Trouble with Managerial Statistics

Friday, January 18, 2013

Farrokh Alemi, Ph.D. alemi@cox.net

John Large, Ph.D.

Arthur Williams, Ph.D.

Pracht Etienne, Ph.D.

Almost all health administration degrees require a course on Managerial Statistics, some outsourced to business school and others taught within the program. Table 1 shows the frequency with which a managerial statistics course is taught. The content of the managerial statistic course is seldom controversial. Seldom are these courses discussed during accreditation visits. There are widely available textbooks in managerial statistics that have standardized content. There are even open source textbooks that have led to widespread use of standard content [OpenIntro, 2012]. These courses attract little attention, have not changed in years, and do not stir controversy. Everyone "knows" what is taught in managerial statistics. It is generally focused on introduction to distributions (mostly Normal and other continuous distributions), hypothesis testing (comparison of paired or independent means), correlation and beginning of regression analysis. In this paper, we discuss the relevance of managerial statistics courses to new and emerging health administrators and discuss how the content of these courses need to change.

Table 1: Statistic Courses in Health Administration Programs at 20 Universities

	AB	CO	CR	GW	JH	MU	UA	LO	UC	UF	UI	UK	UM	MC	MI	MS	UP	SC	T	X
Biostatistics	0	1	0	0	0	0	1	1	0	1	0	0	0	1	0	0	1	0	0	0
Statistics	0	0	0	0	2	0	0	0	1	0	1	0	1	1	1	1	0	0	1	0

AB=Army-Baylor, CO=Columbia, CR=Cornell, GW=George Washington, JH=Johns Hopkins, MU=Marymount Univ, UA=Univ of Arkansas, UC=Univ of Colorado, LO=Univ of California Los Angles, UF=Univ of Florida, UI=Univ of Iowa, UK=Univ of Kentucky, UM=Univ of Miami, MC=Univ of Michigan, MI=Univ of Minnesota, MS=Univ of Missouri, SC=Univ of South Carolina, T=Trinity, X=Xavier, UP= Univ of Pittsburgh

In health administration curricula, some introductory statistics courses, as well as other quantitative courses such as operations research, data mining, decision analysis, are taught as survey of the field. Students are expected to know a little about different methods; none of the methods are covered in enough depth so that the student can apply lessons learned. When these students graduate, they are not expected to use quantitative tools. If they need to, they can bring in consultants to help them with analysis. Other quantitative and statistics courses are taught as a method of understanding the literature. In these courses, students are taught enough to understand the statistics reported in medical literature. Other statistic courses are taught as a necessary evil, a course needed for advancing other topics, but course content is not seen as inherent to the central tasks of management. In this paper, we re-think the role of statistics for managers and suggest a more central role for quantitative analysis.

Few educational leaders advocate for managers to have depth in statistical analysis. This attitude is often reinforced by senior industry leaders, some of whom are members of advisory boards of health administration programs. Industry leaders

typically ask for graduates to have leadership or communication skills but not analytical ones. It is not that they are hostile to analytical skills but they do not see its relevance. Perhaps an accountant might need to randomly sample accounts for inspection, or a finance manager might need to forecast trends, or a marketing person might need to survey customers. These are valid application of statistics to management but these are occasional tasks. They are not common everyday task of the manager. In the view of some industry leaders, real managers focus on people and not numbers. For example, James in Six Habits of Extraordinary Managers writes:

> "Average bosses focus on numbers rather than people. They jiggle revenue and profit numbers, monkey with statistics and data, and spend more time worrying about their spreadsheets than making things happen. Extraordinary bosses know that numbers represent only the history of what's happened–and understand that the best way to have great numbers is to make sure that that the job gets done. They realize that their responsibility is to manage people and their activities so the numbers take care of themselves [James 2012]."

These industry leaders make fun of "management by numbers." They consider managers who do so as "average" and "ordinary." They argue that management is a people business. They point out that leadership is an art and that no amount of numerical analysis can replace the insights of the manager.

This type of attitude shows in health administration advisory boards as well. While many board members are not against statistics, they are not for it either. From these board members we hear statements such as "we want graduates who write well." We seldom, if ever, have heard that we want graduates who have high analytical skills, who can do statistical analysis, who can analyze massive data and provide insights. Effective communication seems to matter most. The rest, even the content of what is being communicated, is secondary. This attitude also manifests itself in faculty meetings, where some faculty members argue that a focus on statistical analysis is needed for Ph.D. programs and management is a professional degree that does not need to emphasize statistics.

There is more than a bit of ironic short-sightedness in these beliefs, even if they provide aid and comfort to the innumerate. Two of the major contributors to people-oriented insights into management were persons with extraordinarily strong and cutting-edge quantitative skills. One is Herbert Simon whose book *Administrative Behavior* is a classic and is still read by some students of management [Simon 1947]. Simon won the Noble Prize in Economics and developed many of the early artificial intelligence algorithms with teams at Carnegie-Mellon University. The other is W. E. Deming whose original work was path-breaking method for statistical quality control [British Deming Association, 1992]. Regrettably, many programs in health administration (and business

schools) give more weight to “people-oriented” management adages that Simon dismissed 80 years ago, as “management parables” by showing that these are inconsistent, self-contradictory, and devoid of any useful content. Similarly, Deming used statistical insights to improve work processes [Salsburg 2002]. These two persons utilized what was learned from quantitative insights to generate changes in management that are deeper and longer lasting, one would suggest, than all the touchy-feely leadership books that disappear from the business and self-help shelves into oblivion.

Revolutionary Changes

Management is undergoing revolutionary changes [McAfee and Brynjolfsson, 2012]. More information is available, inside and outside the organization, than ever before. The availability of massive databases is referred to as “Big Data.” It assumes different forms. Inside organizations, large databases measure every change in the organization’s productivity and operations. Electronic health records have led to availability of data on operations within the organization. All sort of data are available. It is easy to see who is doing what and achieving which outcomes, therefore providers can be benchmarked. Data are available on true cost of operations as every invoice is tracked, so now contracts can be negotiated with real data at hand. Data are available on profitability of different operations. Data are available on pharmaceutical costs and its relationship to various outcomes. The wide availability of massive amount of data has made management with numbers easier and more common. Here are some examples of how health care organizations are gathering massive databases to enable insights to best practices [Jaret, 2013]:

1. The Personalized Medicine Institute at Moffitt Cancer Center tracks more than 90,000 patients at 18 different sites around the country.
2. In any given year, the Veteran Affairs Informatics and Computing Infrastructure (VINCI) collects data on more than 5.5 million patients across 153 medical centers.
3. Kaiser Permanente has a database of 9 million patients.
4. Aurora Health Care system has 1.2 million patients in its data systems.
5. The University of California’s Medical Centers and Hospitals has a database with more than 11 million patients.
6. United States Food and Drug Administration agency has combined medical records of more than 100 million individuals to track effectiveness of medications post release.

Growing use of electronic health record has enabled health care organizations, specially hospitals and insurance companies, to have access to large data.

In addition, data gathers in unexpected places. Patients' preferences, market share and competitive advantages can all be determined from analysis of comments and text left on the web [Alemi, Torii, Clementz and Aron, 2012]. Massive amounts of data are available on the web. These data can help managers identify consumer's preferences. Most of these data take the form of text. While some consider data to be numbers, in reality, both numbers and text are data. Analysis of the Web and emailed comments requires statistical skills and is one area where "Big Data" tools are making a difference.

Change in Competencies

What managers do has been changing. Today's managers rely heavily on continuous quality improvement and Lean management. These techniques use statistical process control. Through these techniques managers bring data to analysis of operations. The use of these techniques is widespread and an essential capability of modern managers [Vest and Gamm, 2009].

The availability of data has enabled managers to go beyond traditional roles and address clinical questions. For the first time, managers can measure comparative effectiveness of difference healthcare interventions. They can talk to physicians about their clinical practices. For years, managers have left decisions on quality of care to clinicians; thereby they have restricted their decisions to a small portion of the operation. They have ignored the central part of their business. Data is changing this equation. Today, managers have data on what is best for patients and can work with their clinicians to change practices. For example, analysts have been able to examine pairs of drugs that cause a side effect not associated with either one. They found that Paxil, a widely used antidepressant, and Pravastatin, a cholesterol-lowering drug, raise patient's sugar, a problem not associated with either one [Tatonetti, Ye, Daneshjou and Altman, 2012]. In this example, and other comparative effectiveness studies, we see an emerging new role for managers.

Data Driven Organizations Perform Better

"Big data" is changing which managers succeed. "As the tools and philosophies of big data spread, they will change the long lasting ideas about the value of the experience and the practice of management [Eshkenazi 2012]." The expectations is that companies that get insights through analysis of big data will do better and therefore these managers will succeed more often. There are many examples of how data driven companies do better. At Mercy Hospital in Iowa City, Iowa managers who benchmark their clinicians and pay them for performance report 6.6% improvements in quality of care [Izakovic 2007].

Many investigators point out that the Veteran Administration was able to reinvent itself because of its focus on measurement of performance [Longman , 2010]. The VA healthcare system was known for poor quality. Kaiser and colleagues set out to change the culture of VA to a data driven company. Over a short interval, they were able to not only change the culture but change patient outcomes. According to Longman, a focus on measurement and data has radically changed the performance of this system of care; it now reports some of the best outcomes for patients anywhere in United States.

A recent study of 330 North American companies showed widespread attitudes towards data. The more companies characterized themselves as data driven, the more they were likely to outperform their competitors on financial and operational results. Data driven companies were 5% more productive and 6% more profitable than less data driven companies [Brynjolfsson, Hitt and Kim, 2011].

In health care, companies that rely heavily on Lean and various process improvements can be classified as data driven, albeit relying on small data sets. These companies use process improvement tools to introduce statistical analysis within the organization. It is not clear that organizations that use statistical process improvement are financially better off or have larger share of the market. Many studies show that when organizations fully implement these tools, including the emphasis on measurement [Nelson et. al, 2000], they deliver better care at lower cost [Shortell, Bennett and Byck, 1998]. The part that is not clear is whether better outcomes leads to more financially sound organizations.

In healthcare, the use of electronic health records has been associated with reductions in medication errors [Stürzlinger, Hiebinger, Pertl and Traurig, 2009]. Mangers have used electronic health records to maximize reimbursement, in ways that have surprised insurers [Abelson, Creswell, and Palmer 2012]. Other managers report analyzing data within electronic health records to reduce “never events” within their facilities and to measure quality of care [Glaser, Hess 2011]. These data suggest that managers are finding ways to use data within electronic health records to improve their organizations. These efforts are expected to continue creating an unprecedented shift to heavy use of data.

The transformation is most apparent in health insurance industry. SeeChange pays its members a cash reward to complete a customized compliance program. The company uses personal health records, claims databases, lab feeds and pharmacy data to identify what actions should be rewarded with cash incentives. Other insurers are using patient satisfaction data and health outcomes to optimize their provider networks.

Transition

It may not be easy to transition to evidence based and data driven management [Barton and Court, 2012]. The data are available, removing the largest barrier; but other barriers remain. The personnel needed for the analysis are not readily available [Davenport and Patil, 2012]. Once the senior health administrators see the benefits that could emerge from data driven companies, the race is on. The use of big data is the natural end of the push to electronic health records. The old adage of "build it and they will come" applies here. Someone has gone ahead and built large databases in health care. What remains is for managers to take advantage of these data. To take advantage of these opportunities, managers need new training programs.

Surprisingly, data warehouse are more useful and more available to middle than top managers. It used to be that middle managers did not have access to the same data as the top level managers, who might have commissioned special data collection efforts. The electronic health record has changed the equation. Data are available to all managers, from top to bottom. The lower the rank of the manager the more time he or she has to delve into the data. There are no organizational restrictions on middle managers to use data within electronic health records. The key barriers are imagination and skills. Middle managers unprecedented access to data is new and changes what skills these managers need.

Data is no panacea. Management intuition, the art of management, effective communication, is still important and useful. A vision is needed before data are analyzed. Organizational leaders must select a direction, set priorities, resolve conflict and do a lot more than data analysis. But data can inform all aspects of the organization, it can support the vision. Data can reduce conflict. What is measured is what is accomplished, data can motivate change. Data can reveal market preferences. Evidence based management can help middle managers gain more insight about what is working well and what needs improvement.

New Breed of Managers

Capitalizing on big data depends on hiring a new breed of managers who have extensive quantitative skills and embedding these skills in managers who have day-to-day responsibilities requiring lesser skills. In investment firms, this new breed of bankers with extensive quantitative skills is known as "Quants" and they rely on their statistical and computer models to trade. Several educational organizations specialize in training and producing "Quants." In health care, it is not clear who trains managers with strong quantitative skills. While some quantitative skills are taught to all managers, as noted, what is taught lacks depth and foci.

For example, consider changes in roles of marketing managers. In the past they have been focused on sales, promotion and market segmentation. They may have been responsible for conducting satisfaction surveys, often actually conducted by outside contractors. Now, marketing managers need to understand "web sentiment." They need to analyze comments made on the web, set benchmarks and understand changes in the Web sentiment. Who can an organization that wants to understand competitive data on the web turn to? Not current students in health administration programs as their exposure to big data is limited and they do not have the statistical skills to analyze data on the Web.

Consider another example. A Lean process improvement effort wants to know if their efforts have led to lower surgical infections. The data on surgical infections are available in the electronic health record and therefore the team wants to analyze these data as they emerge over time. Many graduates of current health administration programs cannot do this simple analysis because (a) they do not know how to get data from electronic health records using Standard Query Language (SQL) and (b) they do not know statistical process control and therefore cannot analyze time-based data.

What Is Taught; What Is Needed

Managerial statistics courses are focused on managerial issues only in name. Many of these courses cover the same topic as a standard introductory course in statistics. The claim that these courses fit management is based on the examples provided and not on the methods taught within these courses. These courses focus primarily on hypothesis testing. Few managers test hypotheses. These introductory courses focus on continuous distributions. Managers have access to data that are discrete. These courses examine repeated samples from stable processes but managers deal with constantly changing processes. Many assumptions of statistical tests are not met in continuously changing processes of care. Managers need to use statistical process control tools to analyze time based data and these procedures are often not covered in introductory managerial statistics courses.

Even before managers analyze their data, there is a more fundamental problem. The data are dispersed in the medical record. Data are in hundreds, if not thousands, of tables within medical records. Data are on the Web in thousands of sites. First and foremost, managers need to organize the data, a set of skills not taught in introductory managerial statistics courses. They need to merge data from several sources, which they often do not know how to do.

Table 2 lists a set of difference between current and proposed managerial statistics courses. From data, to methods of inference and analysis, the tools are different and the skills needed are different. Obviously not all of these tools and

methods can be taught in an introductory course. What is important is to what extent the problems addressed in the managerial statistics course reflect real problems that current managers face as opposed to imaginary problem sets that are designed to make the application of existing statistical tools in the text look easier and relevant.

Table 2: Current and Proposed Content in Managerial Statistics Courses

Topic	What is taught?	What is needed?
Distributions	• Focus is on Normal, Uniform and other continuous distributions with little coverage of discrete probability theory	• Focus on probability distribution in discrete events, including Bernoulli, Binomial and Geometric distributions. Normal and Poisson as approximation of Binomial distribution
Data	• Single measures and not time based data • Time consuming prospective data collection efforts	• Longitudinal, time based, repeated measures • Text and numbers • Rich data with multiple interpretations • Focus on observational and retrospective extant data
Confidence Interval Estimation	• Normal distribution estimation of confidence interval	• Estimation of upper and lower control limits in process control charts • Bootstrapped estimates of variability.
Methods of Inference	• Comparison of mean to population • Comparison of two means • Paired t-test and comparison of dependent means • Analysis of Variance	• Statistical process control tools such as XMR charts, p-chart, time-between chart, Tukey chart • Risk adjusted process control tools • Use of simulations
Methods of Forecasting	• Correlation analysis • Multivariate Linear Regression analysis • Logistic regression	• K-Nearest Neighbor • Local regression • Text analysis • Causal analysis • Survival analysis
Study Design	• Factorial design of experimental studies • Surveys	• Matched case control using observational data • Data visualization for extant data

The use of statistics for forecasting is obviously important but it is usually not covered in introductory managerial statistics courses. Market surveys are obviously important but also not taught in managerial statistics courses. The organization of data

from multiple sources is important but again is not adequately addressed. Finally, analytical tools such as control charts are important but again not addressed in many current managerial statistics courses. They are taught in quality improvement courses but without sufficient depth to (a) conduct risk adjustments and (b) to cover various types of data. In the end, the trouble with managerial statistics courses is that they do not cover the tools that managers need.

Managers need to understand how to look at data. Analysis is messy, with lots of overlapping redundant data that provides contradictory insights. Data-intensive managers are swimming in data. They need to structure large amount of data before they could address a specific problem. The data they work on is rich with text and numbers. It is collected passively. Over time, more data are available and one major task of the manager is to decide what is not relevant. The data itself never stops flowing and the manager must decide what time period he/she would like to examine and why. A course on statistics, even a beginning course in statistics, needs to teach students how to select and manipulate data. At some point, an exposure to Standard Query Language may be useful so that students can organize their analysis from numerous tables of data.

Statistics Through Out the Curriculum

Data scientists and quantitative managers need a lot more than an introductory course in managerial statistics. These types of managers are not made in one course but in use of quantitative methods throughout the curriculum. A review of curriculum shows that many courses are already focused on quantitative analysis. In addition to the courses listed in Table 1, there were many other courses that at first glance do not seem to be a quantitative course, but the modern version of these courses would require different analytics. Marketing courses need to expose managers to sentiment analysis. Accounting courses need to expose managers to fraud detection statistics. Courses on strategy planning need to expose students to casual and network analysis. All of these new approaches require analytical tools. Health administration programs that wish to prepare graduates that excel in quantitative methods need to rethink the content of many courses.

Discussion

Not every management student needs to become a statistician. Not every manager needs to rely on quantitative analysis. Some may need a general introduction to statistical methods. Others need more depth. But all future managers will be ill served if their exposure to the single course on statistics does not fully prepare them to understand the full set of options available. At a minimum, this first course should

include statistical process control. Control charts are the building block of Lean and numerous other modern managerial techniques. Its absence in managerial statistic courses is surprising and a problem for the entire curriculum. In these introductory courses, the focus on hypothesis testing is unnecessary and may do more damage than good. It may turn off the student from pursing more quantitative analysis. Statistical process control tools introduce the concept of hypothesis testing in the context of control limits. As such they provide a rigorous introduction to statistics as well as to applications in management.

Of course, some would want to do more than statistical process control. For these students a concentration in data analytics would properly prepare them for the new emerging managerial roles. Health administration programs that want to prepare managers that rely on data should build quantitative problem solving into the entire curriculum. In marketing, strategic planning and human resource management, students should have hands-on quantitative examples that demonstrates the use of data.

A new breed of managers is needed who can take advantage of the vast amount of data available to them. This future is built on that first course: managerial statistics. If this course is not set straight, then the foundation is problematic. The real trouble with managerial statistics, as it now exists, is that it does not open a window to this new emerging world of quantitative managers. The trouble with managerial statistics is that it has no revolutionary zeal.

References

Abelson R, Creswell J, Palmer G (2012). Medicare Bills Rise as Records Turn Electronic, New York Times. http://www.nytimes.com/2012/09/22/business/medicare-billing-rises-at-hospitals-with-electronic-records.html?_r=1&pagewanted=all&_moc.semityn.www Accessed October 26th 2012.

Alemi F, Torii M, Clementz L, Aron DC (2012). Feasibility of real-time satisfaction surveys through automated analysis of patients' unstructured comments and sentiments. Qual Manag Health Care 21(1):9-19.

Barton D, Court D (2012). Making advanced analytics work for you. Harv Bus Rev. 90(10):78-83, 128.

British Deming Association (1992). A Brief History of Dr. W. Edwards Deming British Deming Association SPC Press, Inc.

Davenport TH, Patil DJ (2012). Data scientist: the sexiest job of the 21st century. Harv Bus Rev. 90(10):70-6, 128.

Brynjolfsson E, Hitt L, Kim H (2011). Strength in Numbers: How Does Data-Driven Decision Making Affect Firm Performance? http://www.a51.nl/storage/pdf/SSRN_id1819486.pdf

Eshkenazi A. APICS (2012). The Association for Operations Management. http://www.apics.org/omnow-landing-page/2012/10/26/joining-the-big-data-revolution accessed on October 29th 2012

Glaser J, Hess R (2011). Leveraging healthcare IT to improve operational performance. Healthc Financ Manage 65(2):82-5.

Izakovic M (2007). New trends in the management of inpatients in U.S. hospitals--quality measurements and evidence-based medicine in practice. Bratisl Lek Listy 108(3):117-21.

James G (2012). Management & Leadership: 6 Habits of Extraordinary Bosses at http://business.time.com/2012/09/27/6-habits-of-extraordinary-bosses/#ixzz2AQmI7YEd September 7th 2012, accessed October 26th 2012.

Jaret, P (2013). Mining electronic records for revealing health data. New York Times, January 14, 2013, http://www.nytimes.com/2013/01/15/health/mining-electronic-records-for-revealing-health-data.html?pagewanted=all&_r=0

Longman P (2010). Best Care Anywhere: Why VA Health Care Is Better Than Yours. Berrett-Koehler Publishers; Second Edition, Revised edition.

McAfee A, Brynjolfsson E (2012). Big data: the management revolution. Harv Bus Rev. 90(10):60-6, 68, 128.

Nelson EC, Splaine ME, Godfrey MM, Kahn V, Hess A, Batalden P, Plume SK (2000). Using data to improve medical practice by measuring processes and outcomes of care. Jt Comm J Qual Improv. 26(12):667-85.

OpenIntro (2012), See for example http://www.openintro.org/stat/textbook.php accessed October 26th 2012.

Salsburg, D. (2002) The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century, W.H. Freeman / Owl Book.

Shortell SM, Bennett CL, Byck GR (1998). Assessing the impact of continuous quality improvement on clinical practice: what it will take to accelerate progress. Milbank Q. 76(4):593-624, 510.

Simon, HA. (1947). Administrative Behavior: a Study of Decision-Making Processes in Administrative Organization (1st ed.). New York: Macmillan.

Stürzlinger H, Hiebinger C, Pertl D, Traurig P (2009). Computerized Physician Order Entry - effectiveness and efficiency of electronic medication ordering with decision support systems. GMS Health Technol Assess. 19; 5: Doc07.

Tatonetti NP, Ye PP, Daneshjou R, Altman RB (2012). Data-driven prediction of drug effects and interactions. Sci Transl Med. 4(125):125ra31

Vest JR, Gamm LD (2009). A critical review of the research literature on Six Sigma, Lean and StuderGroup's Hardwiring Excellence in the United States: the need to demonstrate and communicate the effectiveness of transformation strategies in healthcare. Implement Sci. 1; 4: 35.