database, and analyzes it in a maximally meaningful manner. This is hardly new. But what is new is that (a) instead of undertaking these tasks directly, we provide leadership and guidance to others who have the responsibility for their implementation, (b) we look at the broader implications of problems, taking a holistic corporate view rather than a narrow technical approach, and (c) we strive to help the company avoid problems proactively rather than responding to them reactively—by, for example, building quality and reliability into the design of products.

SPECIFIC QUESTIONS

To bring the previous comments into focus, we now raise some specific questions for the panel.

1. What roles should statisticians in business and industry consider within their own organizations today so that they can be valued contributors in the world of "statistics without statisticians"? What generic "career paths" for statisticians in business and industry are realistic?

2. How can the skills required by today's environment in business and industry be most effectively developed in statisticians?

3. How can we "ride the crest" of IT revolution, including helping ensure more meaningful databases, without letting it dominate our profession the way mathematics has in the past?

4. How should our profession approach dismantling technical boundaries separating us from other professionals?

5. How can we have the same impact on the design process that we have had in manufacturing?

6. How can we make the fields of reliability and SPC more proactive and hence more valuable in management's eyes?

7. How can we take everything we have learned in the traditional application areas and apply this effectively in the growth areas such as finance?

ACKNOWLEDGMENTS

We express our thanks to Will Alexander, Frank Forbes, Dan Thorpe, Mark VanDeven, Bill Wunderlin, the editor, and two referees for their very helpful comments that significantly contributed to the improvement of this article.

[Received January 1998. Revised January 1998.]

REFERENCES

Box, G. E. P. (1990), Commentary on "Communications Between Statisticians and Engineers/Physical Scientists," by A. B. Hoadley and J. R. Kettenring, *Technometrics*, 32, 251–252.

Box, G. E. P., and Kramer, T. T. (1992), "Statistical Process Monitoring and Feedback Adjustment—A Discussion," *Technometrics*, 37, 64–73.

Deming, W. E. (1986), *Out of the Crisis*, Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Hoerl, R. W., Hooper, J. H., Jacobs, P. J., and Lucas, J. M. (1993), "Skills for Industrial Statisticians to Survive and Prosper in the Emerging Quality Environment," *The American Statistician*, 47, 280–291.

Moore, D. S. (1997), "New Pedagogy and New Content: The Case of Statistics," *International Statistical Review*, 65, 2, 123–165.

Vander Wiel, S. A., Tucker, W. T., Faltin, F., and Doganaksoy, N. (1992), "Algorithmic Statistical Process Control: Concepts and an Application," *Technometrics*, 34, 286–297.

Vere-Jones, D. (1995), "The Coming of Age of Statistical Education," *International Statistical Review*, 63, 1, 3–23.