Classification and Evaluation of Examples for Teaching Probability to Electrical Engineering Students

George Nagy and Biplab Sikdar

ABSTRACT

Although teachers and authors of textbooks make extensive use of examples, little has been published on assessing and classifying pedagogic examples in engineering and science. This study reviews various characteristics of examples intended for a course on probability for electrical engineers. Twelve examples are constructed to illustrate some characteristics of the correlation coefficient. A survey incorporating these examples was administered to professors and students at Rensselaer who have taught or taken a course in probability. Statistical tests are applied to determine which examples professors and students prefer, and to what extent they agree in their preferences. New bipolar criteria are proposed to classify objectively a broader set of examples that appear in textbooks. Even though preferences depend on educational background and maturity, textbooks on Probability are sharply differentiated by the proposed classification criteria.

Index terms: education, analogies, learning, probability, statistics.

I. INTRODUCTION

Two aspects of examples used in teaching Probability to electrical engineering undergraduates at Rensselaer Polytechnic Institute are investigated here: (1) Do professors and students have consistent and compatible opinions about the relative value of a dozen examples focused on the characteristics of the correlation coefficient? (2) Can examples in eight popular textbooks on probability (of which all but one include *engineering applications* or *random processes* or *communications* in their title or subtitle) be classified objectively? Methods developed to answer these questions are likely to be useful, if not necessary, for any deeper study of the *quality* of pedagogic examples.

Students *demand* examples, both in the classroom and in their textbooks. About one half of the contents of the twenty or so introductory books on probability that were examined are devoted to clearly demarcated and numbered examples that range in length from two lines to two pages. Since all of the theory of probability follows from a few axioms, any consequence of the axioms can be presented either as a theorem or as an example. For instance, Example 3.2-1 in Stark and Woods [1] and Example 3.23 in Leon-Garcia [2] give exactly the same result for the probability density function and cumulative distribution function of a linear transformation of a random variable as do Theorems 3.19 and 3.20 in Yates and Goodman [3].

The rest of the paper is organized as follows. Section II reviews some hypotheses proposed in the literature about the nature of good examples. Section III describes the survey constructed to investigate preferences of professors and students, and a statistical analysis of the results of the survey. Section IV contains an analysis of textbook examples and Section V presents the concluding remarks.

II. WHAT IS A GOOD EXAMPLE?

Examples and exemplification have received much more attention in the teaching of languages [4] than in engineering. Yelon and Massa [5] say that "good examples are accurate, clear, attractive and transferable." Sweet [6] suggests the following characteristics for teaching grammar:

1. *They illustrate or confirm the rule clearly. They are unambiguous.*
2. *They are understandable without more context.*
3. *They are as concrete as possible, the more concrete the better – especially in giving words and vocabulary for beginners.*
4. *They do not contain difficult or rare vocabulary or irregular forms that are not involved in the particular rule being illustrated.*

The first two points are directly applicable in the context of teaching probability while the third point suggests using numerical rather than symbolic expressions. The last point would proscribe mathematics irrelevant to the concept being illustrated, such as cumbersome arithmetic, difficult differentiation or integration, or completing the square.

The desirability of not letting complex mathematical manipulation obscure engineering or physical intuition is emphasized by Faria, who develops a series of examples to illustrate high frequency analysis of conductance-grounding effects [7]. All of his examples are deliberately based on simple two-conductor transmission line theory. An earlier article that makes the same point demonstrates a progression of three examples on the application of the method of moments (a numerical procedure for solving a linear operator equation by transforming it into a system of simultaneous algebraic equations) to electromagnetic problems [8]. In probability, many concepts can be illustrated simply and intuitively, for instance, with uniform probability densities on the real line or square.

In contrast, Blair, Conte and Rice argue that the derivations associated with analytic expressions are instructive [9]. To illustrate homomorphic signal processing, the authors show mathematical "tricks" for deriving the equations that characterize the distortion of signals in the frequency domain and for approximating the signal in the time domain to arbitrary accuracy. The approximation process is then illustrated for two input signals. Approximations based on transformations of variables (i.e., moment generating functions and characteristic functions) are also important in probability. Section IV presents some statistics about the occurrence of both interrelated examples and advanced mathematical techniques in textbooks on probability.

Insights from software engineering may be applicable to the *coverage* of examples. Cordy recommends input partition testing, where program inputs are divided into equivalence classes that correspond to every possible path through the program [10]. The notion of coverage is also fundamental in VLSI fault testing [11]. Educational software for generating problems and solutions from a set of templates has been available for at least four decades, but problems that test *mastery* of the material have different characteristics from examples for *instruction*. In contrast to test generation, the accent in teaching is on illustrating correct solutions rather than on comprehensive coverage.

Other queries about the quality of an example:

1. Is topicality important? Are examples based on contemporary affairs (e.g., elections) preferable to examples related to prerequisite or co-requisite engineering courses? Most early studies of probability were inspired by games of chance, which remain popular in textbook examples.
2. When is an example *misleading?* Will an example of a probability mass function with a range consisting of three possible values suggest to students that all mass functions are defined on exactly *three* values? Only one of the examined textbooks gives deliberate examples of common student mistakes [12].
3. Should an example illustrate only a single procedure or concept, or several? Should examples be interdependent? Should examples of probabilistic notions be based on mathematical abstractions or on concrete phenomena? What accounts for the sustained popularity of Feller's [13] and Papoulis's [14] examples?

III. STUDENT AND PROFESSOR EVALUATIONS OF EXAMPLES

As the above questions suggest, the notion of what constitutes a good example may be subjective. Some may consider illustrating fundamental concepts to be more important than applications of probability to real problems. Personal differences in the way different people learn can also affect their judgment of the pedagogic effectiveness. Thus it is of interest to learn if indeed there are certain qualities of examples that are universally admired or deprecated. Therefore a survey was constructed to discover any consensus, within and among groups with diverse educational backgrounds, on a set of twelve examples (E1 to E12) that follow the most popular prototype (generic, numerical) encountered in text books.

All examples in our survey illustrate the same procedure for determining the correlation coefficient and the statistical dependence between a pair of discrete random variables with given joint probability mass functions. Different values of the marginal and joint probability mass functions produce contrasting characteristics with respect to criteria mentioned in the literature, like simplicity and coverage. The examples, including the results of the calculations, were displayed in a six-page survey with a short introductory note (the survey instrument is available from the authors).

Responses to the survey were collected from 46 subjects consisting of three groups: (1) 23 undergraduate students majoring in electrical or computer engineering at RPI who were in the last week of a senior level course on probability, (2) 10 graduate teaching assistants for the same course, and (3) 13 professors of electrical and computer engineering (6 from outside RPI) who regularly use probability in their research or have taught an undergraduate probability course. Respondents received no incentive, and came under no pressure, to complete the survey. Less than half of the undergraduate students in the two classes (one year apart) to whom the survey was distributed returned it. Each respondent answered the following questions.

Q1: From the 12 examples, pick the *four* you like the best (we call these A), the four you like second best (B) and the four you like the least (C) (without ordering the examples within each of the three groups).

Q2: If you were to select *two* examples from these 12 to show in a class, which would they be?

Q3: If you were to pick *three* examples to show in the class, which would they be?

The first question ascertains preferences with regard to individual examples. The second and third questions should indicate whether examples that may not be ranked as the best in the set of twelve may still be picked when pairs or triplets of examples are to be shown to a class. Pairs or triplets may illustrate different concepts (such as positive and negative correlation, or uncorrelated but dependent variables), even though these examples may not be considered best when viewed in isolation.

A. Results from the Survey

Table I shows how many respondents assigned each example to each category. The table also shows the overall rank of the examples, calculated by assigning a weighted score $s(i)$, $1 \leq i \leq 12$, as in [15],[16] to each example:

$$s(i) = 1 \cdot n_A^i + 0 \cdot n_B^i + (-1) \cdot n_C^i = n_A^i - n_C^i \qquad (1)$$

where n_A^i, n_B^i and n_C^i are the number of responders of a group classifying Example i in category A, B and C, respectively. The choice of –1, 0 and 1 of the weights was made in accordance to the Likert scale commonly used in survey research [17]. The twelve examples were ranked according to these scores, with the example with the highest score considered the best. Ties in the scores were broken randomly. The order of the examples in the surveys given to different responders was not randomized. While this may lead to a bias in the responses, the results do not show any conclusive evidence of ordering effects.

Table I. Group responses, scores and ranks for Question 1.

Ex.	Professor					Graduate					Undergraduate				
	A	B	C	Score	Rank	A	B	C	Score	Rank	A	B	C	Score	Rank
1	5	5	3	2	2	6	4	0	6	2	14	9	0	14	1
2	6	2	5	1	4	7	1	2	5	3	16	1	6	10	2
3	4	3	6	-2	9	4	2	4	0	7	6	9	8	-2	8
4	4	4	5	-1	7	2	4	4	-2	8	4	10	9	-5	10
5	3	5	5	-2	8	1	1	8	-7	12	4	5	14	-10	12
6	3	3	7	-4	11	4	4	2	2	5	7	8	8	-1	6
7	3	6	4	-1	6	1	3	6	-5	9	5	11	7	-2	7
8	4	6	3	1	5	0	4	6	-6	11	8	8	7	1	5
9	11	2	0	11	1	5	4	1	4	4	9	7	7	2	4
10	4	7	2	2	3	7	3	0	7	1	10	10	3	7	3
11	3	3	7	-4	12	0	5	5	-5	10	3	7	13	-10	11
12	2	6	5	-3	10	3	5	2	1	6	6	7	10	-4	9

The two key questions to be answered are whether groups with different levels of exposure to a topic evaluate an example consistently, and whether there are certain characteristics that define a good example. The following statistical hypothesis tests were carried out to answer the first question.

1. *Test 1. Are the ratings of an example consistent among the groups?*

To answer this question, the likelihood that the observations from two groups are generated from the same population was computed. The classification of an example as an A, B or C in response to Question 1 by members of *two* groups forms one set of samples of observations. The medians of the ratings of the two groups are compared with the non-parametric Mann-Whitney test [18]. The null hypothesis is that the two samples are drawn from the same population and therefore the medians are equal. The test is repeated for each of the 12 examples in the survey.

Result of Test 1. Table 2 reports the observed significance levels between the ratings of the professors against the two student groups. The professors and graduate students differ significantly in their opinion (with a p value of less than *0.10*) in four of the examples while the professors and undergraduates differ significantly on two of the examples. The professors disagree with both the student groups on E9, but only with the graduate students on E8 (which suggests that the undergraduates and graduates disagree on E8).

Table 2. Mann-Whitney significance values for consistency among groups.

Example	Professor-Graduate	Professor-Undergraduate
	p	p
1	0.1733	*0.0824*
2	0.2644	0.2288
3	0.6894	0.7792
4	0.7412	0.6466
5	*0.0711*	0.2609
6	0.1471	0.3427
7	0.1690	0.9716
8	*0.0318*	0.9163
9	*0.0699*	*0.0059*
10	*0.0494*	0.5046
11	0.7313	0.7132
12	0.2828	0.9017

2. *Test 2. Are examples consistently classified as good or bad?*

The number of times Example i is classified as A, B and C, n_A^i, n_B^i and n_C^i, respectively, represents a multinomial experiment with three possible outcomes. How a group assesses an example can be judged by the number of times it is classified in category A as opposed to C. Classification into category B, does not add any further information on whether it is good or bad. Thus to test whether the respondents assess the example favorably is equivalent to testing whether they select bin A over bin C more often than might happen by chance. The estimator for the category probabilities of Example *i* is $\hat{p}_x^i = n_x / n$, where $n = n_A^i + n_B^i + n_C^i$ and x is one of A, B or C. The null hypothesis $\hat{p}_A^i - \hat{p}_C^i = 0$ is tested against the alternative hypothesis $\hat{p}_A^i - \hat{p}_C^i > 0$.

Result of Test 2. As shown by the observed significance levels in Table 3, for the professors the null hypothesis can be confidently rejected only for E9, suggesting that they consider this example a good one. Graduate students like E1, E2, E9 and E10, while undergraduate students like E1, E2 and E10. The results indicate that there is less of a consensus among the professors than within and between the student groups. Also, while all three groups think alike on E11, there is considerable disagreement on E5 and E8, among others.

Table 3. Significance levels for within-group agreement on individual examples.

Example	Professor	Graduate	Undergraduate
1	0.2354	*0.0001*	*0.0000*
2	0.3811	*0.0249*	*0.0087*
3	0.7396	0.5000	0.7047
4	0.6311	0.8009	0.9263
5	0.7646	0.9997	0.9966
6	0.9116	0.1990	0.6020
7	0.6480	0.9908	0.7196
8	0.3520	0.9999	0.3980
9	*0.0000*	*0.0283*	0.3076
10	0.2009	*0.0000*	*0.0169*
11	0.9116	0.9992	0.9983
12	0.8838	0.3257	0.8467

3. *Test 3. Are questions labeled the best by a responder (Q1 of the survey) also likely to be picked as the two examples to show in class (Q2)?*

To answer this question, a two-way contingency table was generated where the variables are (i) how many times an example is selected in Category A of Q1 and (ii) how many times it is selected in Q2. Fisher's exact test [19] was applied to the null hypothesis that these two variables are independent against the alternative hypothesis that they are not.

Result of Test 3. Table 4 shows the observed significance levels for the three groups: in all cases the null hypothesis can be confidently rejected, showing that there is strong correlation between the examples ranked in Category A of Q1 and those included in Q2. Table 4 also shows the degree of correlation as given by Kendall's or Spearman's correlation coefficient and its asymptotic standard error (ASE).

Table 4. Significance levels for selections in Q1 and Q2.

Group	*p*	Kendall/Spearman coefficient	
		Value	ASE
Professor	2.243E-10	0.5230	0.0679
Graduate	3.374E-06	0.4427	0.0854
Undergraduate	2.776E-15	0.4881	0.0536
Combined	4.694E-28	0.4881	0.0379

B. Characteristics of Examples

The other important question that the survey was designed to answer was what are the characteristics of a good or a bad example. One indication of the quality of an example is its overall ranking. Additional indications are given by comments on the example provided by the responders. The characteristics to be evaluated are listed in Table 5 and described below.

Table 5. Examples with specific characteristics

	Characteristic	Examples
1	Simple numbers	E1, E2, E6, E9, E10, E 12
2	Illustration of key concept	E9, E10
3	Oversimplification	E2, E4
4	Extreme conditions	E5, E6, E12
5	Same X and Y values	E1, E2, E6, E9, E10, E12

1. *Simple numbers:* Examples with simpler numbers are preferred by many respondents, cutting across professors, graduate and undergraduate student groups. Here, "simpler numbers" refers to the use of few digits, integers and fractions that are easy to manipulate. Of the six examples with simple numbers, four (E1, E2, E9 and E10) were ranked as the top four by each group. On the other hand, among the six examples with complex numbers, four were ranked in the bottom half by all three groups. Five undergraduates explicitly indicated their preference for simple numbers in their comments.
2. *Illustration of a key concept*: A key concept that was to be illustrated through these examples was that of correlation and dependence between random variables (as explicitly mentioned in the survey). While a student can easily understand that correlated random variables are dependent and independent ones are uncorrelated, the same student may find the fact that independence is not implied by absence of correlation to be far from obvious. E9 and E10 were the two examples that showed dependent but uncorrelated random variables. Among the professors, E9

was by far the most popular choice with a score of 11, while E10, with a score of 2, was tied for second place. Graduate students ranked these two examples in positions 4 and 1 respectively, with scores of 7 and 4. While undergraduates also ranked E9 and E10 in the top four with ranks of 4 and 3 respectively, their scores were only 2 and 7. In contrast, undergraduates had scores of 14 and 10 for their top two examples.

3. *Oversimplified examples*: Two professors and two undergraduates commented that they consider examples with all zero rows or columns as bad. Among the two such examples in the survey, one was ranked 7, 8 and 10 by the professors, graduates and undergraduates, but the other was ranked 4, 3 and 2. Thus other qualities may compensate for having all zero rows or columns.
4. *Examples showing extreme conditions*: Three of the examples in the survey had a correlation coefficient of 1 and all the non-diagonal elements in the matrix for their probability mass function were zero. Two of these three examples were ranked in the bottom half by all groups while the third example had ranks of 11, 5 and 6. This suggests that examples showing perfect correlation are not highly valued by any of the groups.
5. *Use of symbols instead of numbers:* Our survey used numerical values in all the examples. Two of the graduate students pointed out that it might be better to use symbols rather than numbers, because it would shift the focus entirely on the equations and eliminate any effects that the choice of numbers may have on the results. However, numbers would still need to be substituted to show directly that a correlation coefficient of zero does not mean independence.

6. *Individual preferences:* Finally, individual preferences may also affect how examples are viewed and ranked. For example, one undergraduate noted that Example 6 was the best while its overall ranking by the undergraduate population is only 6. Also some respondents (one professor, one graduate and two undergraduate students) mentioned that they found it difficult to differentiate between good and bad examples.

C. *Pair Selection*

In Q2 of the survey, the respondents were asked to pick two examples that would be shown to the class in order to illustrate the concepts of dependence and correlation between pairs of random variables. Table 6 shows how many times various types of pairs of examples were chosen by the professors, graduate and undergraduate students. A respondent may select 2 examples out of 12 in 66 different ways, but the respondents opted for only a very small subset of this sample space. Only 3 of the 13 professors (23% of the population) failed to pick any of the two examples showing a dependent but uncorrelated pair of random variables (E9 and E10). Among the graduate students, 3 of the 10 respondents (30%) picked neither E9 nor E10. However, this number increases to 14 out of 23 among the undergraduate students (61%). Also, while only one professor and one undergraduate picked a pair with *independent and therefore uncorrelated*, and *dependent yet uncorrelated* random variables, 4 of the 10 graduate students selected such a pair.

Table 6. Selected types of paired examples.

Group	Pair type			
	At least one showing dependent but uncorrelated	One independent, uncorrelated and other dependent, uncorrelated	One correlated and other uncorrelated	One positively correlated and other negatively correlated
Professor	10	1	9	1
Graduate	7	4	5	1
Undergraduate	9	1	17	1

A common trend among all groups was selecting pairs composed of one example showing correlated, and another showing uncorrelated, random variables. In the survey, 9 of the 13 professors, 5 of the 10 graduate and 17 of the 23 undergraduate students selected such pairs. Only one respondent each from the three groups picked a pair showing both negatively and positively correlated random variables. Table 6 and comments provided by the respondents suggest that the following factors affect the choice of examples that different groups like to see in a class:

1. *Examples illustrating contrasting situations:* There was a marked respondent preference for selecting contrasting pairs, such as (i) one example showing *correlated* and the other showing *uncorrelated* random variables, (ii) one showing *dependent* and the other showing *independent* random variables, and (iii) one showing *uncorrelated and independent* and the other showing *dependent but uncorrelated* random variables.
2. *Illustration of a key concept:* Examples that illustrate a key concept are favored more by the professors and graduate students as compared to the undergraduate students. Examples of key concepts consisted of *dependent but uncorrelated* random variables.

3. *Examples that demonstrate the significance of a parameter:* Some respondents prefer a set of examples that can illustrate the effect of a parameter on the conclusion drawn from the example. For correlation and independence, a pair of examples can show the effect of increasing the variance of one of the variables on the correlation coefficient (while keeping its mean constant).
4. *More than two examples:* In some cases, more than two examples may be necessary to clarify all aspects of the problem. Two professors stated that a triplet would be better than a pair because it could include one example showing *dependent and correlated* random variables, one showing *dependent but uncorrelated* random variables, and a third example showing *independent and uncorrelated* random variables.

The results for the answers to Question 3 of the survey, where the respondents were asked to pick three examples to show to the class, were similar.

Consensus in rating examples. One of the surprising observations from the survey was that students have more unified opinions than the professors. An indication of the degree of agreement on the relative value of an example is provided by its score as given by Equation (1), and reported in Table 1. Table 1 shows that only one example (E9) had a high score from the professors, and three had scores lower than –3. The remaining 8 examples have scores between –2 and 2, suggesting that there is no consensus among professors on whether they are good or bad, or even that they are neither. In contrast, the graduate students give four examples scores higher than 4, and scores lower than –5 to another four. This trend is also reflected in the scores of the undergraduates who give three examples large positive scores, four examples large negative scores (<-4) and the rest scores between –2 and 2. Significantly, while undergraduates ranked examples with simpler numbers as the best, both professors and graduate students ranked an example which shows a key concept (correlation coefficient of zero does not mean independence) as the best example.

IV. CLASSIFICATION OF TEXTBOOK EXAMPLES

This section presents the eight binary criteria that were used to classify several hundred examples from popular textbooks, and also briefly mentions some criteria that were discarded. The criteria were chosen without any presumption that some categories are pedagogically superior to others.

Since having mutually exclusive categories would clearly simplify the comparison of textbooks on the basis of their examples, an attempt was made to sort examples directly into five classes:

1. Definition or an instantiation thereof: *$f(x)=e^{-x}$, $x>0$ is a pdf.*
2. Procedure for evaluating some quantity: *for Bernoulli trials,*
   $VAR = E[x^2]-(E[x])^2 = p - p^2 = pq$.
3. Application to an engineering problem: *given the pdf of voltage at the terminal of a resistor, compute the pdf of the power.*
4. Design: *optimize an 8-bit quantizer, -10V to +10V, with maximum SNR.*
5. Illustration of a phenomenon that obeys the laws of probability: *a pair of fair dice.*

Many examples fell into more than one of the predefined classes, and the authors disagreed on many assignments. Therefore instead of defining classes, the following unambiguous bipolar criteria (i.e., *dichotomies*) were formulated.

D1. Almost all examples in textbooks on probability pose a question or problem that has some (known) solution or result. The solution may be quantitative or not. A quantitative solution can be expressed in either symbolic or numerical form. Dichotomy *D1* determines whether the solution is *quantitative* or *not quantitative*.

D2. Dichotomy *D2* determines whether an example is *numerical*. (A quantitative but non-numerical conclusion *must* be symbolic.) Numerical solutions include graphs, computer printouts, and also functions with numerical parameters, like k^2 , $k= 3 , 4, \ldots, 10$, or *the [10, 25) interval on the real line.*

D3. Is the example based on some *concrete professional* process, the understanding of which requires college-level study of science or engineering? Professional processes require some knowledge of mechanics, circuit or signal analysis, device physics or fabrication, thermodynamics, information theory, etc. Thus resistive power dissipation, A/D converters, delimiter circuits, modulation, and photodiodes represent professional processes.

D4. Is the example based on a *concrete common-sense process* that requires no technical background? Coins, dice, urns, roulette wheels, temperature or stock records, pass/fail testing of circuits or chips, and the timing of telephone calls and email, are all concrete common-sense processes. In this framework, a *set of 10 uniformly distributed integers between 1 and 6* is not concrete, but *tossing 10 dice* is. If there is neither a professional nor a non-professional process, then the example is *abstract*, and the answer to both *D3* and *D4* is NO (0).

D5. Even if a problem has a solution that is itself valuable or noteworthy, sometimes the process of getting there is of even greater interest. In probability, there are often opportunities for introducing useful mathematical steps. Dichotomy *D5* therefore distinguishes examples that present or emphasize some mathematical technique or “trick,” from those that don’t.

D6. Is the solution stated? An example may either state a problem and illustrate its solution, or simply state a problem without providing the solution. An example of the latter might be: *Consider the difference between a binomial distribution with* $p = 0.4$ *and* $n = 100$*, and a Gaussian probability density function with* $\mu = 40$ *and* $\sigma^2 = 16$.

D7. Is the solution invited? If the solution is not given, it might be invited: For instance: *The Chebyshev Inequality often gives a loose bound. Find the Chebyshev bound on the probability that a Gaussian random variable is more than two standard deviations above its mean.*

D8. Some examples are self-contained, while to understand others, the reader must refer to an earlier example. *D8*, which separates self-contained from continued examples, is the simplest of the tests.

In each of eight textbooks [1] - [3], [13], [14], [20] - [21] the chapter(s) on mathematical expectation, including mean, variance and covariance, was selected. In some books the same chapter introduces random variables. In others, the chapter on expectations also covers moment generating and characteristic functions. All of the examples in the selected chapter(s) were categorized.

Although the binary criteria listed above appear simple, additional clarifications had to be developed to ensure consistent consensus. The authors independently answered eight yes/no questions about each example. In addition to the authors, a set of randomly selected examples was scored by two others. Disagreements were rare. Not all of the 2^8

combinations are possible (if a quantitative solution does not exist, then it can be neither numerical nor symbolic). The consolidated results are presented in Table 7.

As expected, most of the examples offer quantitative solutions. Two texts, Haddad and Yates & Goodman, stand out with respect to the number of numerical solutions. None of the texts require any college-level knowledge of engineering or science to understand most of the examples, but over one third of Haddad's are in this category. Abstract problem statements – without reference to any underlying physical mechanism – are favored by Fine and Stark & Woods, and, to a lesser extent, by Miller and Childers. Yates & Goodman, Feller, and Haddad, favor classical probabilistic settings like coins, dice, and balls in urns.

Table 7   Classification of textbook examples

Textbook	Feller [13]	Fine [20]	Haddad [21]	Leon-Garcia [2]	Miller & Childers [22]	Papoulis & Pillai [14]	Stark & Woods [1]	Yates & Goodman [3]
Number of examples classified	25	27	17	71	46	68	43	88
Numerical solution given	44%	44%	94%	41%	48%	51%	33%	70%
Symbolic solution given	48%	56%	6%	58%	50%	49%	67%	27%
No quantitative solution	8%	0%	0%	1%	2%	0%	0%	2%
Abstract problem statement	48%	81%	18%	54%	73%	56%	81%	14%
Prof'l/engineering process	0%	11%	35%	7%	11%	7%	2%	17%
Common sense process	52%	7%	47%	39%	16%	37%	16%	69%
New or instructive maths	40%	7%	0%	18%	35%	19%	51%	2%
Continued example	48%	11%	18%	17%	22%	6%	26%	44%
Solution invited	0%	0%	0%	0%	7%	0%	0%	0%

Stark & Woods, Feller, and Miller and Childers go out of their way to present examples that introduce interesting mathematical techniques (or perhaps they simply do not avoid such examples.) Although not measured by the dichotomies, Feller, Fine, Miller & Childers and Stark & Woods require more advanced mathematical preparation than the other books.

V. CONCLUSIONS

Unlike many engineering problems where the optimality of solutions can be proved or the effectiveness of an approach may be quantified, the effectiveness of pedagogic examples in a classroom and the classification of such examples is difficult to judge. The difficulty arises because of inherent differences in the learning styles of individuals, their background, and interest in the subject. Nevertheless, the evaluations are far from random. Undergraduate engineering students in a Rensselaer course on probability prefer simpler examples to those examples that may show key concepts, a choice which contrasts with that of graduate students and professors. Contrasting pairs of examples are appreciated by all groups. Student opinions on an example are more consistent than the opinions of professors. These observations were pronounced enough for our relatively small sample to provide statistically significant results. Whether they apply to universities with different demographics would require further experimentation.

According to eight binary criteria, different authors show marked differences in their choice of various types of examples in their textbooks. The type of examples that were selected for the survey, generic problems with numerical solutions, represents the most popular category of examples in the textbooks that were examined.

The dichotomies for characterizing textbook examples could be applied, perhaps with minor changes, to texts for other mathematics courses in engineering curricula. In combination with additional criteria (topical coverage, length, vocabulary, notation) they may be valuable to publishers, authors and teachers. They could also be embedded into automated search techniques to harvest examples from the web.

The survey instrument, and the corresponding statistical tests, could also be applied to evaluate examples in other domains. Alternatively, other aspects of examples could be studied in a similar manner. In view of the results of the textbook classification, it would be interesting to find out whether situating examples in an application context would enhance them in the view of the various constituencies. For instance, examples for computing the correlation coefficient could include variables encountered in practice, like temperature and switching delay in a CMOS inverter. Then any positive or negative correlation could be linked to some underlying physical phenomenon.

What the proposed methodology lacks so far is any direct assessment of whether any examples are more conducive to students learning the exemplified procedure and concept. Examination results will not reveal the extent of learning if the examinations are based on test problems that are too similar to the examples presented in the course. The taxonomy of examples and the methodology for assessing the preferences of students and professors developed here may provide a starting point for studies of the *effectiveness* of examples.

ACKNOWLEDGMENTS

We are grateful to Cathy Johnson, Love Library, University of Nebraska-Lincoln and Prof. James Cordy, School of Computer Science, Queens University for suggestions and pointers, and to our students and colleagues for participating in the survey. Thoughtful suggestions of four reviewers improved both the analysis and the presentation.

REFERENCES

[1] H. Stark and J.W. Woods, *Probability and random processes with applications to signal processing,* Third Edition, Prentice-Hall, Upper Saddle River, NJ, 2002.

[2] A. Leon-Garcia, *Probability and random processes for electrical engineering,* Second Edition, Addison-Wesley, Reading, MA 1994.

[3] R. D. Yates and D. J. Goodman, *Probability and Stochastic Processes (A Friendly Introduction for Electrical and Computer Engineers),* John Wiley and Sons, New York, 2005.

[4] P. Byrd, C-W. C. Liu, A. Mobley, A. Pitillo, S. Rolf Silva, and S.S.-Wen, “Exemplification and the Example,” Issue 1 of the *Journal of English Grammar on the Web.* *http://www2.gsu.edu/~wwwesl/issue1/extitle.htm* (consulted on July 5, 2006).

[5] S. Yelon, and M. Massa,. “Heuristics for creating examples,” *Performance and Instruction*, pp. 13-17, October 1987.

[6] H. Sweet, *The practical study of languages: A guide for teachers and learners.* London: Oxford University Press, 1964.

[7] J.A. Brandao Faria, “A pedagogical example of the application of transmission-line theory to study conductor grounding effects,” *IEEE Trans. Edu.* 47, pp. 141 – 145, Feb. 2004:

[8] E.H. Newman, “Simple examples of the method of moments in electromagnetics,” *IEEE Trans. Edu.* 31, pp. 193 – 200, Aug. 1988.

[9] W.D. Blair, J.E. Conte, T.R. Rice, "An instructive example of homomorphic signal processing," *IEEE Trans. Edu.* 38, pp. 211 – 216, Aug. 1995.

[10] J.R. Cordy, "Notes on Program Testing and Correctness, CISC 211," Report, Queens University, Kingston, Ont., 2005.

[11] IEEE Computer Society, *Proceedings of the 20th IEEE VLSI Test Symposium* (VTS 2002), Monterey, CA, USA. 2002.

[12] C. Ash, *The Probability Tutoring Book*, IEEE Press 1993.

[13] W. Feller, *An introduction to probability theory and its applications, Volume 1,* Second Edition, John Wiley and Sons, New York, 1960.

[14] A. Papoulis and S.U. Pillai, *Probability, random variables and stochastic processes*, Fourth Edition, McGraw Hill, Boston, 2002.

[15] T. Sumner, M. Khoo, M. Recker and M. Marlino, "Understanding educator perceptions of "quality" in digital libraries," *Proceedings of ACM/IEEE Joint Conference on Digital Libraries,* 2003.

[16] C. Cheng, A. Kumar, J. Motwani, A. Reisman and M. Madan, "A citation analysis of technology innovation management journals," *IEEE Trans. Eng. Manag.* 46, pp. 4-13, February 1999.

[17] R. Likert, ``A technique for the measurement of attitudes," Archives of Psychology, vol. 140, no. 55, 1932.

[18] H.B. Mann and D.R. Whitney, "On a test of whether one of 2 random variables is stochastically larger than the other," *Annals of Mathematical Statistics*, vol. 18, pp. 50-60, 1947.

[19] R.A. Fisher, "On the interpretation of $\chi 2$ from contingency tables, and the calculation of P," *Journal of the Royal Statistical Society*, vol. 85, no. 1, pp. 87-94, 1922.

[20] T.L. Fine, *Probability and Probabilistic Reasoning for Electrical Engineering,* Prentice Hall, Upper Saddle River, NJ, 2006.

[21] S.L. Miller and D.G. Childers, *Probability and random processes with applications to signal processing and communications,* Elsevier, Amsterdam, 2004.