Teaching Logic and Decision Making Using Probability & Statistics Course

Roes Arief Budiman
Department of Mechanical and Manufacturing Engineering
University of Calgary
Calgary, Alberta T2N 1N4, Canada
Email: rbudiman@ucalgary.ca

BACKGROUND
Probability and statistics course for second year engineering students is often taught as an applied mathematics course by emphasizing using ready formulas and memorizing procedural problems. This teaching approach is reinforced by textbook format that overwhelms students with too much information. Having taken calculus and linear algebra courses in first year, such approach can further diminish students' appetite in linking mathematics with real world engineering problems. Probability and statistics course needs to be taught with a different emphasis to improve students' analytical thinking skills.

PURPOSE
Probability and statistics course can be taught to include aspects of logic, such as inference and deductive reasoning, if problems presented are drawn from real world problems. These problems naturally teach students how to make decisions using mathematical equations under incomplete information. This paper describes an ongoing development of an engineering probability and statistics textbook in 2012-2014, which strives to teach logic and decision making as well.

DESIGN/METHOD
There are two approaches explained. First, the textbook format resembling a novel with detailed explanations for equations and worked examples will be described. Ideas that motivitate the textbook format in its three editions will be explained: connectivity and applications of formulas. Real world engineering problems as exam questions will be shown. Second, students' survey data on effectiveness of the textbook in teaching logic and decision making will be discussed.

RESULTS
Students have indicated that probability and statistics course is useful for their career. This is a significant change since previously students regard the course as a 'dry' math course. Data showing the effect of class size will be shared.

CONCLUSIONS
The ongoing course delivery and textbook development will be discussed in terms of success level in teaching logic and decision making for engineering students. Future additions to the course, such as group projects and case studies, will be discussed.

Keywords
Probability and statistics. Teaching logic and decision making. Critical thinking skills.

1. INTRODUCTION

Probability and Statistics for Engineers ENGG 319 course at the Schulich School of Engineering, University of Calgary is a second-year common-core course for all programs but Electrical Engineering that offers a similar course for its third year students. There were about 530 students taking ENGG 319 in Fall 2013 term, representing 75% of the total engineering cohort. ENGG 319 is an introductory probability and statistics course that covers notions of probability; conditional probability; discrete and continuous probability distribution functions; confidence intervals; hypothesis tests; joint probability distributions including correlation coefficient and maximum likelihood estimator.

ENGG 319 students can be introduced to elements of logic when notions of probability are discussed. The introduction to logic can be performed using the perspective that we assign a probability value to a sentence or statement that represents an argument leading to a conclusion, while in logic we are more concerned with assigning true or false value. For instance, one deductive reasoning prescribes that if "a statement p is true, then a statement q follows", then "the negation of q will necessarily require the negation of p." Such example can also introduce conditional probability concept and sets up a scenario to discuss hypothesis test.

Teaching probability and statistics with logic as outlined above cannot be aligned with nor is supported by standard textbooks of probability and statistics for engineers (Johnson, 2011; Walpole, 2012; Montgomery, 2011). A standard textbook covers the course topics sequentially with little or no discussion on logic. Engineering students at the University of Calgary can take an introductory logic course as a non-science optional course. Such logic course, however, is focused on formal aspects of logic and does not provide connections between logic and probability.

The standard textbooks focus on helping students understand example problems by using a variety of figures, concept summaries, and equations. They use different formats to highlight figures, concepts, and equations from the main text. Important equations are placed in boxes with coloured background. Concept summaries are placed in separate paragraphs with lines separating them from the main text. These different formats for figures, concept summaries, and equations slow down the reading process. Although the formatting strategy succeed in highlighting them, it prevents a textbook from developing one good, coherent story that connects all course topics. Students instead focus on the highlighted parts and learn to solve exercise problems by memorizing mathematical expressions of the distributions and use their exercise problem database to search for a solution template. As a result, students are implicitly trained to collect as many exercise problems as possible and to learn by *inductive* reasoning.

It is not surprising that engineering students develop pattern-matching skills when solving problems since most texbooks present the course topics as separate stories. Instead of providing mathematical and logical connections from hypergeometric to binomial to normal distribution, most textbooks present these distributions separately. The absence of connections gives an impression to students that the course topics are a collection of *distinct* mathematical formulas that can only be memorized.

This paper describes an ongoing project since 2012 to teach probability and statistics for engineers in a *deductive* approach, whereby students learn the concepts and the connections among the course topics and apply them to problem solving. Under this teaching approach, it is expected as a learning outcome that, for example, a student can explain *why* a certain distribution function is the most appropriate for a problem before he solves the problem. It is also expected as another learning outcome that students can interpret what the probability value means qualitatively for a given problem. This probability value interpretation is not easy to teach since students have to understand the flow of information leading to solution and the general features of the relevant equations. Introductory logic is presented during first tutorial and it provides an initial point to continue inserting logic into the course topics. Decision making aspects of probability and statistics are discussed throughout the course by blending the numerical probability answers with their economic and cost consequences. This course redesign requires several steps to implement as explained below.

2. COURSE REDESIGN

2.1 TEXTBOOK

The main teaching tool for the course redesign is a suitable textbook. Acosta (2000) stated that engineering statistics course has been dominated by formal mathematical components. This characterization is not entirely accurate, as alluded in the Introduction, since mathematical statistics textbooks (Feller, 1968; Hogg, 1995) depart markedly from engineering statistics textbooks (Johnson, 2011; Walpole, 2012; Montgomery, 2011) in presenting the course topics. Mathematical textbooks explain concepts and give proofs of relevant theorems, effectively providing connections between the course topics, but they don't provide a sufficient number of examples to reflect upon. On the other hand, engineering textbooks focus on using mathematical formulas and figures in order to solve example problems.

A textbook plays a dominant role in determining teaching style. I therefore decided to write a probability and statistics for engineers textbook in order to achieve two goals: (i) Presenting a coherent story starting with notions of probability and ending with the joint probability distribution; (ii) Simplifying the writing formats so that the textbook reads more like a novel than a standard textbook.

Probability and stastistics for engineers course is essentially an applied mathematics course for audience who is not receptive to reading mathematical proofs and theorems. Proof of central limit theorem is thus contained in an appendix; students don't have to read it to gain a qualitative understanding of the reasons why normal distribution becomes the distribution for sampling activities. Having such proof in the textbook helps convince students that confidence intervals are a logical result of central limit theorem. Derivations of chi-square and t distribution are available in appendices. Derivations that simplifies hypergeometric distribution into binomial distribution are provided in details in the main text, however, since students need to see how one distribution function can develop into another when an assumption, which in this case is a constant probability for a trial, is inserted. Detailed derivations to yield normal distribution from binomial distribution is also provided in the main text in order to introduce normal distribution as a continuous version of binomial under 50% trial probability and infinite trial number assumption.

The redesign textbook format is simpler than the standard textbooks'. It has no boxes for important equations and separate paragraphs for concept summaries. It is single-column printed on double-sided A4 papers and has hand-drawn figures and illustrations. There are 10-20 solved example problems per chapter and they occupy about 50% of the pages. Each example is used to discuss further concepts and equations developed in the pertaining sections. To include decision making aspects of the course, several example problems are selected for further discussion.

As illustration, one example problem that gives further discussion on decision making is as follows.

> "An airline carrier notices that 3% of reserved seats result in no-shows at the airport gate for whatever reason. It has a fleet of airplanes with 220 seats each at full capacity. What is the minimum number of reservations it should set for the airplane so that it has a 95% probability that it will be fully seated?"

If the airline reserves the number of passengers equal to the number of seats, then there is only $0.97^{220} = 0.1\%$ probability for the airplane to be fully seated. The probability for 215 to 220 occupancy is given by binomial distribution:

$$\sum_{x=215}^{220} \frac{220!}{x!\,(220-x)!} 0.97^{x} 0.03^{220-x} = 35\%;$$

thus, it is unlikely (35% probability) to have close to full occupancy (215 to 220) if only 220 reservations are taken. Decision making perspective of an engineer comes about when he translates the probability information into profit-loss estimation. For instance, even though on average there are $0.97 \times 220 = 213.4$ passengers fill up the airplane, this average passenger number implies a steady-state cash flow condition and the airline may not be in that financial stage.

If the airline is not financially healthy, it may be more realistic to start a profit-making initiative by having reservation number above 220. It will ensure a much higher probability of a full flight at a possible cost of having to financially compensate stranded passengers. In this case, binomial distribution is not valid anymore since the variable is the number of reservations in order to have 220 passengers. The correct distribution is negative binomial and the question is answered by

$$P(220 \le k \le m) = \sum_{k=220}^{m} \binom{m-1}{220-1} 0.97^{220} 0.03^{k-220} = 0.95.$$

This equation is solved iteratively by increasing m until the probability reaches at least 95%. The solution is $m = 231$, which yields 95.2% probability. The average number of reservations for a full flight is given by the negative binomial average: $220/0.97 \approx 227$. Profit-loss estimation can be made by incorporating fixed cost, revenue per person, and compensation cost.

2.2 TUTORIALS AND ASSESSMENTS

While lecture hours (150 minutes per week) are used to present theories qualitatively and to explain solved example problems, tutorials (90 minutes per week) are the more appropriate venues to introduce connections to logic and decision making after relevant probability and statistics topics are covered in class. These connections are best made using long exercise problems and probability experiments, e.g., drawing beads randomly from a jar filled with beads of two different colours in order to illustrate sampling with or without replacement and their respective probability density functions.

To provide incentives to students in developing their analytical thinking skills, all course assessment tools – three quizzes, one midterm exam, one final exam – are open notes and open textbook. Because students can bring their lecture notes and textbooks to exam rooms, they spend their study time exclusively on reviewing problem solving techniques with minimum memorization. Almost all students taking the course have never written an open book exam before, and to prevent students from taking for granted the difficulties of such exam format proper study strategy is discussed at the first class. There has been no objection from the students on the open book exam format.

3. STUDENT SURVEY RESULTS

Students taking the course in Summer 2013 took the online anonymous student survey with 78% participation rate (32 out of 41 students). There is an almost 2:1 margin on whether the textbook explains the concepts well, as shown in Figure 1. In Summer 2013 term, I used the textbook's First Edition that doesn't have solved example problems.

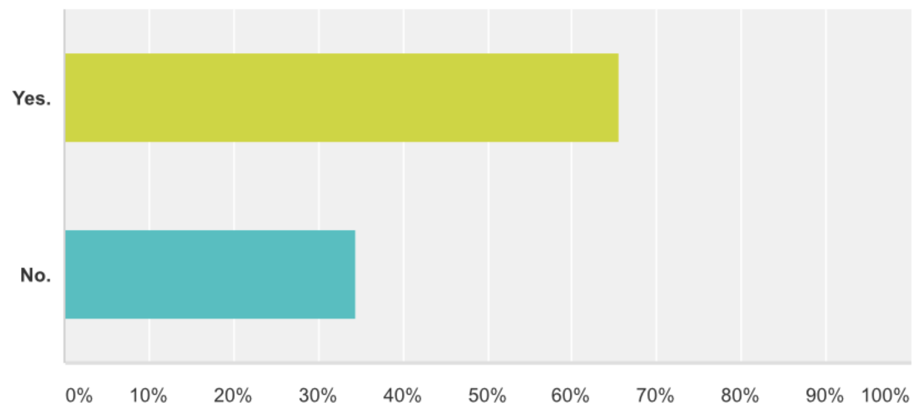

Figure 1. On whether the textbook explains the concepts well (Summer 2013 term).

Figure 2 shows an almost opposite response when the textbook's Second Edition was used in Fall 2013. There were no major changes between First and Second Edition. Same lecture notes were used. Participation rate was 20% for the Fall 2013, and out of 100 students, only 29 responded Yes.

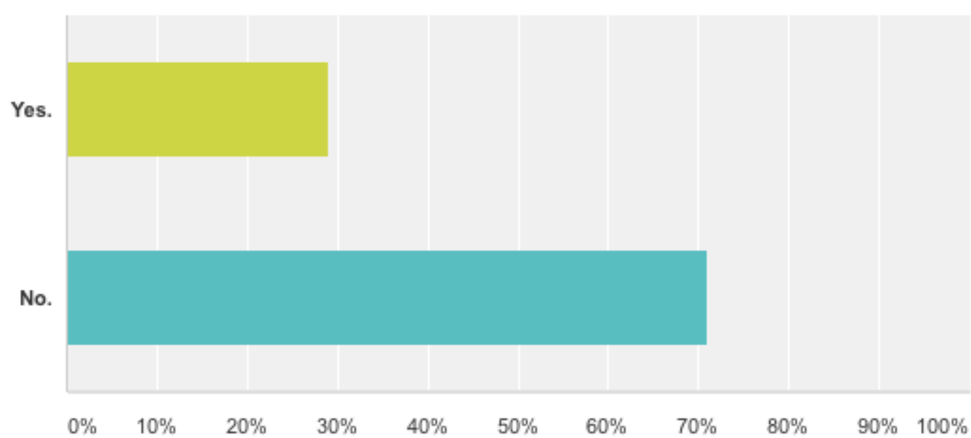

Figure 2. On whether the textbook explains the concepts well (Fall 2013 term).

Figure 3 shows the Summer 2013 student response on whether lectures apply the concepts well through class examples. 28 out of 32 responded positively to the question. Figure 4 shows the Fall 2013 student response, where 84 out of 100 responded positively to the question. The two response numbers are statistically identical and go to show that the lecture delivery and content did not change.

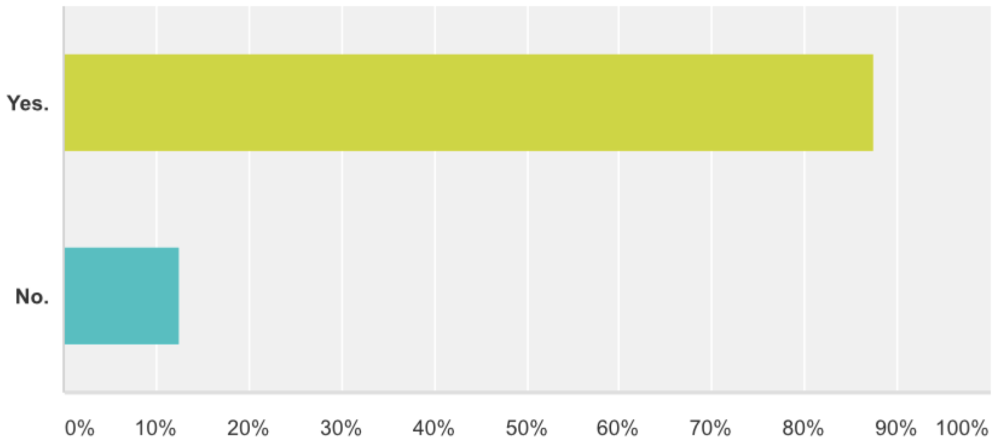

Figure 3. On whether lectures apply concepts well through class example problems (Summer 2013).

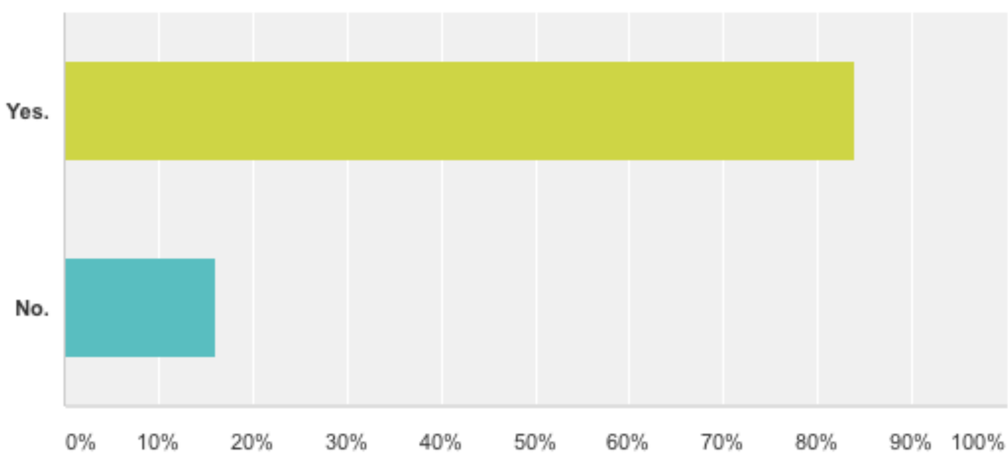

Figure 4. On whether lectures apply concepts well through class example problems (Fall 2013).

What changed between Summer and Fall 2013 was the class size. The Summer class had 41 students, while the Fall course had 4 classes with about 130 students in each class. The three-fold student number increase does not affect the learning experience adversely in class, but it affects the *reading experience* adversely for a large class.

The First and Second Edition did not have solved example problems. Both editions present the concepts and explain connections between the course topics through mathematical derivations. This textbook format did not pose a major issue for a small class size in Summer 2013 since students had ample time to interact with course instructor and, correspondingly, pay attention more to topics discussed in class and tutorial. Such interactive environment effectively disappears in a large class size. Students in a large class size, therefore, require more reading supports to reiterate what they have heard in class. To respond to the large class' need, the Third Edition has incorporated solved example problems. This Summer and Fall student survey comparison has resulted in a major revision of the textbook.

Students' perception to the course is positive, as shown in Figure 5. There is almost a 2:1 margin on whether the course is useful for engineering career. Students wanted to learn and paid attention to the course material. This positive perception was not apparent in previous years as the course was perceived as a uninteresting (dry) math course, heavy with formula memorization.

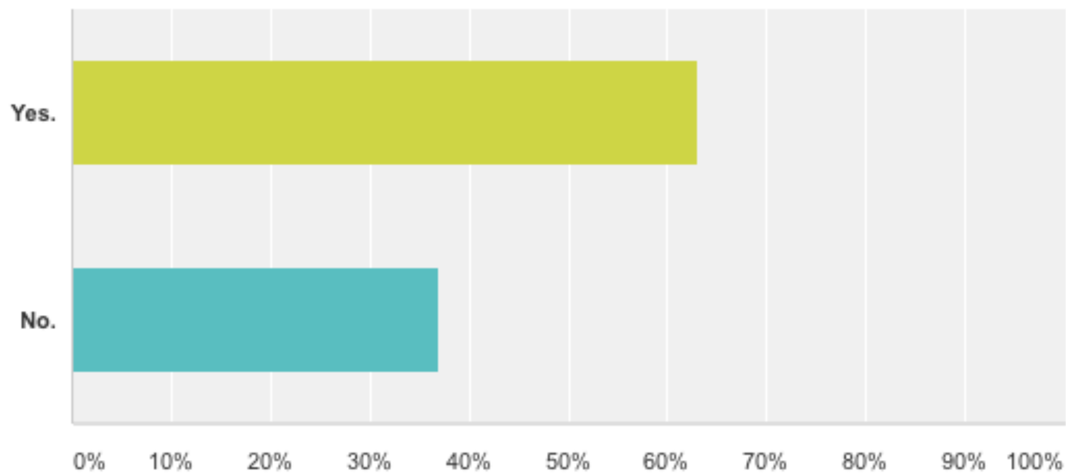

Figure 5. On whether ENGG 319 Probability and Statistics for Engineers course will be useful for engineering career (Fall 2013).

One more way to change the course perspective is to give more practical-oriented exam question. This can be done without sacrificing the rigor and difficulty level of exam questions. In Fall 2012, when the course was taught using Montgomery & Runger's textbook (Montgomery, 2011), the most difficult midterm exam question was

> "A joint density function is given as $f(x,y) = 6(1-y)$ for $0 < x < y < 1$. Find $E[x]$,"

while the most difficult Fall 2013 midterm question, when the redesign textbook was adopted, was

> "The two chemical engineers are evaluating the probability of localized corrosion which will create a small hole of several millimeters in diameter and will cause an oil leakage. Given that there is a 4% probability of one-corrosion induced hole formation per 5 km due to metallurgical and electrochemical factors, how long should the pipeline be before there is a 99% probability that there is one hole?"

The most difficult Fall 2012 final exam question was

> "A learning experiment requires a rat to run a maze (a network of pathways) until it locates one of 3 possible exits. Let Y_i denote the number of times exit i is chosen in successive runnings. Assume that the rat chooses an exit at random on each run. For general n, find $Cov(Y_2, Y_3)$,"

while the most difficult Fall 2013 final exam question was

> "Electrical components have failure statistics that obeys Weibull distribution. Tests indicate that 5% of the electrical components fail within the first 2 years of continuous use and that 5% are still functioning after 10 years of continuous use. Assuming that inspected components were installed at the same time, determine the probability that during an inspection at 5 years of continuous use there are at most 4 components that must be tested before one failed component is found."

The more practical-oriented exam questions improved average midterm marks by 10%, which I believe was not caused by easier exam questions, as can be observed by the two midterm exam questions above. Further data collections are required to verify this hypothesis.

4. FUTURE WORK AND CONCLUSIONS

The next step in improving teaching probability and statistics for engineers course are to design group design projects where students work together on probability or statistics experiments. Example projects are (i) statistical distribution of customers visiting coffee shops and (ii) determining distribution parameters using real data to compare two related probability distribution functions.

Appropriate textbook is required for a significant redesign of probability and statistics for engineers course in order to teach students analytical thinking skills. It appears that most standard textbooks for the course are not compatible with this teaching goal as these textbooks emphasize students' knowledge mastery instead of problem solving mastery. One solution offered in this paper is to simplify the textbook format to having two content types: (i) presenting the course topics as one connected, coherent story and (ii) solved example problems that explain in details how concepts are applied for problem solving.

REFERENCES

Johnson, R. A. (2011). *Miller & Freund's Probability and Statistics for Engineers.* Boston: Prentice Hall.

Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). *Probability & Statistics for Engineers & Scientists.* Boston: Prentice Hall.

Montgomery, D. C., & Runger, G. C. (2011). *Applied Statistics and Probability for Engineers.* Hoboken: John Wiley & Sons, Inc.

Acosta, F. M. A. (2000). Hints for the improvement of quality teaching in introductory engineering statistics courses. *European Journal of Engineering Education*, 25(3), 263-280.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Volume I*. Hoboken: John Wiley & Sons, Inc.

Hogg, R. V., & Craig, A. T. (1995). *Introduction to Mathematical Statistics*. Upper Saddle River: Prentice-Hall, Inc.