Statistics Education of Practicing Engineers

Jorge Luis Romeu, Ph.D.
Research Professor, Syracuse U.
http://myprofile.cos.com/romeu
jlromeu@syr.edu

Conference in Honor of Dr. Dudewicz
July 2008

Outline

- Statement of the Problem
- Background & Engineering Survey
- Research and Survey Results
- Professional Courses/Certifications
- Mentoring and other methods
- Proposed Solution/Consequences
- Academe-Industry Institutes
- Discussion and Conclusions

Statistical Education Research:

- For Primary school curriculum
- Secondary school curriculum
- Undergraduate curriculum
- Various graduate curriculums
- Service courses curriculums
- Statistics in sciences/humanities/Bio
- But NO studies on post-grad Stats Ed!

Statistics in Engineering:

- Many engineering measures of interest:
  - Are random in nature: e.g. life, quality, stress
- Engineers measure, compare, test them:
  - reliability/maintainability studies
  - performance evaluation studies
  - optimization and simulation studies
  - continuous improvement studies,
  - SPC, DOE, Six Sigma, etc.

But Engineering Statistics

- Insufficiently taught in undergraduate
  - Very few courses (mainly one or two)
  - Or none at all (no courses offered)
- Engineers have to study it on their own
  - As best they can, using different means
  - As practicing professionals (after college)
- Finding they need/use it in their work
  - And have not studied enough of it!

Same syndrome occurs in other Engineering Careers

- Example: Corrosion studies
- Recent National Workshop held
- At the US Academy of Sciences
- Corrosion not taught sufficiently
- Very few (if any) courses
- Billions in losses, because
- Corrosion not considered in Design

Examples of Undergraduate Engineering Statistics:

- Mechanical Engineering
  - One course, at Math Dept.
- Civil Engineering
  - One course, at Math Dept.
- Electrical Engineering
  - One course, at Math Dept.
- Computer Science
  - One course, taught internally

Undergraduate Statistics Curriculum

- Descriptive Stats (Chs. 1& 2): Examples of uses of statistics in problem solving Frequency distributions, Pareto, Dot, Stem-and-leaf and other diagrams and graphs; descriptive measures and their calculations. Case study.
- Probability (Ch. 3): sample spaces, events, counting rules, axioms of probabilities, elementary theorems, conditional probability, Bayes theorem, mathematical expectation. Case study.
- Distributions (Ch. 4): random variables, discrete distributions: Uniform, Binomial, Hypergeometric, Geometric, Multinomial, Poisson. Approximations. Chebyschev' theorem. Applications.
- Densities (Ch. 5): continuous random variables and distributions: Normal and its approximation to the Binomial, Uniform, Exponential, Log-Normal, Gamma, Weibull. Joint distributions. Checking for Normality. Variable Transformations.
- Sampling Distributions (Ch. 6): populations and samples, distributions of the mean and the variance; Student t, F and Chi Square distributions.
- Inferences Concerning the Mean (Ch. 7): point and interval estimation.
- Some applications in Reliability engineering (Ch. 15). Text: Johnson's.

Graduate Statistics Curriculum

- Review of Probability, Random Variables, Probability Distributions (Ch. 3 to 6): Discrete distributions; Uniform, Binomial, Multinomial Hypergeometric and Poisson. Continuous distributions: Normal, Exponential, Gamma, Weibull, Approximations.
- Sampling Distributions (Ch. 8): populations and samples; parameters and statistics; sampling distributions (t, F, Chi-Square).
- Point and Interval Estimation (Ch. 9, 16): estimation of mean, proportion and variance of a single sample; paired samples; difference between two means/proportion; ratio of two variances. Quality Control. Applications.
- Hypothesis Testing (Ch. 10): theoretical development and framework, tests for the mean, proportion and variance of a single population; tests for two means, two proportions and two variances. One and two sided tests. Goodness of Fit; Sample size.
- Correlation and Linear Regression (Ch. 11 and 12): simple linear regression, including model verification, residual analysis, multiple regression, selection of variables, choosing the best model. Lack of Fit. Variable transformations.
- Analysis (ANOVA) of Variance (Ch. 13): one-way and two way ANOVAs, randomized experiments; random blocks; Factorial designs; model verification and residual analysis.
- Other topics (Design of Experiments; Non Parametrics) as time allows (Walpole/Myers)

Quality Engineering: Weekly Topics

Intro; Basic Concepts; Gurus; Company-wide Q; COPQ

Quality improvements: Roadmap (Juran); Intro Six Sigma

Detailed Six Sigma (DMAIC); Old Tools; Process Capability.

Design for Quality (DFSS); New Tools: QDF & other Matrices

Advanced Statistical Analysis: Intro to DOE; Case Studies

First Midterm; Test review and Course analysis

Supplier Relations: Supply Chain Mgmt., Lean Manuf.

Inspection, tests and measurements: Gage R&R

Winter Break; no classes

Acceptance Sampling; OC function and applications

Statistical Process (SPC) Control: Theory and Organization

Practical applications of Control Charts; Intro Reliability

Reliability & maintainability: FMEAs & Fault Trees

Second Midterm Test; ISO/Baldrige

Strategic Quality Management; Audits; Assurance.

Certification Statistics B.O.K.

American Society for Quality

- Certified Quality Engineer
  - statistical content of the exam (50%+)
  - http://www.asq.org/certification/quality-engineer/bok.html
- Certified Reliability Engineer
  - statistical content of the exam (40%+)
  - http://www.asq.org/certification/reliability-engineer/bok.html

How can engineers bridge the Gap between both levels?

This is the Main Topic of this Talk!

Move Statistical Content:

a) Intro: Up-stream (grade school)

b) College: inference/modeling

c) Lifelong Learning: special topics

The survey on how engineers learn Statistics on their own:

http://lcs.syr.edu/faculty/romeu/SurveyICOTS.html

Provides some direction regarding the <u>Means</u> used by engineers in learning statistics.

We can then expand and improve: On Such Means and Topic Areas.

Survey Methods of Self-Learning

(1) reading books, journals, manuals or other hard copy,

(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 colleagues

(8) other sources: e.g. hands-on (practical) working experiences, and taking Six Sigma training

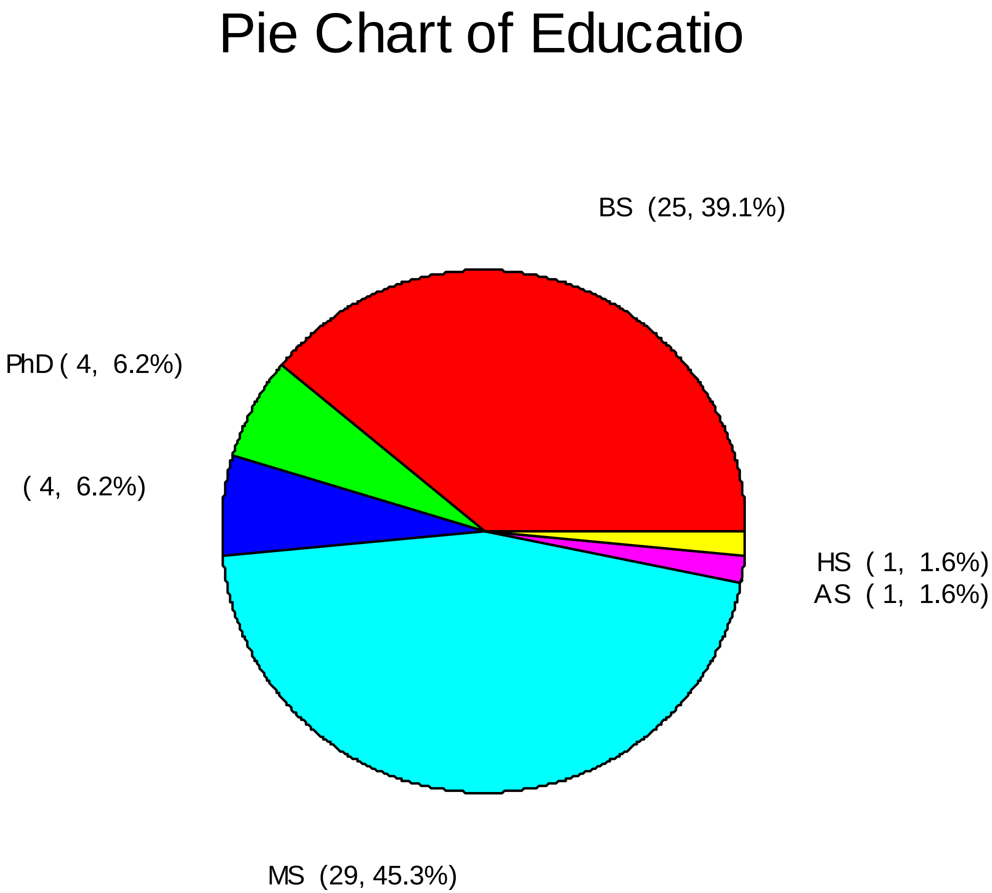
ICOTS Survey on Practicing Engineers Statistical Education (%).

HardR	WebR	Tutor	ProfMg	Certif	ShortC	Mentor	Other	Educ	Area	Spclz	Cours	Years
15	10	5	10	10	30	20	0	BS	I	ME	2	10
35	35	0	0	0	10	20	0	BS	I	SW	2	1
70	10	10	0	0	5	5	0	PhD	I	Chm	3	7
35	15	0	0	0	25	25	0		I	Chm		
5	10	5	5	20	10	15	30		I	Chm		
5	0	0	2.5	10	2.5	0	80	BS	I	IE	1	13
25	0	0	10	15	0	10	40	MS	I	IE	2	11
10	0	0	0	75	15	0	0	BS	I	EE	0	25

Survey Sample Description

- 64 responses received (and counting)
- 61 from the US
- 3 from abroad.
- 8% were females
- 56% had graduate degrees
- 60% had 16+ years of experience
- 90% were from industry.

Educational Levels



Survey Results Regarding College Statistical Training

I) Among all surveyed, 16% have not taken any statistics courses in college (33% among BS), 38% took only one (38%) and 26% have taken 2 courses (24%).

II) 1/3 of those with a BS degree only, have never taken a single statistics course in college; another 1/3 of them have taken only one course. Hence, 2/3 engineers of all surveyed had either none, or very little statistical training (i.e. taken a single course).

III) Engineers that pursue graduate school have a larger opportunity of taking statistics. Only 7%, in our sample, have never taken a statistics course.

Survey Results on Methods (I)

Methods of Acquiring Stats Knowledge after College

Method	HardCopy	Internet	Software	Chapt. Mtgs.
Prop.	0.272	0.099	0.02	0.067

Note: for details of this Survey implementation and its results see our paper in the 2006 ASA Joint Research Conference, Knoxville, Tennessee (June of 2006).

Survey Results o Methods (II)

Methods of Acquiring Stats. Knowledge after College

Method	Certif.	Courses	Mentor.	HandsOn
Prop.	0.199	0.127	0.15	0.066

NOTE: for details of the statistical analysis of significant correlations between survey variables and implications, see our paper in the 2006 JSM/ASA, Seattle (August).

Methods Preferred

I) ***"Readings" constitute the preferred means of learning: books and journals, as well as web tutorials, provide 38% of statistics knowledge. The use of web tutorials (10%) is increasing with time: older engineers prefer hard copy, whereas younger ones read web-based material.***

II) Short courses, exam preparations for the professional certifications, and Black Belt training, are also important methods of learning statistics (33%).

III) mentoring received from more experienced colleagues and hands-on (learning by doing), also constitute frequent learning activities (22%).

Hard Copy and Web Readings

- Most popular method (almost 40%)
- Younger prefer Web; older, hard copy
- Web is faster, more economic medium
- Web tutorials: dispersed, unclassified
- Most material in English -third world?
- Access becoming complex (Browsers)
- Best option for the future, though.

Some Problems with Readings

- Web material, more and more relevant
  - need for creating a data base of existing ones
  - need for cataloguing and assessing it
  - need for sequencing such reading material
  - need for developing more/filling “holes”
- Hard Copy material, also widely used
  - Also needs to be sequenced and assessed
  - Catalogued to find weak areas and fill them.

Professional Courses

- About 20% used these as means in learning
- Intensive, short, to the point, practical
- Single topic, no inter-relationships, uneven
- Student body also very heterogeneous
- Background and assumptions often missing
  - or checking them is poorly stressed
- Some end up teaching SW and formulas
  - that are then questionably applied

Professional Certifications

- About 20% of knowledge, learnt this way
- Professional Societies: ASQ, SME, IEEE
  - offer several statistically-based certifications
  - Quality (Control); Reliability; Logistics, etc.
- Self-study materials (manuals, questions)
- On-line and classroom courses
- Require periodic updating/follow-ups

Mentoring/Hands on

- Learning from older colleagues/engineers
- Learning from relatives (spouses, etc.)
- Trial and error/Hands on experiences
- Very few use on-line tutorials
- Attending professional conferences
- Chapter meeting presentations
- Other methods ... (wife, brothers, etc.)

Methods that substitute prof. practice and experience:

Multivariate Regression Analysis (Percents):

The regression equation is:

YearsPract = 26.7 - 0.407 WebRead - 0.159 Certificats - 0.141 OtherWays

55 cases used 9 cases contain missing values

Predictor	Coef	StDev	T	P
Constant	26.676	2.327	11.46	0.000
WebRead	-0.4070	0.1023	-3.98	0.000
Certific	-0.15919	0.07110	-2.24	0.030
OtherWay	-0.14147	0.06306	-2.24	0.029

S = 9.524 R-Sq = 26.5% R-Sq(adj) = 22.2%

ICOTS Survey on Practicing Engineers Statistical Education (Ranks)

HardR	WebR	Tutor	ProfMg	Certif	ShortC	Mentor	Other	Educ	Area	Spclz	Cours	Years
1.5	5	3	5	5	8	7	1.5	BS	I	ME	2	10
7.5	7.5	2.5	2.5	2.5	5	6	2.5	BS	I	SW	2	1
8	6.5	6.5	2	2	4.5	4.5	2	PhD	I	Chm	3	7
8	5	2.5	2.5	2.5	6.5	6.5	2.5		I	Chm		
2	4.5	2	2	7	4.5	6	8		I	Chm		
6	2	2	4.5	7	4.5	2	8	BS	I	IE	1	13
7	2	2	4.5	6	2	4.5	8	MS	I	IE	2	11
6	3	3	3	8	7	3	3	BS	I	EE	0	25
7.5	7.5	3	3	3	3	6	3	PhD	I	Chm		
4	4	1.5	4	7	8	6	1.5	BS	I	Matr	0	21
5.5	5.5	2.5	2.5	2.5	2.5	7	8	PhD	I	Matr	0	10

Note: the highest the rank, the most preferred method.

Sign confidence interval for median of rank data (1):

	N	N*	Median	Confidence	Confidence Interval	Pos
HrdRd	64	0	6.000	0.9392	(5.500, 7.000)	25
				0.9500	(5.500, 7.000)	NLI
				0.9664	(5.500, 7.000)	24
WbRd	64	0	4.000	0.9392	(3.500, 4.500)	25
				0.9500	(3.500, 4.641)	NLI
				0.9664	(3.500, 5.000)	24
OnLnTut	64	0	2.500	0.9392	(2.500, 3.000)	25
				0.9500	(2.500, 3.000)	NLI
				0.9664	(2.500, 3.000)	24
PrfMtgs	64	0	3.500	0.9392	(3.000, 4.500)	25
				0.9500	(3.000, 4.500)	NLI
				0.9664	(3.000, 4.500)	24

Note: Ranks are 1 through 8, the highest being the most preferred.

Sign confidence interval for median of rank data (2):

	N	N*	Median	Confidence	Confidence Interval	Pos
Certif	64	0	4.000	0.9392	(3.000, 5.500)	25
				0.9500	(3.000, 5.500)	NLI
				0.9664	(3.000, 5.500)	24
ShrtCrs	64	0	4.750	0.9392	(3.500, 5.500)	25
				0.9500	(3.500, 5.641)	NLI
				0.9664	(3.500, 6.000)	24
Mentor	64	0	6.000	0.9392	(4.500, 6.000)	25
				0.9500	(4.500, 6.000)	NLI
				0.9664	(4.500, 6.000)	24
OthrWys	64	0	3.500	0.9392	(3.000, 4.500)	25
				0.9500	(2.859, 4.641)	NLI
				0.9664	(2.500, 5.000)	24

Note: Ranks are 1 through 8, the highest being the most preferred.

Preferred post-college education method depends on Level.

Non-Parametric (Mood) Comparing Education (Ranks)

Response	p-value	Lowest	Highest
Stats Courses	0.028	PhD	MS
Hard Read	0.080	BS	PhD
Web Read	0.052	MS	PhD
Certifications	0.043	PhD	BS
Short Courses	0.114	PhD	MS

Hence, the main problem:

Most Engineers

- Don’t study enough statistics in college
- No room to add more stats to curriculum
- No way to know what topics they will use
- Topics will differ, on case by case basis
- But they Need to lean stats for their work!

Proposed Solution:

- Teach statistics in more than one level
- Move it up and down the learning stream
- Upstream: intro stats from grade schools
- Downstream: advanced stats as “life-long”
  - By Creating Academe-Industry Institutes
- Then, leave college for *formative* stats:
  - Statistical inference (testing and C.I.)
  - Statistical modeling (regression/ANOVA)

Academe-Industry Institutes

– local partnerships w/industry & prof. societies
– internships in industry for college students
– teach practical short courses and seminars
– (re)certification training/courses, consulting
– “free” assessments to small/medium size orgs.
– mentoring and developing new web materials
– work with secondary Ed. Technology teachers
– help create a “community of statistics users”

Institutes Solve Three Problems

- Past Situation
  - Practicing engineers w/o statistics training
  - Obtain it in workshops, evening courses
- Present Situation
  - Current engineering students learn statistics
  - Also hands-on experience via internships
- Future Situation
  - High School students exposed to statistics
  - High School teachers also receive training

Some of their Functions

- Serve small/medium size industry
  - Providing free/inexpensive assessments
- Train practicing engineers in Q&R
  - And other Industrial Statistics Methods
- Enhance undergraduate education
  - By using college interns in their functions
- Prepare H.S. students for engineering
  - By incorporating teachers and students

Stake Holders/Benefits

- Industry/Services
  - Competitiveness, profits, survival
- Academe/University
  - Better teaching and research
- All Government Levels
  - Tax base increase, economic growth
- The Public at Large
  - More Jobs, better services & quality of life

Possible Income Sources

- Federal Government grants
  - NSF: educational function (engineering)
  - Other agencies sponsoring job development
  - Such as Economic Development Agency
- State and Local Government grants
  - Help local industry remain competitive
  - Save local jobs; revert regional emigration
- Local Industry and Academe support
  - Office space, phone, computers, students.

Institute Board of Advisors

- Integrated by all Stake holders
  - State and Local Government
  - Institute industry customers
  - Academic faculty/administration
  - Interns and Students
  - Industry Donors ($$)
- Help define directions to pursue
  - Focusing on problem-solving activities
- Help Find new Customers and Services

Institute Networking

- With Other Industry-Academe Centers
  - Of different type, in the region
  - Of the same type, in the nation
  - To enlarge and refine activities
  - To conduct joint activities
  - To exchange students and faculty
  - To teach joint Q&R courses
  - Other mutually beneficial activities

Discussion

- Two objective and unquestionable facts:
  - Many/most engineers need stats in their work
  - But don’t learn it in school: no room to teach it!
- Need creative ways to provide this learning
  - moving stats forward/backward in curriculum
  - teach elements of stats, from grade school
  - statistics core AND THINKING, in college
  - on-going, life long learning after college work

Conclusions

- Statistics education of practicing engineers
  - a serious issue with important repercussions
- Government, Industry, Academe
  - must all contribute to address such issue
- Professional societies also stake holders
  - And have a leading role to play
- Meta Model is quite general
  - Statistics is just one instance (as Corrosion)
- That has to be faced by society at large.