Comment

Diane E. Duffy

1. INTRODUCTION

Professor Banks has successfully produced a provocative review of industrial statistics in the U.S. In much of this article, however, I found myself thinking, "This is not industrial statistics as I know and practice it." I have never used a control chart, never calculated a process capability index, never applied Taguchi methods. Industrial statistics is an umbrella term covering statistical activity in fields as diverse as manufacturing, communications and marketing. The common ground that these diverse areas share is the negative distinction of being non-academic. Banks writes, "Most of this paper relates these issues to the manufacturing sector, but a large part is pertinent to service industries; very little aims at statisticians working in such areas as high finance." I found rather less than "a large part" of this paper to be directly relevant to my experiences in the telecommunications industry. The diversity of statistical practice across fields and industries is such that I am doubtful of the existence of a single common body of critical issues. This is not to say, however, that I found the points in this review uniformly orthogonal to my perceptions and concerns.

Banks writes as "an academic statistician who has enjoyed broad consulting experiences and a relatively large number of interactions with various companies." I write as an industrial statistician who has enjoyed a sometimes fascinating, sometimes frustrating, often rewarding and always challenging career in the telecommunications industry. My views are rooted in my experiences, but I make no claim as to their wider relevance for industrial statisticians either within or beyond the field of telecommunications. Rather than emphasize points of agreement with Banks, in the ensuing paragraphs I deliberately focus on issues that look much different from my corner of the world of industrial statistics. In addition, I address a couple of critical topics omitted from his article.

2. THE ROLE(S) OF INDUSTRIAL STATISTICS

Banks touches upon the role of industrial statisticians in several places. For example, in Section 4 he writes, "[C]ompanies rarely need academic researchers; rather they must have access to Ph.D. level statisticians internally, who may not be developing new theory, but can comfortably command the old." Banks' Section 5 is relevant as well, as it describes his view of the skills and knowledge industrial statisticians should have in order to be effective.

Glaringly missing throughout this article is mention of, much less serious discussion of, collaboration and interdisciplinary research. Collaboration is not the same as consulting. Consulting involves a consultant and a consultor, if you will—one who gives expert advice and one who asks for it. The distinction between these two roles implies a fundamental asymmetry in the consulting relationship. Collaboration involves a team of individuals with a common goal or objective, shared ownership of problems and issues, joint responsibility for resolutions and a shared stake in the final results or outcomes. Not all of the practice of industrial statistics is collaborative, and there are certainly interactions of a consulting nature. However, the real meat and excitement are in collaboration. Essentially any important industrial problem is multidisciplinary and any successful attack requires an integrated approach. The opportunity to work with others on a variety of problems across a range of different fields and subfields is one of (if not the) major attractions of industrial statistics.

As collaborators, statisticians may make careful or sophisticated application of existing methodology to messy real-world problems. They may be required to extend or revise available statistical approaches, either slightly or significantly. They may develop new theory and methods spurred by real problems which escape tidy textbook classification. Certainly the abilities to ask the right questions, to quickly glean the critical domain information and to aid in defining the problem, are as important as a command of the key statistical issues. To be positioned to be effective collaborators, industrial statisticians need to stay up-to-date in their profession, and industrial statistics departments need to develop new approaches where needed.

What part of this is research? The short academic answer, rooted in the all-powerful metric of publishability, goes along the following lines: developing new methodology is research; extending existing methodology is research when the extensions are significantly complex; applying existing methodology may be research if the application area is of sufficient current interest. Rather than give a comparable industrial answer, I would argue that in industrial statistics, and certainly in management of industrial statistics, the critical

Diane E. Duffy is Director of the Statistics and Data Analysis Research Group, Bellcore, 445 South Street, MRE-2Q-342, Morristown, New Jersey 07962-1910.

question is not "Is it research?" but "What is its value?" The value of work in industrial statistics depends not only on its technical quality, but also on the strategic importance of the problem and the business impact of the solution.

Finally, it should be noted that there is an educational by-product to successful collaborations in industrial statistics. In effective collaborations, each contributor broadens his or her knowledge of the areas of expertise of other contributors. Ongoing effective collaborations involving statisticians serve to increase the level of quantitative literacy in the company in a natural and informal way. I consider this an important part of my job. Banks writes in subsection 3.1 that "industrial statisticians would probably do their companies better service by teaching large numbers of low-level employees how to use simple statistics, rather than by spending the same amount of time developing a trend resistant run order for a partially balanced incomplete block design with four associate classes." I do not necessarily disagree with Banks' implied assessment of the importance of the given, highly specialized and arcane topic in experimental design, but I find his view of co-workers completely at odds with my own experiences.

3. RESEARCH TOPICS IN INDUSTRIAL STATISTICS

Banks reviews four conventional topics of research in industrial statistics. His summary is: "The methodologies discussed in Section 3 are either intellectually stale, or quickly becoming so. One does not expect that new research in these areas will make astonishing discoveries that can provide enormous benefit to industry." Banks' dismissal of these fields as intellectually stale not only smacks of academic bias, but also comes dangerously close to making a major error in assessing what benefits industry. New research (in these or other areas) does not have to make astonishing discoveries in order to provide enormous benefit to industry. New research may involve a creative extension, a complex application or an innovative combination of ideas—if this leads to significant revenue increases or cost savings, it is, by definition, of enormous benefit. As a secondary issue, I find the tone of this statement and of Banks' later comment that "there are dozens of statistical problems that have intrinsic mathematical interest whose solution would provide enormous profit potential to industry" more than a bit amusing. Perhaps it is in some ways a matter of personal style, but I view most advances as incremental and evolutionary in nature—revolutionary progress and truly astonishing discoveries are quite rare. In addition, I wonder if Banks believes that a statistical problem must have intrinsic mathematical interest in order to qualify as research. Lastly, high-impact industrial problems are fundamentally multidisciplinary—statistics alone is not going to provide enormous profit potential.

In the two subsections below I discuss two active research areas in industrial statistics in more detail. I take the liberty of illustrating some of the points with examples of work of which I am aware. My purpose is simply to make the points more concrete, and I apologize for any Bellcore provincialism that these examples may suggest.

3.1 Software Reliability

Software reliability is discussed as one of the newer topics in industrial research. The following comments primarily serve to broaden Banks' discussion of this important topic.

Banks writes that software is "as unique as artwork, since each program is written for a specific task, . . . there is only one chance to get it right." Individual programs do have specific tasks, but (many of) the modules comprising large complex programs are far from unique in their basic functionality. Fast-growing interest in and application of object-oriented programming languages and platforms have been fueled in no small part by the desire to increase code reuse (and by the strong belief that much code can be reused). Thus far it is primarily small applications (no more than 30K lines of code) that have been developed entirely in object-oriented environments. The extent to which the object-oriented paradigm will influence and improve large systems (millions of lines of code) remains to be seen. There are many other efforts in the computer science field aimed at code reuse. As one example, colleagues at Bellcore have been experimenting with the use of customized information retrieval technology to aid software developers in locating potentially useful modules from previously written systems.

There has been substantial additional work by Dalal, Mallows and colleagues on the problem of deciding when to stop testing a piece of software, and on related problems. The most direct follow-up to Dalal and Mallows (1988) is work by Dalal and McIntosh (1992) on extending and applying their theory to telecommunications software development projects. Notable among the needed extensions to the methodology is the ability to accommodate code churn (which may be as frequent as daily and which makes the software under test a moving target). Dalal and McIntosh also discuss metrics for assessing testing effort and they develop graphics to guide testing managers in their decision making. The approach described in Dalal and McIntosh (1992) has been packaged as a CASE (computer-aided software engineering) tool and used to aid decision making in several large software systems (i.e., systems with seven to eight million lines of code in total for which new releases include 40K to 400K lines of new code).

Recently a Bayesian variant of this methodology for deciding when to stop testing was used to help plan the test cycle for a large system release before testing had begun (i.e., when the code was still under development). In separate work, Dalal and Mallows consider the related problem of estimating the quality of externally provided software (Dalal and Mallows, 1992).

I am aware of several other efforts aimed at judicious incorporation of statistical thinking in the software development process, and I am sure there are many more such efforts of which I am unaware. While these may not offer "the sweeping solution" that Banks, at least, hopes to find, they hold promise for meeting the more realistic objective of steady progress in this field. For example, there is a great deal of current interest in the use of experimental design methodology (e.g., orthogonal arrays and fractional factorial designs) and sampling techniques to reduce the number of test cases used on a piece of software without sacrificing the quality and coverage of the testing. To see widespread use, these techniques for reducing testing effort must be made available as automated tools in software testing labs. Dalal and Patton at Bellcore (and no doubt others at other research labs) are currently experimenting with these approaches. The initial results look promising and larger-scale trials are needed to make general recommendations. These ideas are also likely to be useful for testing finite state machines. (As with any relatively new domain of application for a maturing field, there is the potential for rejuvenation as new problems are motivated by the particular constraints of the application domain. Banks' sense that "the bloom is off the experimental design rose" may be premature.)

Dalal, Horgan and Kettenring (1993) provide an up-to-date review of statistical techniques of potential use in various phases of the software development process for the purpose of improving software quality and reliability. Among the topics they discuss are approaches that can be aimed further upstream in the software development cycle, where potential savings can be greater. One specific example of such an approach is recent work by Eick and coworkers (1992) on using capture–recapture methods to estimate the number of faults outstanding in a software requirements document. This estimate can be used to decide if the requirements are well enough formulated to proceed with software development or if further review and revision of the requirements to reduce the number of outstanding faults is warranted.

3.2 Statistics and Computing

Banks is careful to state that his examples of new research topics in industrial statistics are not meant to be either exhaustive or representative. However, the omission, here and throughout the article, of discussion of the myriad ways in which computing and statistics are interacting in industrial research strikes me as egregious. The last 20 or 25 years has seen explosive growth in both the availability and the capability of computing. Just as it is the truly rare statistician who makes no use of computing in his or her work, it is the truly rare industry in which computers have not had significant impact on the workforce.

The following paragraphs attempt to highlight some areas of current industrial research related to statistics and computing. Like Banks, I make no attempt at being complete; however, I have tried to somewhat fairly represent the scope of statistical computing research topics. The purist may question whether some of these topics are "really statistics." I confess to having little concern for such distinctions. Particularly in an industrial context, quibbling over discipline definitions and boundaries does little to advance progress in solving real problems.

The trend toward ever larger and more complex data sets has spurred interest in several aspects of statistics and computing. Data management strategies including databases tailored for statistical applications is one such area; high-performance computing and communications, particularly high-speed networking and large bandwidth to disk, is another area. For example, ongoing studies of telephone signaling network traffic yield a data set of approximately 100 million observations. Customized software permits flexible interactive analyses of these data despite the size and complexity. Other studies have focused on traffic on Ethernet local area networks (LANs), which are three orders of magnitude faster than telephone signaling links (Leland et al., 1993). These LAN studies produce data sets in the gigabyte range and, to date, only a (carefully selected) fraction of their data has been explored. Among the other contexts in which very large data sets can arise are medicine (imaging data), aerospace (satellite data), marketing (panel data) and meteorology (environmental data).

The explosion of computing presence and power has had the effect of greatly increasing the number and size of databases which corporations maintain, sell and utilize. This in turn has motivated some interesting problems in the interface between database research and statistics. For example, increasingly software applications need to access disparate databases. This requires linking or merging information sources, none of which can usually be taken to be error free. Bayesian decision analysis seems, on the face of it, to be of potential use here. As a second example, data mining is a term often used by researchers in the artificial intelligence community to describe the process of exploring large databases with the goal of extracting patterns. Data mining has clear connections to statistics; however, when very large databases are involved, computational issues and algorithms may be at least as important.

Statistical graphics and data visualization is a very active area of research. Industrial statisticians have played a notable role, motivated by the fact that graphical techniques can be very effective both at quickly highlighting key features in data and at communicating these to subject matter experts and to managers. Contributions from industry have ranged from groundbreaking research in the foundations of graphical perception and in plotting paradigms for exploratory data analysis, to the development of state-of-the-art systems for high-interaction dynamic graphics [e.g., Bellcore's XGobi software, developed by Swayne, Cook and Buja (1991)]. (Banks, in Section 5, states that graphics are useful for presentation to upper management. This narrow view considerably underestimates the value of statistical graphics in industrial environments, and the quality and quantity of industrial research in this field gives a clear signal of the broader usefulness of graphics.)

Modern statistical methodologies invariably require computing. To incorporate these methodologies into application domains successfully, it is critical that they be packaged in software that is easy to use. Well-designed graphical user interfaces (GUIs) are much easier to learn and to use than classical text-based or command line interfaces. Designing good GUIs for statistical software is very much an art form, and there are many open research questions in this field.

Lastly, the availability of powerful computing has made simulation an increasingly important technique across a wide range of problems; locally these range from studies of network performance aimed at improving network reliability and integrity, to models of telephone workflows aimed at reengineering work processes for greater efficiency. Statistics has much to contribute to simulation studies. In the design phase, probability models are frequently needed, and experimental design techniques may be used to choose parameters for different runs. (Note Banks' reference in subsection 3.1 to research in Latin hypercubes motivated by computer models.) In the analysis phase, statistics are used for investigating, summarizing and presenting results. When auxiliary data are available, statistics plays a significant role in validating and/or improving the simulation model.

4. THE EDUCATION OF INDUSTRIAL STATISTICIANS

As anticipated at least in part by Banks, the topic of statistical education spawns seemingly endless discussion. I confine myself to several brief comments.

While I agree with Banks that computational skills are essential, I am uneasy with the blanket requirement of serene competence in at least one major package. If a student is headed for a certain field of applied statistics, and if there is a standard package in that field, then competence in that package is a major asset. However, the heterogeneous world of hardware platforms, computing environments, programming languages and software packages demands broader computer literacy. A student may be better off familiarizing him- or herself with a variety of packages running on different platforms, at the expense of being a master of any one package. This is especially true given the need to adapt to the constant change in all phases of computing.

In Banks' opinion, "survey sampling is not ubiquitously used..." in industrial statistics. My experience is much the opposite—lurking under many interesting problems are critical sampling issues concerning the representativeness of the data and the potential for sampling biases. A solid understanding of the fundamental ideas of sampling theory (e.g., stratification, clustering, probability sampling and weighting, probability sampling proportional to size) is important for any applied work; extensive experience deriving and applying survey sampling formulae for variance estimators in complex designs is not.

Banks' view of the very low demand for discrete multivariate analysis (MVA) methods may be appropriate for some industrial domains, but it is certainly not universally appropriate. In fields such as biostatistics and marketing, familiarity with contingency tables and related techniques is absolutely essential. (Interestingly enough, the quality movement's focus on customer satisfaction has led to increasing emphasis on measuring customer perceptions of quality and service. Data on customer satisfaction is frequently discrete multivariate data collected by sample surveys!)

5. MISCELLANEOUS

Banks raises concerns about the ritualized use of TQM tools. Applied statistics is rife with examples of the blind application of methodology without thoughtful assessment. I have wondered whether improvements in interface design for statistical software could simultaneously promote sensible and thoughtful analysis and meet users' demands for ease-of-use and efficiency.

Banks takes amusement in the trendiness of American business as illustrated by the race to hop on the TQM bandwagon. I am not sure that statistics is so different; Banks' use of words such as "hot" and "fashionable" to describe topics in statistics illustrates a similar propensity to be ever enamored by that which appears to be new.

In Section 6 Banks deplores "[t]he occasional jingoistic tone of corporate rhetoric . . . ." Occasional jingoism is, however, just that—occasional. What is not occasional in industrial environments, and hence of greater significance in the relationship between industrial and academic statisticians, is economics—costs, revenues, profits, patents, intellectual property rights. Proprietary restrictions on data and results contrast sharply with many academic statisticians' experiences and, coupled with a bottom line emphasis, are sometimes viewed as unscientific.

Banks writes in Section 5: "[I]t is my impression that students at all levels should anticipate a credential deflation upon graduation." The ensuing paragraph appears to accuse the typical non-academic employer of underestimating the expertise of a newly employed statistician, while exempting academic employers of such underestimation. I have a fairly good memory of my start at Bellcore, and credential deflation was not an issue!

ACKNOWLEDGMENTS

I thank Andreas Buja, Sid Dalal, Jon Kettenring and Walter Willinger for reviewing draft versions of this commentary. I am particularly grateful for Sid's comments on software reliability and Andreas' comments on computing.

Comment

Gerald J. Hahn

Wow! This article is a commentator's delight! It packs in more provocative ideas—and probably succeeds in offending more readers (from CEOs to Californians)—than any paper I have ever read in a technical journal; moreover, it does so in an IMS journal—not exactly a hotbed of radicalism! Its broad title "Is Industrial Statistics Out of Control?" is, in fact, too modest, and somewhat rhetorical. The author takes on academic statistics as well, and throws in five book reviews to boot! The article makes enticing reading—like a good novel, I found it hard to put down. And, amazingly, I found myself agreeing with many of the author's points! I am more competent to discuss Banks' remarks on industry than on academia—but that does not stop me from commenting on both.

FIRST SOME PLAUDITS

Add my "hear, hear" to:

- "Intelligent use of simple tools will achieve about 95% of the knowledge that could be obtained through more sophisticated techniques."
- "The real danger in the ubiquitous Taguchi cult is that industry users may copy the inefficiencies as well as the insights." I also welcome the increased appreciation of planned experiments that has accompanied the Taguchi fervor, but, like other statisticians, am dismayed at the combative attitude toward past work of some of Taguchi's more avid supporters.
- "Companies ... must have access to Ph.D. level statisticians, who may not be developing new theory, but can comfortably command the old." What is frequently needed are extensions of current methods tailored to the problem at hand.
- "Contrary to most students' beliefs, defining the problem can be the greatest service the statistical consultant provides; this task is often fluid, difficult and without unique solution. Unless students see this process for themselves, they usually think in terms of end-of-the-chapter exercises and struggle to find an exact match between a recently taught tool and the client's naive statement of the problem."
- "New industrial statisticians should be prepared to quickly build a very deep understanding of the industry that employs them." A solid scientific background is extremely helpful in this regard.
- "Superficial exposure to a great range of methods (cluster analysis, ridge regression, CART, the bootstrap, etc.) may be better education than the current drill in the mathematical consequences of sufficiency, Fisher information and the strong law of large numbers." Modern statistics is much broader than what can be embedded in one, or even many, Ph.D. programs. While academics can, to some degree, choose their own playing field, industrial statisticians cannot. Thus, the most industry can expect from an academic program is that it builds fundamental understanding, exposes students to a wide variety of important areas—and then helps them develop the skills to rapidly learn more as needed.
- "Some academic statisticians regard most indus-

Gerald J. Hahn is Manager of the Management Science and Statistics Program at the GE Research and Development Center, Schenectady, New York 12301.