TEACHING ENGINEERING STATISTICS TO PRACTICING ENGINEERS

Jorge Luis Romeu
Syracuse University, United States
jromeu@ecs.syr.edu

Teaching statistics to engineers is a challenging task. First, lacking space, many engineering curricula include few or no statistics courses, and these are often packed and highly theoretical. Thence, students don't perceive statistics as a part of their engineering toolkit, but as a nuisance to endure. On the other hand, engineering is a two-part endeavour. One consists in building or modifying systems. The second is measuring/assessing system performances, which are nothing but random variables. Therefore, there can be no engineering work without statistical analyses. In this paper we discuss and assess ways to enhance the insufficient statistical education that many engineers receive once they have left college. Such methods, designed for practicing professionals include (print and electronic) materials produced for self-study, short training courses and the development of industry-academe organizations to help practicing engineers by "looking over their shoulders." Finally, a selection of related free Web Sites are presented.

INTRODUCTION

The statistical education of practicing engineers is of pressing importance for two main reasons. First, statistical methods are intimately linked to sound engineering practices. Second, statistics is insufficiently, inadequately and too theoretically taught in engineering schools.

On the other hand, statisticians have spent big efforts on the education of school children, undergraduate and graduate students. But this author has found very little about the problem of after-college engineering education. Due to the brevity of this paper we can only overview a few educational efforts: first, among statistics organizations, then, among individual educators.

Centers for Statistical Education, designed to study how school students learn statistics have been organized at ASA (http://www.amstat.org/education/index.cfm?fuseaction=main) and the RSS (http://www.rsscse.org.uk/). IASE (http://www.stat.auckland.ac.nz/~iase/) the education section of ISI, has studied how statistics is taught and learned at every level, from early childhood to the graduate school, publishing their research findings in their electronic journal SERJ. The CAUSEWEB Consortium (http://www.causeweb.org/) was instrumental in organizing USCOTS, the American version of ICOTS, and includes much material on undergraduate education. Other good education projects are Merlot (http://taste.merlot.org/), on multimedia teaching and learning, ARTIST on course assessment (https://app.gen.umn.edu/artist/) and Technomath, about IT and math training in the industrial workplace (http://www.ioe.ac.uk/tlrp/technomaths/).

At the individual level, we mention the work of Batanero, Bisgaard, Bickel, Bailar, Blumberg, Chance, Garfield, Joliffe, Locke, MacGillivray, Moore, Ottaviani, Pearle and Wilde, among the many researchers, whose paper references can be found through the above-mentioned web pages, journals, organizations and institutions, or through Google, or its new Scholar version.

Improvements in the way statistics is taught in engineering schools have been proposed by Hogg *et al*. (1985) and Kettenring (1995), among others. Spedding (1998 and other papers) has proposed media improvements, and Romeu has advocated the use of practical applications, student projects, group learning, and technology and statistical software, as far back as 1986 and as recently as 2004. His SU course web page http://web.syr.edu/~jlromeu/ecschedul.html includes examples of the use such practical approaches in graduate engineering statistics classes.

Notwithstanding the excellence in the mentioned work, they do not address the problem of after college statistical education (Technomath pursues workforce training, including stats, but at basic level). We find, after our 30 years teaching in engineering departments, and consulting in quality and reliability, that many engineers have difficulty understanding concepts like confidence intervals, hypothesis tests or modelling. For example, some assess regression via its Index of Fit (R^2), instead of using the t or F statistics or the p-values, and fail to check all model assumptions.

In the rest of this paper we show some ways practicing engineers use to cope with their lack of statistics background, learning the subject on their own, on the job, we discuss some problems involved in some of these learning endeavours, and we propose some solutions.

AN ENGINEERING SURVEY

In order to support our working hypothesis, that the statistical education of engineers is deficient and inadequate, and further investigate the efforts to solve this, we implemented a pilot survey. Our questionnaire (http://web.syr.edu/~jlromeu/SurveyICOTS.html) was submitted to the members of several quality, reliability, manufacturing etc. engineering chapters in New York and Florida (via the Isostat ASA group), and to members of ENBIS (http://www.enbis.org/).

We ask how practicing engineers have acquired their post-college statistical knowledge: (1) reading books, journals, manuals or other hard copy materials, (2) reading Web and Internet materials, (3) following on-line courses or learning software, etc., (4) attending conferences and chapter meeting talks, (5) pursuing preparation for professional certifications, (6) taking short training courses, (7) receiving mentoring from more experienced work colleagues and (8) other sources, such as hands-on (practical) working experiences, and taking Six Sigma training.

The survey data collected so far has yielded the following results:

Table 1: Methods of Acquiring Statistical Knowledge after College

Method	HardCopy	Internet	Software	Chapt. Mtgs	Certification	Courses	Mentoring	HandsOn
Percent	0.272	0.099	0.02	0.067	0. 199	0.127	0.15	0.066

To characterise the survey taker we asked their (9) education level, (10) area (academe, industry, government), (11) specialty (mechanical, electrical, industrial, etc.), (12) total number of stats courses taken in college, (13) years practicing engineering, (14) country and (15) gender.

To date we have received 62 responses, of which 59 are from the US, 3 from abroad. Of these, 8% were females, 56% had graduate degrees, 60% had 16+ years of experience and 90% were from industry. Among all, 16% have taken no stats courses in college (33% among BS), 38% took only one (38%) and 26% have taken 2 courses (24%). We continue collecting data.

We recognize that this is a "self-selected," but not a random sample. But for a ground-breaking, pilot study as this one, we expect that it is large enough to adequately illustrate (and help characterize and support) the working hypotheses and problems discussed in this paper.

The first issue to signal out is that 1/3 of engineers with a BS degree have never taken a single statistics course in college, and that another 1/3 of them have taken only one. This means that the majority of BS level engineers (2/3 in our survey) have either no statistical training or a very deficient one (a single course). Engineers that pursue graduate school (MS, PhD) have a larger opportunity of taking stats, and only 7%, in our sample, have never taken one.

Survey also shows how "readings" constitute the preferred means of learning: books and journals, as well as web tutorials, provide 38% of stats knowledge. Interestingly, the use of web tutorials (10%) is increasing with time: older engineers prefer hard copy, whereas younger ones read web-based material. Short courses, and certifications and Black Belts training, are also two important methods of learning (33%). Finally, the mentoring received from more experienced colleagues and hands-on (learning by doing) also constitute frequent learning activities (22%).

In the next sections we analyze these survey results and discuss how they relate to the education methods that practicing engineers are currently using to learn statistics after college, and how these methods, under the present circumstances, constitute a deficient remedial system. We finally present specific proposals on how to improve them, and on how to fund such efforts.

HARD COPY AND WEB-BASED TUTORIALS

The most popular way for engineers to enhance their statistical knowledge is through individual readings of hard copy (28%) and web-based tutorials (10%) containing detailed, step-by-step numerical examples on specific statistical topics with practical applications. Hard-copy (books, journals, etc.) readings are still prevalent, but the ubiquitous Internet is taking over.

We have been developing precisely these types of web tutorials for the past several years, supported by the government-sponsored Reliability Analysis Center. The reader can find over 20 such stats tutorials in http://web.syr.edu/~jlromeu/urlstats.html, covering distributions, confidence interval derivation, testing, SPC, reliability and availability modelling and data analysis, etc.

There are other Web Sites with extensive developments in the theory and practice of statistics. We present in the Appendix, as a resource for practicing engineers, a selection of their URLs. Some web tutorials can be found through professional organisations such as the American Statistical Association (ASA), American Society for Quality (ASQ), Inter American Statistical Institute (IASI), or the Royal Statistical Society (RSS). Other sites with statistics education content belong to peer-reviewed publications such as SERJ, Chance or JSE, or projects like CAUSEWEB. Others can be found in personal pages of statistics educators (Leon, R.). There are entire books about statistics on the Web, such as the handbook by the National Institute of Standards and Technology (NIST) or the one by Math World, both in English, or the textbook by J. Lejorza, one of the very few statistics materials we have found in Spanish.

Tutorials are easily downloaded and perused but their quality varies. Not all are self-explanatory or have detailed, step-by-step numerical examples. In addition, we find three serious problems associated with most Web-based tutorials: identification, language, and internet access.

The first problem is the lack of a single point of departure which may list, classify and sequence such tutorials. Often, there is excellent material available that engineers don't even know exists –let alone where to find them, or in what order to read them. The usual procedure is to explore a topic, via Google or other such search engine, and sort through the many results that one obtains, some relevant and other that are not, a very slow and inefficient procedure.

The second problem is that most of this material is in English, which has become the international language of science and peer-reviewed publications (for, authors want to reach the broadest audience possible). However, many Developing World engineers and statisticians do not have adequate English reading skills and, therefore, cannot use such excellent and free materials.

The third problem is that, in order to access such materials, potential users require an up-to-date computer and operating system, Internet access and sophisticated web browsers. Again, Developing World engineers, the ones who need most such information (for they cannot afford buying the books or subscribing to journals and professional associations) may also miss out on these opportunities because their hardware and software is often older, and even obsolete.

Not withstanding the above constrains, Web Tutorials, can be very effective for the practicing engineer who, from office or home, and on their own time, can obtain the necessary statistical background that was not provided during their university curriculum, but that is crucial for their professional practice. Web tutorials are poised to become the learning tool of the future.

PROFESSIONAL COURSES

Intensive, short, professional courses are a very popular alternative to taking the longer, more structured sequence of courses in graduate school. The sample reflected that 20% of the statistics knowledge of practicing engineers is obtained via short courses and Black Belt training. But courses often focus on a single topic, developing it at scant length, and seldom in-depth. They are expensive and vary in quality. With the proper content and background, they provide a useful experience, readily applicable by the practicing engineer who is usually in need of a tool that solves an immediate problem. However, such immediacy constitutes its main drawback.

For example, an engineer may take a course in regression, without adequate background in testing and estimation. The result may well be just a training course on how to use regression software. In addition, regression limitations and underlying assumptions, or the procedures to check them, are often not included, or not fully explained in the short course, for lack of time.

On the other hand, if topics are taught at full length, then short courses become Academe-like, and defeat their purpose of providing a valid alternative. A middle-of-the-road course of action needs to be devised that, while still teaching statistics properly, does it in a manner that circumvents the long detours that characterize the Academic environment.

This author has designed and taught "reliability engineering statistics" courses since the mid 1980s, when it was a five day course and covered many topics. In the 1990s we redesigned it into three days (http://src.alionscience.com/pdf/RAC-1ST/Statistics_Training_Course.pdf), due to the pressures of the times. The first day, it covers EDA and the basics of the main distributions used in this area. The second day, it covers interval estimation and hypothesis testing, and the third, an introduction to regression and ANOVA. We teach 6 hours a day, plus 2 hours of Lab.

Our course provides the statistics base for all engineering courses that our organization offers, such as electronic, mechanical and software reliability, accelerated testing, reliability growth and reliability centred maintenance, all of which use statistics intensively. But most engineers still register directly to the specific courses they are interested in and, lacking the proper statistics background, miss out in the many statistical aspects that these necessarily cover.

The main problems we have found teaching short statistics courses are the disparity in student levels, as well as the large differences in their expectations. We had senior engineers who had not taken a formal statistics course in twenty years, but who frequently work with this topic, as well as novices with little idea of what statistics is. We had managers and data collectors, only interested in recognizing a statistical problem, or in using the information they are gathering, as opposed to analysts, who actually implement the statistical methods on real data.

Courses are taught in open format, where anyone can register, adding to the large heterogeneity. Others are in-house, where students know each other, share data and discuss their applications. Regarding software, *Minitab* is easy to learn and use but very often unknown or unavailable to students. *Excel*, in spite of its technical limitations, is in every student's computer.

We use our web tutorials on distribution assessment (via Goodness of Fit tests as well as empirically) to emphasize Exploratory Data Analysis (EDA). We check the validity of each assumption before implementing confidence intervals, Chi-Square, and *t* and *F* tests. We teach the importance of checking, at least graphically, regression and ANOVA model assumptions (independence, normality and homoskedasticity), and emphasize modelling the process, and not the data.

We implement *Minitab* Execs to illustrate specific problems (e.g., how regression and ANOVA models change, as sample size and error variance change, etc). We gather class data (age, sex, etc.) to show issues in data collection, and then use them in simple, group exercises.

Engineering professional organizations (ASQ, IEEE) also teach short courses during their annual meetings, as well as for obtaining and keeping their reliability, quality, six sigma etc. certifications. But they also suffer from the above-discussed lack of unifying structure problems.

PROFFESIONAL ORGANIZATIONS AND ACADEME-INDUSTRY INSTITUTES

A third way to provide statistical knowledge to practising engineers is via professional organizations activities such as chapter meetings and conferences (7%), certification training and exams (20%). However, such activities are as uncoordinated as the independent web or hard copy readings, lacking a unifying thread that allows a structured acquisition of statistical knowledge.

Hence, we proposed the organisation of "Academe-Industry-Professional Institutes." The partnerships would serve practicing engineers as well as students still in college, to gain valuable, practical statistical knowledge. Such local institutes could provide the missing structure in courses and tutorials and hands-on experience (see our piece in the Amstat News of April, 2005).

The importance of industrial statistics has been readily recognized by those organizations who can afford it. For example, the US military require an assessment of quality and reliability in their equipment. They fund projects such as the Reliability Analysis Center (now, RIAC) to provide consulting, internet and in-house training, and other statistical services, many of them free or at nominal cost, for agencies of the US Department of Defense and for its contractors.

Large industries can also afford a statistics department. But small or medium engineering companies usually cannot. An Institute would provide some of these services, either free or at a very low cost, to qualifying small and medium-size companies who cannot afford a full time industrial statistician. Such Institutes would operate in the following way:

Engineering students taking industrial statistics courses would work as interns or research assistants. Hence, operating costs would be very low (mainly student stipends). Under a professor's guidance, students would provide free statistical assessments and assist qualifying industries in implementing their recommendations. Students would get hands-on experience, faculty would gain better insight about industry needs, and industries that cannot afford it would receive this service. Finally, the engineering professional organizations could use these professors and students as mentors and instructors in their presentations and distance learning (DL) courses.

Such an idea is not unrealistic. The US Department of Energy supports such institutes (one is at Syracuse University: SU-IAC). It provides free audits in the area of energy that save industry (and the country as a whole) substantial money. Institutes would not compete with consultants or for-profit organizations, for they do not engage paying customers. Moreover, Institutes would provide a "proof of concept," as well as awareness on the subject, to smaller organizations, thus increasing the possible pool of future industrial statistics consumers.

Practicing engineers could then take well structured and sequenced short courses and workshops at these local Institutes, and also attend periodical presentations on statistical subjects. Such workshops could use the freely available material already in the Internet in their courses, facilitating their reading by the practitioners. In addition, engineers would have a place where they could periodically go for consultation with professors or advanced students, and where they would network and establish contacts that would mutually help each other grow professionally.

Finally, the survey shows how mentoring from more experienced colleagues, hands-on trial and error experiences, etc. contributes the remaining 15% of how engineers acquire statistical knowledge. The said institutes could also contribute as an organized mentoring system.

DISCUSSION

There are two facts regarding practicing engineers and industrial statistics that have to be confronted and resolved. One is that engineers need to apply statistics in their work, to measure and compare product characteristics that are intrinsically random variables. The second fact is that many engineers, for lack of space in their curriculum, do not receive enough, or adequately taught statistics in college, a situation that is not likely to change in the near future. As a result engineers are forced to seek, once they reach the workplace, alternative ways of acquiring the statistical knowledge that they lack, in order to capably function in their careers.

We have discussed several means of doing so: professional courses, hard copy and web tutorial readings, and Professional-Industry-Academe "institutes." These are not necessarily the only ones, or the best approaches; just the most widely used at this time.

On the other hand, two main problems arise when engineers seek statistical knowledge on their own, after leaving college. The first is that the existing material and procedures (readings, courses, certifications, etc.) constitute a collection of unstructured, unorganized, uneven elements that, in order to become efficient, require a thorough reorganization and classification. The second issue is that, to provide such required structure and reorganization, substantial funding is needed.

Unfortunately these types of activities, even when they are of general benefit to society, are usually not very profitable. Thus, they have to be sponsored by socially-minded organizations such as the government and its research agencies, universities, professional societies and larger industry that have the skills, resources, organization and authority to carry them out successfully.

For example, in the United States the armed forces and the departments of commerce and education support research agencies that play a prime role in such work (e.g., NIST, NSF, NIH, NCES, and DOD-IACs). Such government research agencies and organizations have provided the policies, direction and support to pursue innovations in many fields that flourish today.

In Britain, the Institute of Education of the University of London supports the already-mentioned Techno-Math Literacy in the Workplace project. It trains workers in the plant floors with some basics in applications of IT, mathematics and statistics. Their interesting project differs from our proposal in scope (general math vs. industrial statistics) and level (basic vs. specialized).

In addition, there are several other organizations that are able to contribute to this effort. Universities, especially via their DL training and life-long education departments, and via the proposed Professional-Academe-Industry institutes, also have a very important role to play. They will gain the most through these efforts, by better educating their own students and faculty as well as by drawing more individual and corporate support, from their own communities.

National and international professional engineering and statistical societies such as the statistics divisions of ASQ and IEEE, and the physical sciences and quality divisions of ASA, in the US, and their equivalent in other countries, also have a very important role to play. They can become local organizers, supporters or mentors, providing statistical oversight, structure, books, materials, etc., and specialized faculty for teaching better structured short courses and workshops

For developing and poor countries, lacking these organizations and resources, there are international projects such as the Fulbright Speakers Specialist Roster or the Juarez Lincoln Marti that provide free, to qualifying universities and organizations, faculty specialists that travel to the host country and work, for several weeks. The only obligation of the host organization is to lodge and feed the Specialists during their stay. Our experience working as Fulbright Speakers and Juarez Specialist in Latin America has been very rewarding. Further information about how these operate, and how to contact them, can be found in their respective Web Pages (Appendix).

Finally, big industry also has an important stake in helping raise the statistical knowledge of practicing engineers. Many large industries subcontract part of their work to smaller industrial units: their suppliers. They then have to spend a large amount of time and resources controlling and assessing their products' quality and reliability. Better statistically trained engineers in their supplier industries, can help decrease these larger corporations' production costs, increasing their product quality and reliability, and therefore, also increasing their sales and profits.

Better statistical education of practicing engineers benefits all: the engineers themselves the corporations, government and the public. All stake-holders must be conscious of these benefits and be prepared to contribute the means to undertake the work discussed in this paper. When all beneficiaries realize this, the required reorganization will finally take off.

ACKNOWLEDGEMENTS

We gratefully appreciate and recognize the comments and suggestions of our three reviewers, that helped improve the presentation of this paper.

REFERENCES

Hogg, R. *et al*. (1985). Statistics education for engineers: An initial task force report. *The American Statistician*, 39, 168-175.

Kettenring, J. (1995). What industry needs. *The American Statistician*, 49, 2-4.

Romeu, J. L. (1986). Teaching engineering statistics with simulation. *Journal of the Institute of Statistician (RSS Series D)*, 35(4).

Romeu, J. L. and Land, M. (2004). Group learning, contextual projects, simulation models and student presentations in enticing engineering statistics students. In *2004 Proceedings of the Statistical Education Section*, Toronto. Alexandria, VA: American Statistical Association. More papers in http://web.cortland.edu/romeu/research.html.

Spedding, T. (1998). Teaching statistics to engineers: Towards developing a multi-media teaching environment. *Proceedings of ED MEDIA and ED TELECOM-98*, (pp. 1333-1338). More information in http://www.uow.edu.au/commerce/mgmt/pdffiles/trevor.pdf.

APPENDIX: URLs of Selected Statistical Education Organizations:

American Statistical Association. http://www.amstat.org; its Statistical Journals on-line http://hermia.asa.catchword.org/vl=12097200/cl=48/nw=1/rpsv/home.htm and its Section on Engineering and Quality and Productivity (http://www.amstat-online.org/sections/qp/).

American Society for Quality, http://www.asq.org. Free material is available through its Statistics, http://www.asq.org/statistics/l and Reliability (http://www.asq.org/reliability/) Divisions.

Chance Magazine. http://www.amstat.org/publications/chance/

Fulbright Speakers Specialist Program: http://www.iie.org/cies/Specialists

International Statistical Institute. Main page: http://www.cbs.nl/isi/ Also check its Business and Industry Section and the Intl. Assoc. for Statistical Education pages.

Instituto Inter Americano de Estadisticas. http://www.indec.mecon.ar/iasi/iasi/publicaciones.htm

Juarez Lincoln Marti Int'l. Education Project. http://web.cortland.edu/matresearch

Math World: http://mathworld.colfram.com/topics/ProbabilityandStatistics.html

Minitab resources. http://www.minitab.com/resources/articles/StatisticsArticles.aspx

Lejarza, Juan. Libro de Texto en Castellano http://www.uv.es/~lejarza/hip3/indexh.html

Leon, Ramon V. http://web.utk.edu/~leon/talks/default.html

National Institute of Standards and Technology. http://www.itl.nist.gov/div898/handbook/

Statistical Education Research Journal of the International Association for Statistical Education. http://www.stat.auckland.ac.nz/~iase/publications.php

Syracuse University DOE Industrial Assessment Center. http://iac.syr.edu/carranti.htm

Royal Statistical Society. Journals for members. http://www.jstor.org/journals/09528385.html Also visit the Center for Statistical Education and Business and Industry Section.