Quality Engineering, 27:111–121, 2015

ISSN: 0898-2112 print / 1532-4222 online
DOI: 10.1080/08982112.2015.968054

Taylor & Francis
Taylor & Francis Group

Statistics: A Life Cycle View

Ron S. Kenett
KPA Group, Ra'anana, Israel

ABSTRACT Statistics has gained a reputation as being focused only on mathematical modeling, data collection, and data analysis. This article is about an expanded view of the role of statistics in research, business, industry, and service organizations. Such an approach provides an antidote to the narrow view of statistics outlined above in an effort to reposition it as the leading profession in the analytics domain. This view creates a need for focused research activities and the development of new methods and tools, often in collaboration of statisticians with experts in other disciplines (Kenett 2012). Specifically we discus here a "life cycle view" consisting of (1) problem elicitation, (2) goal formulation, (3) data collection, (4) data analysis, (5) formulation of findings, (6) operationalization of findings, (7) communication, and (8) impact assessment. These eight phases, with internal iterations, combine the inductive–deductive learning process studied by George Box (1997). Covering these phases, beyond the data analysis phase, increases the impact of statistical analysis and enhances the level of generated knowledge and information quality. The envisaged overall approach presented here is that the research and practice of applied statistics needs to involve a trilogy combining (1) a life cycle view, (2) an analysis of impact, and (3) an assessment of the quality of the generated information and knowledge. We begin with a section introducing the problem, continue with a review of the information quality (InfoQ) concept presented in Kenett and Shmueli (2014), and proceed with a description of the eight life cycle phases listed above. Adopting a life cycle view of statistics has obvious implications to research, education, and statistical practice. These are presented in the context of several examples. We conclude with a discussion of such implications.

KEYWORDS applied statistics, life cycle view, impact analysis, information quality (InfoQ), practical statistical efficiency (PSE), statistical engineering, trilogy of applied statistics

INTRODUCTION

Applied statistics is about meeting the challenge of solving real-world problems with mathematical tools and statistical thinking. However, as eloquently stated by David Cox, "Much fine work in statistics involves minimal mathematics; some bad work in statistics gets by because of its apparent mathematical content" (Cox 1981, p. 295). This article is about *fine work in statistics* and an expanded view of the role of statistics in research, business, industry, and service

Address correspondence to Ron S. Kenett, KPA Group, P.O. Box 2525, Ra'anana 43100, Israel. E-mail: ron@kpa-group.com

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lqen.

organizations. The life cycle view expands the work presented in Kenett and Thyregod (2006) and aims at improving the position of statistics as a scientific discipline with wide relevance to organizations and research activities. The life cycle envisaged here is presented in Figure 1. It includes internal iterations that combine the inductive–deductive learning process studied in Box (1997). The main point is that considering a life cycle, with iterations, brings attention to usually neglected but crucial components in the application of statistics. Moreover, the proper application of statistics, focused on generating knowledge and information, requires such an approach.

Such a life cycle view is even more critical in the context of big data analytics and modern technology used in running basic and sophisticated data analysis method.

This view is part of a trilogy that includes the following:

- **Measuring impact of statistical work**. For the concept of practical statistical efficiency (PSE), see Kenett et al. (2003).
- **Assessing the quality of information generated from a given data set**. For the concept of information quality (InfoQ), see Kenett and Shmueli (2014).
- **A life cycle view of statistics**; see Kenett and Thyregod (2006).

An endorsement of this view is provided by the INFORMS Certified Analytics Professional (CAP) certification program (https://www.informs.org/Certification-Continuing-Ed/Analytics-Certification). The domains of CAP certification are (1) business problem (question) framing, (2) analytics problem framing, (3) data, (4) methodology (approach) selection, (5) model building, (6) deployment, and (7) model life cycle management. These domains were derived from an extensive job task analysis study of the current knowledge, skills, and abilities that must be demonstrated by effective and successful analytics professionals. The trilogy outlined above covers all of the CAP domains, and more. The next section expands on specific phases in a life cycle view of statistics.

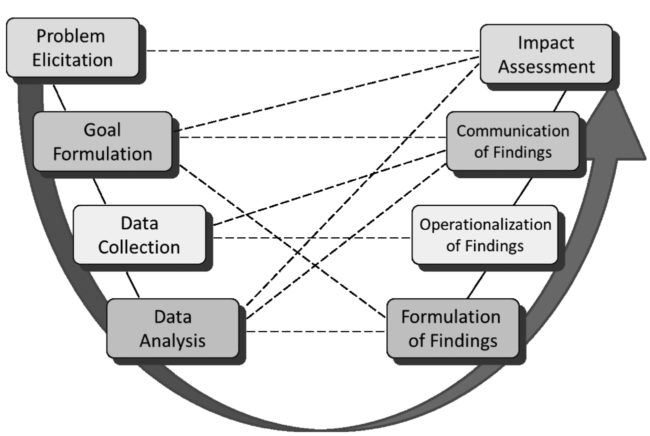

FIGURE 1 The life cycle view of statistics.

THE LIFE CYCLE VIEW OF STATISTICS

The life cycle view we elaborate on here includes eight phases, namely: (1) problem elicitation, (2) goal formulation, (3) data collection, (4) data analysis, (5) formulation of findings, (6) operationalization of findings, (7) communication, and (8) impact assessment. These eight phases are introduced below, with examples.

Problem Elicitation

Greenfield (1987) presented humoristic examples of encounters illustrating the widespread opinion of many clients that a consultation with a statistician is not very different from a consultation with a dentist: you give him a hint about your symptoms, you are placed in the chair, she looks into your mouth, and she diagnoses and solves the problem, all in less than an hour. The reality is, of course, more complex and requires skills in active listening. The experience statistician knows how to listen and ask probing questions in order to check his own understanding and bring out facts and assumptions. At the problem elicitation phase, the statistician should keep his customer focused and yet challenged in order to bring forward relevant details. Process maps and site visits are often useful tools for structuring the investigation. Walking around in the process area under investigation typically provides insights on possible sources of variation. A famous quote by Yogi Berra states: "You can observe a lot by watching." One should also read the client's body language because he might have a hidden agenda and keep some relevant information undisclosed. His goal might be to have the statistician cover up for some internal blunders or settle an internal dispute in the company in the role of impartial umpire. Handling unstructured problems and problem elicitation in general are critical applied statistics activities that are not supported by tools and methods and mostly addressed informally without a system. Most statistics textbooks do not even discuss this phase.

Goal Formulation

Sometimes, the problem being addressed is not clear. Goal formulation can be facilitated by the intervention of an experienced statistician. Eventually, this may even be judged to be his major contribution (Daniel 1969). Bill Hunter described two chemists entering his office who, when asked to describe the problem they came to get help for, entered into a lengthy discussion that eventually led to a common reformulation of the problem. Once they had the problem clearly formulated, they knew how to solve it and did not need the help of a statistical consultant any more. They left his office thanking him abundantly for giving them the opportunity to clarify the problem. Bill only asked them to explain their problem and sat the whole session, saying nothing (W. G. Hunter, Department of Statistics, University of Wisconsin, Madison, personal communication, October 20, 1979).

Data Collection

Cobb and Moore (1997, p. 810) argued that "statistics requires a different kind of thinking because data are not just numbers, they are numbers with a context." Sometimes the term *metadata* is used to denote the context domain. However, very few textbooks address the mapping between the context domain of a problem and the analytic domain where abstract statistical model are used. This mapping is discussed in Kenett and Shmueli (2014) as an initial step in identifying goals that enable the assessment of the quality of the information generated by a specific statistical analysis. The importance of *context* is illustrated by an exercise in a textbook for fourth-grade pupils in a Danish school. For each day in July an ice cream vendor recorded the number of ice creams sold and the data were shown on a simple run chart without indication of the actual dates on the chart. In July it was very hot for a period of nine consecutive days and the students were asked to identify these days and also to determine which days were Sundays. By itself, the graph is just a representation of 31 numbers. However, given the context, and the associated implicit information (weekly calendar cycle with parents more inclined to offer ice cream treats in weekends and on sunny days) it is possible, even for these young children, to separate signal from noise in this set of numbers. The pupils are able to solve this elementary statistical exercise because they have the ability to distinguish between the effects of different patterns of variation (Kenett and Thyregod 2006). In general, the context domain of data is revealed where data are generated, on the shop floor, in the laboratory, etc. Therefore, in order for a statistician to be able to respond properly to the needs of his customer, it is important for him to take part in the data collection or at least have the opportunity to watch data being collected or generated. The statistician should not just rely on his customer telling the whole story. To such customers, several features of the data are just as obvious as the repetitive pattern of the week in the example above and therefore they are not mentioned. Usually, the customer is considered the subject matter expert with specific knowledge and experience. The statistician, on the other hand, is an expert in treating variation in measurements. When visiting the customer premises, the statistician is alert to possible sources of variation. He literally, or mentally, sketches an Ishikawa cause-and-effect diagram listing these sources and indicates their nature and their dynamics such that they can be adequately represented in the data analysis (Kenett 2007). One should carefully consider what sources can be controlled, or not, and whether one wants to include the potential effect of this variation, or not. In addition, one should make sure that traceability between physical artefacts and data is secured. Individual parts should be traceable to the work order, measurements should be traceable to the measuring device, and the calibration history of measuring devices should be recorded and retrievable.over time.

The statistician should not confuse high-frequency and low-frequency variation. In control chart applications, for example, it is sometimes seen that a subgroup of four items is used to represent an hour's production but that all four are taken consecutively at the same time instant. In such cases, the within-subgroup variation only represents the (small) instantaneous high-frequency variation. The "natural" variation of the process might better be represented by the low-frequency variation, with samples taken at equal intervals within, say, an hour. Often the customer has data relating to the issue that gives valuable insight in the order of magnitude of variation due to various sources.

Today's big data has been characterized by volume, velocity, and variety (Russom 2011). These three Vs have often been an impediment to data analysis. Industrial organizations established data warehouses were meta-tagged data were stored and retrieved in cubes. Machine-generated data or unstructured data from text or voice recordings in service centers were often not part of the warehouse, mostly because of the three Vs. For more on integrating structured and unstructured data see Kenett and Raanan (2011).

Data Analysis

This section is about data analysis with a practical perspective. It builds on two case studies originally presented in Kenett and Thyregod (1996) and Kenett and Shmueli (2014).

The Automobile Part Example

An automobile part is produced by injection molding followed by baking and other finishing process steps. The molding is performed in three parallel streams. Suddenly the reject rate of the part due to cracks increased dramatically and, as a result of this increase, it was decided to investigate the process. The engineers from the plant had just completed a course in design of experiments and were eager to practice their newly acquired skills. Therefore, they designed a rather comprehensive 32-run experiment with 11 factors focusing on injection molding machine factors like speed, clamp time, etc. Very little was learned from the experiment that could reduce the scrap rate and therefore a second experiment was planned to focus on differences in molds. This experiment clearly demonstrated that most of the scrap came from one mold, and in the search for the root cause it was discovered that during routine maintenance some minor changes had been made to one of the molds and that these changes were the source of the problem. In hindsight, the team realized that because each part carried a mark indicating which mold produced the part, the problem could have been solved without experimentation. Because the plant was having high rates of scrap, it would have been rather easy to use the existing data and utilize the traceability of the parts to the mold. A histogram of scrap by mold would have revealed that scrap was not evenly distributed over the molds and clearly identify the culprit. Although the production in parallel streams strongly suggested that stratification would be a meaningful analytical tool, the power of this simple analysis on existing data was overlooked in the eagerness of the team and the statistician working with them to perform an experiment. Alternatively, running control charts on number of cracks and dimensions would have shown the same information and provide an online statistical process control (SPC) mechanism with early warnings of shifts and effective diagnostics. For a web-based cloud-hosted SPC system, see http://www.spclive365.com. for more on modern industrial statistics methods see Kenett and Zacks (2014).

The eBay Internet Auction Example

Katkar and Reiley (2006) investigated the effect of two types of reserve prices on the final auction price on eBay. They specifically studied the effect of two types of reserve prices on the final auction price. A reserve price is a value set by the seller at the start of the auction. If the final price does not exceed the reserve price, the auction does not transact. On eBay, sellers can choose to place a public reserve price that is visible to bidders or an invisible secret reserve price (bidders only see that there is a reserve price but do not know its value). Their data came from an experiment selling 25 identical pairs of Pokémon cards, where each card was auctioned twice, once with a public reserve price and once with a secret reserve price. The data consist of complete information on all 50 auctions. The authors used linear regression to test for the effect of private/public reserve on the final price and to quantify it. The utility was statistical significance to evaluate the effect of private/public reserve price and the regression coefficient for quantifying the magnitude of the effect. The authors concluded, "A secret-reserve auction will generate a price $0.63 lower, on average, than will a public-reserve auction" (Katkar and Reiley 2006, p. 17).

Formulation of Findings

Statistical analysis produces outputs with various features such as *p*-values, standard errors, regression models, principal components, discriminant score functions, analysis of variance tables, control charts, descriptive statistics, etc. In order to translate statistical results into relevant information, careful formulation of findings is required. In the formulation of findings, we need to translate features into benefits and eventually core values. Reynolds and Gutman (1988) provide an introduction to "laddering," a marketing methodology used to design and position products and services. The methodology relies on cognitive science where means-end chains have been studied for achieving such a mapping. The approach can also be used to formulate findings from statistical data analysis. For example, when a control chart indicates a point out of control, the process engineer needs to launch an investigation to characterize the possible out-of-control condition that this signals. The immediate benefit for the process engineer is that, by acting effectively on early warning signals, he performs his role professionally and gets well-deserved recognition. If that engineer has, as a core value, satisfaction from his professional responsibility, we have completed the means-end chain, linking feature to benefit and value. The formulation of findings is a conscious effort to translate statistical reports into the language of the customer. It involves identifying consequences of the findings of importance to the customer and requires an understanding of the context of the problem investigated. It also requires effective communications and intensive interaction with the customer. In that context, statistical consultants are many times drawn into organizational internal issues and their findings are instrumental in settling internal disputes over responsibility for problems and budget allocation.

Operationalization of Findings

To operationalize findings means to put some new insights into operation. Operational definitions are used to determine a system state in terms of a specific, publicly accessible process of preparation or validation testing, which is repeatable at will. A recipe for making a cake is an operational definition used in the household kitchen. Properties described in this manner must be sufficiently accessible, so that persons other than the definer may independently measure or test for them at will. An operational definition is generally designed to model a theoretical definition. The binary version produces either the result that the object exists or that it does not. This applies to the classification of production items as pass or fail. Goodman and Hambleton (2004) reviewed school students test score reports in the United States and showed good and bad examples of how test reports are communicated and made actionable. A prime example of operationalization of findings is provided by the National Education Goals Panel that recommended that test reports help answer four questions asked by parents: (1)How did my child do? (2) What types of skills or knowledge does his or her performance reflect? (3) How did my child perform in comparison to other students in the school, district, state, and, if available, the nation? and (4) What can I do to help my child improve? These are examples of operationalization of data-based reports. In a life cycle view of statistics, the statistician is also responsible for the operationalization of findings.

Communication

It is obvious that the statistician should effectively communicate his findings to the client. For the communication to be effective, it is important to use the language of the customer and not to distract attention from the problem. Concepts and notation from mathematical statistics will usually restrain communication and should therefore be avoided. For most people, graphical displays are more effective communication tools than mathematical formulae. Findings that cannot be presented in a graph are probably not worth communicating. One should, however, keep graphs and slides simple and avoid the temptation to overload them with unnecessary "information," like logos, fancy symbols, etc. Tufte (1997), in his seminal book about envisioning information, discusses the misrepresentation of data related to the 1986 *Challenger* disaster. He criticizes several of the data displays used by NASA and suggests a rather simple display with temperature as the x-variable and O-ring condition on the y-axis to explain what happened. The chart, based on a logistic regression, demonstrates that O-ring failure had occurred on all launches below 18°C. The temperature at launch was 3°C. The information about the problem was therefore there; it was, however, not properly communicated to the decision makers. The report of the investigation into the loss of the Space Shuttle *Columbia*, 17 years after the *Challenger* disaster, concluded that the NASA management practices constituted a major cause of the accident (Columbia Accident Investigation Board 2013). One of these practices was "engineering by viewgraphs," where technical documentation consisted of slide presentations where every single text slide used bullet outlines with four to six levels of hierarchy and then another list of bullets starts afresh for a new slide. On page 191 of the report, the Board makes the following observation: "In this context it is easy to understand how a senior manager might read this PowerPoint slide and not realize that it addresses a life-threatening situation" and, further, "The Board views the endemic use of PowerPoint briefing slides instead of technical papers as an illustration of the problematic methods of technical communication at NASA." Godin (2001) emphasized that using a cognitive style for presentation differs from the documentary style in technical reports. Communication of findings from a specific study could therefore consist of a range of publications, each one adapted to a different audience.

Impact Assessment

To be relevant, statistics needs to measure its practical impact in addressing real problems. PSE is an approach to assess the impact of some statistical work. PSE is a function combining

- V{D} = value of the data actually collected for the problem at hand. A synonym to data quality.
- V{M} = value of the statistical method employed. This is the classical notion of Pitman efficiency.
- V{P} = value of the problem to be solved, usually in economic terms.
- V{PS} = value of the problem actually solved. Usually only a portion of the whole pie.
- P{S} = probability that the problem actually gets solved. Meaning an initiative beyond the PowerPoint and the delivered report.
- P{I} = probability the solution is actually implemented. Deploying solutions often hit unexpected difficulties.
- T{I} = time the solution stays implemented.
- E{R} = expected number of replications.

Computing PSE can be done, for example, by multiplying a 1–5 rating of the eight elements listed above. An example of this approach is provided by a case study in Kenett et al. (2003) describing an old family-owned tea packing company that decided to modernize. Key performance indicators including efficiency, staff absence, and waste were adopted. Basic run charts displaying performance of packing machines were placed near the machines and operators could now observe their performance. This led to a clearer understanding of causes of downtime. The company had two identical packing lines and yet one produced with 54% efficiency and the other with 68%, with the same number of workers. The cause of the problem was eventually identified as the availability of engineers to mend the line when problems occur. A change in work priorities of engineers was implemented to remedy this situation. The PSE components for this case study have been assessed qualitatively on a 1–5 scale as follows:

- V{D} = value of the data actually collected = 4.
- V{M} = value of the statistical method employed = 3.
- V{P} = value of the problem to be solved = 4.
- V{PS} = value of the problem actually solved = 4.
- P{S} = probability level that the problem actually gets solved = 5.
- P{I} = probability level the solution is actually implemented = 5.
- T{I} = time the solution stays implemented = 3.
- E{R} = expected number of replications = 4.

The PSE for the machine packing improvement project was therefore evaluated as 57,600. Calculating PSE values provides an approach to compare projects and identify best practices in statistical consulting practices. We need, however, a theory and better methods for assessing PSE (see Kenett 2012).

A prime example of a project with huge PSE is the U.S. National Demonstration Project for Quality Improvement in Healthcare. Quoting Godfrey (2012):

> Twenty-five years ago we launched an interesting experiment, "The National Demonstration Project for Quality Improvement in Healthcare." It was a modest experiment bringing together twenty-one healthcare providers with twenty-one top industrial companies to explore whether industrial quality methods would work in healthcare settings. The results of this experiment were published as Curing Health Care: New Strategies for Quality Improvement. The statistical methods used by most of these healthcare providers were fairly basic tools of quality improvement; yet, many of the improvements were significant (Berwick et al. 2002).

The term *significant* is clearly a massive understatement because the impact was at the local hospital level, at the national level, and global. Health care is an area were economics, science, social responsibility, and ethical considerations are merged in a unique way. Statisticians should be concerned with the impact of their work and routinely evaluate PSE or any other measure of impact. The next section is about another measure of statistical analysis outcomes, the assessment of the quality of the information generated by statistical work.

AN INTRODUCTION TO INFORMATION QUALITY

Research and statistical analysis generate information. Some analysis is rich and provides answers to the research questions and goals. Some studies consist of applications of statistical tools without significant added value and thereby generating information of low quality. A comprehensive study of the elements that determine the level of information quality provided by an analytic study has been conducted in Kenett and Shmueli (2014), who proposed the concept of InfoQ described herein.

InfoQ determine the level of information quality provided by a specific statistical analysis effort. Formally, InfoQ(f, X, g) = $U(f(X|g))$, where g is a specific analysis goal, X is the available data set, f is an empirical analysis method, and U is a utility measure. To achieve high InfoQ, eight dimensions need to be considered. They determine the specific level of InfoQ and can be assessed independently for any applied work involving statistical methods. The InfoQ dimensions are (1) data resolution, (2) data structure, (3) data integration, (4) temporal relevance, (5) generalization, (6) chronology of data and goal, (7) operationalization, and (8) communication. Statistical projects that have been poorly communicated, and therefore ignored, are in abundance. Sophisticated data analysis based on data of inappropriate resolution with useless results is also common. To be of value, the planning and deployment of statistical work should ensure high InfoQ. A coarse-grained approach to assess InfoQ is to rate each dimension on a 1–5 scale, with 5 indicating high achievement in that dimension. The ratings (Y_i, i = 1, …, 8) can then be normalized into a desirability function for each dimension ($0 \le d(Y_i) \le 1$), which are then combined to produce an overall InfoQ score using the geometric mean of the individual desirabilities. Specifically:

$$\text{InfoQ Score} = [d_1(Y_1) \times d_2(Y_2) \times \ldots \times d_8(Y_8)]^{1/8}.$$

As an example, we evaluate the eight InfoQ dimensions of the eBay auction price study by Katkar and Reiley (2006) described previously. For demonstration purposes, we use a 1–5 scale and generate an InfoQ score based on a desirability function set up so that $d(1) = 0$, $d(2) = 0.25$, $d(3) = 0.5$, $d(4) = 0.75$, $d(5) = 1$.

Data Resolution

The experiment was conducted over 2 weeks in April 2000. We therefore have no data on possible seasonal effects during other periods of the year. Data resolution was in U.S. cents but individual bids were dropped and only the final price was considered. Other time series (e.g., the cumulative number of bids) were also aggregated to create end-of-auction statistics such as "total number of bids." Given the general goal of quantifying the effect of using a secret vs. public reserve price on the final price of an auction, it seems that the data were somehow restrictive. The 2-week data window allows for good control of the experiment but limits data resolution for studying a more general effect. Hence, we rate the data resolution $Y_1 = 4$ (high).

Data Structure

The data included only information on the factor levels set by the authors and the three outcomes final price, whether the auction transacted, and the number of bids received. These data were either set by the experimenters or collected from the auction website. Although time series data were potentially available for the 50 auctions (e.g., the series of bids and cumulative number of bidders), the authors aggregated them into auction totals. Textual data were available but not used. For example, bidder usernames can be used to track individual bidders who placed multiple bids. With respect to corrupted data, one auction winner unexpectedly rated the sellers, despite their request to refrain from doing do (to keep the rating constant across the experiment). Luckily, this corruption did not affect the analysis due to the study design. Another unexpected source of data corruption was eBay's policy on disallowing bids below a public reserve price. Hence, the total number of bids in auctions with a secret reserve price could not be compared to the same measure in public reserve price auctions. The authors resorted to deriving a new "total serious bids" variable, which counts the number of bids above the secret reserve price .Given the level of detailed attention to the experimental conditions but the lack of use of available time-series and textual data we rate this dimension $Y_2 = 4$ (high).

Data Integration

The authors analyzed the 2-week data in the context of an experimental design strategy. The integration with the DOE factors was clearly achieved. No textual or other semantic data seemed to have been integrated. We rate this dimension $Y_3 = 4$ (high).

Temporal Relevance

The short experiment duration and the experimental design assured that the results would not be confounded with the effect of time. The experimenters tried to avoid confounding the results with a changing seller rating and therefore actively requested winners to avoid rating the seller. Moreover, the choice of Pokémon cards was aligned with timeliness, because at the time such items were in high demand. Finally, due to the retrospective nature of the goal, there is no urgency in conducting the data analysis shortly after data collection. We rate this dimension $Y_4 = 5$ (very high).

Chronology of Data and Goal

The causal variable (secret/public reserve) and the blocking variable (week) were determined at the auction design stage and manipulated before the auction started. We rate this dimension $Y_5 = 5$ (very high).

Generalizability

The study is concerned with statistical generalizability: do effects found in the sample generalize to the larger context of online auctions? One possible bias, acknowledged by the authors, is the seller's rating of zero (indicating a new seller), which limits the generalizability of the study to more reputable sellers. In addition, the authors limited the generality of their results to low-value items, which might not generalize to more expensive items. We rate this dimension $Y_6 = 3$ (acceptable).

Operationalization

The authors consider two theories that explain the effect of a secret vs. public reserve price on the final price. One is psychological, where bidders can get "caught up in the bidding" at low bid amounts and end up bidding more than they would have had the bidding started higher. The second theory is a model of rational bidders: an auction with a low starting bid and a high secret reserve can provide more information to bidders than an auction with a high starting bid. Though these two theories rely on operationalizing constructs such as information and caught up in the bidding, the authors limited their study to eBay's measurable reserve price options and final prices. We rate this dimension $Y_7 = 3$ (acceptable).

Communication

This research study communicated the analysis via a paper published in a peer-reviewed journal. Analysis results are presented in the form of a scatterplot, a series of estimated regression models (estimated effects and standard errors), and their interpretation in the text. We assume that the authors made additional dissemination efforts (e.g., the paper is publicly available online as a working paper). The paper's abstract is written in a nontechnical and clear way and can therefore be easily understood not only by academics and researchers but also by eBay participants. The main communication weakness of the analysis is in terms of visualization, where plots would have conveyed some of the results more clearly. We therefore rate this dimension $Y_8 = 4$ (high).

As mentioned above, a coarse approach to determine the overall InfoQ score is by computing the geometric mean of the desirability values associated with the eight InfoQ dimensions scores. For this research the InfoQ score is 73%, a level considered high.

IMPLICATIONS OF THE TRILOGY OF APPLIED STATISTICS

Adopting a life cycle view of statistics and practicing PSE and InfoQ assessments have implications on how projects are conducted or evaluated. To show this, we briefly revisit the case studies from the perspective of a life cycle view, impact assessment, and InfoQ evaluation; that is, the trilogy of applied statistics. We conclude this section with a mapping of the trilogy to statistical engineering.

The Automobile Part Example

The project started as a firefighting effort due to an unexpected increase in scrap. The effort focused on a short-term fix and was not driven by a more structural process improvement goal. Another implicit goal seems to have been the opportunity to practice design of experiments methodology on a local problem. This distinction is instrumental in how the project was handled. From a life cycle perspective, the project could have benefited from an expanded problem elicitation and goal definition clarification. Eventually the use of a designed experiment was overkill and a basic analysis of data with histograms would have revealed the cause of the problem. Upstream, the project seems lacking in operationalization of findings, communication, and impact assessment. In addition, the information quality of the project was not evaluated. An InfoQ assessment would have been a valuable exercise generating lessons learned for future improvement efforts.

The eBay Internet Auction Example

This is a research project on a trendy topic. The research goals were clear and the problem description was comprehensive. From a statistical life cycle view it covers all phases, including the communication and impact assessment phases. Moreover, the project InfoQ score was 73%, so we can clearly attribute to this project information of high quality. The two areas where more research could be done are operationalization and generalization of findings. This might involve collaborative work with psychologists and economists.

The Trilogy and Statistical Engineering

The goal of the work reported in this article is to expand the use of statistical thinking and methods and increase the effectiveness of statistical studies in the process. Recently, a methodology called *statistical engineering* has been proposed for dealing with large, complex, and unstructured problems. Hoerl and Snee (2010a, p. 52) define statistical engineering as "the study of how to best utilize statistical concepts, methods and tools and integrate them with information technology and other relevant sciences to generate improved results." As already mentioned, the development of the trilogy requires the integration of various tools and methods supporting its three building blocks. The life cycle approach provides a broad frame for solving a wide variety of problems from a diversity of fields of study. In order to expand and improve the contributions of statistics, a method to measure the impact of the statistical methods used is needed. A way to measure the quality of information produced is also needed.

The trilogy addresses a very big challenge: expanding the use of statistical thinking and methods. Success requires not only the integration of the three building blocks (life cycle approach, impact of statistical methods and information quality) but also the linking and sequencing of numerous statistical and nonstatistical tools within each building block. Statistical engineering provides needed concepts, roadmaps, and methods for achieving this. The development of the three building blocks was begun before the formalized statistical engineering approach was available and there is ample opportunity for synergistic integration. This will enable the use of the experiences and ideas of others working on solving large, complex, and unstructured problems and, in the process, speed up the development and effectiveness of the trilogy of applied statistics (Anderson-Cook and Lu 2012).

SUMMARY AND DISCUSSION

This article is an attempt to demonstrate, with examples, how applied statistics can expand its role in research, business, industry, and service organizations. The overall concept is that a trilogy consisting of a life cycle view, impact assessment, and an evaluation of information quality is necessary to fulfill this role. Adopting a life cycle view of statistics and practicing PSE and InfoQ assessments has obvious implications to research, education, and statistical practice. A promising development in this direction is the development of statistical engineering, which has recently gained worldwide popularity (Hoerl and Snee 2010b).

The main point is that effective statistical work is much more than properly applying statistical methods. In particular, one needs to emphasize that statistical analysis is a collaborative venture whose success depends essentially on the effectiveness of the communication between the statistician and the client. The statistician needs to exercise social and communicative skills for his work to have an impact. In addition, a systematic assessment of the impact and the quality of information generated by the statistical analysis requires additional activities not usually performed by statisticians.

To achieve these objectives, one needs tools. PSE and InfoQ are examples of such tools, but additional tools and methods need to be developed. Furthermore, joint research of statisticians with experts in cognitive science, computer science, and other disciplines on topics such as problem elicitation, communication, and formulation of findings can lead to breakthrough developments.

Our main message consists of three complementary recommendations:

1. Embrace a life cycle view of the role of statistics and develop skills and tools for problem elicitation and communication of results.
2. Learn how to assess the impact of work done with statistics. PSE is only an initial step.
3. Become better at generating knowledge and teach InfoQ in classes of statistics.

Developing the trilogy of applied statistics involves research initiatives and has implications to statistical practice and education. Moreover, it can form the foundation of a comprehensive theory of applied statistics.

ABOUT THE AUTHOR

Professor Ron S. Kenett, Chairman and CEO of the KPA Group, Research Professor, University of Torino, Turin, Italy, International Professor Associate, Center for Risk Engineering, NYU-Poly, NY, USA, Past President of the Israel Statistical Association (ISA) and of the European Network for Business and Industrial Statistics (ENBIS). Ron authored and co-authored over 160 papers and 10 books on topics ranging from industrial statistics, customer surveys, multivariate quality control, risk management and statistical methods in healthcare up to performance appraisal systems, information quality and integrated management models. He held academic positions at the University of Wisconsin-Madison, the State University of New York, Binghamton, Tel Aviv University and Bell Laboratories in New Jersey. The KPA Group he formed in 1990, is a leading Israeli firm focused on generating insights through analytics with international customers such as hp, 3M, Teva, Perrigo, Roche, Intel, Amdocs, banks and health care providers. Ron was awarded the 2013 Greenfield Medal by the Royal Statistical Society in recognition for excellence in contributions to the development and application of Statistics.

REFERENCES

Anderson-Cook, C. M., Lu, L. (2012). Statistical engineering—Forming the foundations. *Quality Engineering*, 24(2):110–132.

Berwick, D, Godfrey, A. B., Roessner, J. (2002). *Curing Health Care: New Strategies for Quality Improvement*, Jossey-Bass.

Box, G. (1997). Scientific method: The generation of knowledge and quality. *Quality Progress*, 30:47–50.

Cobb, G. W., Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly*, 104:801–823.

Columbia Accident Investigation Board. (2013). Volume I, Washington, DC: U.S. Government Printing Office. Available at: http://www.nasa.gov/columbia/home/CAIB_Vol1.html (accessed October 20, 2014).

Cox, D. (1981). Theory and general principle in statistics. *Journal of the Royal Statistical Society - Series A: Statistics in Society*, 144(2):289–297.

Daniel, C. (1969). Some general remarks on consulting in statistics. *Technometrics*, 11:241–245.

Godfrey, A. B. (2012). Foreword. In Faltin, F., Kenett, R. S., Ruggeri, F., Eds. *Statistical Methods in Healthcare*. Chichester, UK: John Wiley & Sons.

Godin, S. (2001). Really bad PowerPoint (and how to avoid it). Available at: http://www.sethgodin.com/freeprize/reallybad.html (accessed October 20, 2014).

Goodman, D., Hambleton, R. (2004). Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education*, 17(2):145–220.

Greenfield, T. (1987). Consultant's cameos: A chapter of encounters. In Hand, D. J., Everitt, B. S. Eds. *The Statistical Consultant in Action*, Cambridge, UK: Cambridge University Press, 11–25.

Hoerl, R. W., Snee, R. D. (2010a). Closing the gap: Statistical engineering links statistical thinking, methods and tools. *Quality Progress*, 43 (5):52–53.

Hoerl, R. W., Snee, R. D. (2010b). Moving the statistics profession forward to the next level. *The American Statistician*, 64(1):10–14.

Katkar, R., Reiley, D. H. (2006). Public versus secret reserve prices in eBay auctions: Results from a Pokémon field experiment. *Advances in Economic Analysis and Policy*, 6(2), Art. 7: 1–23.

Kenett, R. S. (2007). Cause and effect diagrams. In Ruggeri, F., Kenett, R. S., Faltin, F., Eds. *Encyclopedia of Statistics in Quality and Reliability*. Chichester, UK: John Wiley & Sons.

Kenett, R. S. (2012). A note on the theory of applied statistics. Available at: http://ssrn.com/abstract=2171179 (accessed October 20, 2014).

Kenett, R. S., Coleman, S., Stewardson, D. (2003). Statistical efficiency: The practical perspective. *Quality and Reliability Engineering International*, 19(4):265–272.

Kenett, R. S., Raanan, Y. (2011). *Operational Risk Management: A Practical Approach to Intelligent Data Analysis*. Chichester, UK: John Wiley & Sons.

Kenett, R. S., Shmueli, G. (2014). On information quality. *Journal of the Royal Statistical Society - Series A: Statistics in Society*, 177(1):3–38.

Kenett, R. S., Thyregod, P. (2006). Aspects of statistical consulting not taught by academia. *Statistica Neerlandica*, 60(3):396–412.

Kenett, R. S., Zacks, S., and contributions by D. Amberti. (2014). *Modern Industrial Statistics: With Applications in R, MINITAB and JMP*. 2nd ed. Chichester, UK: John Wiley & Sons.

Reynolds, T. J., Gutman, J. (1988). Laddering theory, method, analysis and interpretation. *Journal of Advertising Research*, 28:11–31.

Russom, P. (2011). Big data analytics. TDWI Best Practices Report, Q4. Renton, WA: TDWI.

Tufte, E. R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, CT: Graphics Press.