A PROJECT-ORIENTED COURSE IN PROBABILITY AND STATISTICS FOR UNDERGRADUATE ELECTRICAL ENGINEERING STUDENTS

Maurice F. Aburdene and Richard J. Kozick

Electrical Engineering Department
Bucknell University, Lewisburg, PA 17837
aburdene@bucknell.edu, kozick@bucknell.edu

ABSTRACT

An introductory course in probability and statistics for third-year and fourth-year electrical engineering students is described. The course is centered around several computer-based projects that are designed to achieve two objectives. First, the projects illustrate the course topics and provide hands-on experience for the students. The second and equally important objective of the projects is to convey the relevance and usefulness of probability and statistics to practical problems that undergraduate students can appreciate. The benefit of this course is to motivate electrical engineering students to excel in the study of probability concepts, instead of viewing the subject as one more course requirement toward graduation. The authors co-teach the course, and MATLAB is used for most of the computer-based projects.

1. INTRODUCTION

At Bucknell University, electrical engineering students are required to take a probability course taught by the mathematics department, or an elective random signals and noise course taught by the electrical engineering department. The mathematics department offers the probability course every semester, and electrical engineering students take the course with mathematics majors. The electrical engineering department requires students to take the probability course, while other engineering departments require a half-course titled "Statistics for Engineers". The electrical engineering course has been listed in the catalog as an elective, but has not been taught recently. We decided to offer the electrical engineering course using the approach described in this paper since we observed that many students were dropping the mathematics course. In the spring of 1997, 26 students enrolled in the electrical engineering course, and approximately three students enrolled in the mathematics course. The mathematics department and electrical engineering department have an excellent relationship and are continuously striving to improve the education of the students. In this paper, we will describe some elements of the recent offering of the random signals and noise course.

It is widely acknowledged that the area of probability and statistics presents challenges for students and teachers alike. As evidence, we submit the following quotes from two authors of probability and statistics textbooks.

> "Probability is a hard subject. Students find it hard, textbook writers find it hard." (Carol Ash in [1])

> "Probability and its applications to engineering problems is perhaps one of the more difficult bodies of knowledge for the undergraduate engineering student to grasp."
> (Rodger Ziemer in [2])

> "I have taught probability and random processes to engineering students at both the undergraduate and graduate levels several times over the span of 30 years of teaching. I have yet to teach the ideal course." (Rodger Ziemer in [2])

Although many texts on probability and random signals are available, we found it difficult to choose an appropriate text that appeals to both faculty and students. Some texts are very mathematical, and some very descriptive. We have decided to develop a course based on computer projects using Matlab [3] that are drawn from electrical engineering to demonstrate the concepts included in most introductory texts. The projects introduce the basic and fundamental probability and statistics concepts through engineering examples and applications using a computational package. Many textbooks describe approaches for introducing disciplinary fundamentals. This paper is more concerned with projects that are related to EE to help students understand probability and statistics, and to help students see the relevance of probability and statistics.

The projects were chosen by the two faculty members teaching the course. As with any course, the choice of topics and examples is influenced by the background, experience, and interests of the faculty members. Such is the case in this course. The projects can serve as appropriate models for the introduction of the fundamental subjects in probability and statistics. While any set of projects may be considered incomplete by some instructors, the projects presented are valuable either as a set, or individually. Instructors can benefit by incorporating aspects of this approach into their courses. The first four course projects are described in this paper.

2. A MODEL FOR A DISK ACCESS SYSTEM

The first project focuses on the concept of the probability of an event. The project develops an intuitive understanding of probability. The students are introduced to uniformly distributed random numbers, the ideas of histograms, scatter diagrams, mean, variance, and median. We also show students how to use Matlab efficiently by using matrix and vector formulations, as opposed to using for-loops or while-loops as in programming languages. Other random number generators are introduced, such as exponential and Gaussian. Students are asked to look at the histograms, scatter diagrams, mean, and variance for these other distributions.

Some of these concepts are introduced through a model for data access from a computer disk. The model assumes that data is read from the disk correctly with probability P on any trial. Students investigate various aspects of the model, including the probability of reading the data correctly in N trials, and the effects of P and N on the speed and reliability of the disk access. This model can be used for other success/failure type systems, including message transmission systems and Internet access systems.

A sample Matlab script for simulating the disk access system is shown in Figure 1 on the following page. Students are asked to write a simulation program on their own. Most students without prior Matlab experience use for-loops and if-else statements. We use the script in Figure 1 to introduce students to more advanced features in Matlab, which are useful for the later projects in the course.

The students are asked to plot a family of curves to display the effects of P and N on the total probability of success in N trials, which is $P_N = 1 - (1 - P)^N$. An example is shown in Figure 2. These curves are used for system design based on this model. Also, students are asked to investigate the sensitivity of P_N with respect to N and P analytically and with Matlab. As a variation on this project, the students are asked to investigate (using simulation and/or analysis) the average number of trials N that are required for successful disk access.

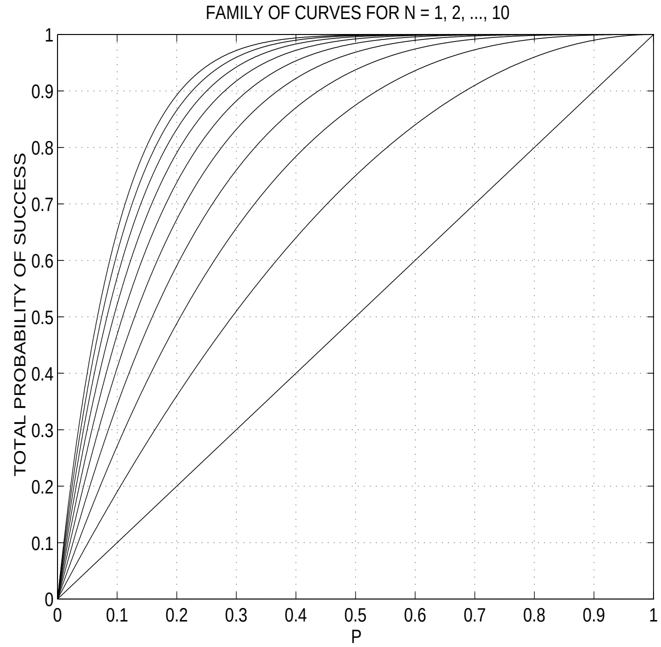

Figure 2: *Family of curves for total probability of success* $P_N = 1 - (1 - P)^N$ *versus* P *for* $N = 1, 2, \ldots, 10$.

3. LINEAR AVERAGING

A sine wave is sampled to produce a data set that is then used to explain the concepts of sample mean, variance, and standard deviation. Histograms of the data are produced to motivate the concept of a distribution function. Random samples from various probability distributions are viewed to build an intuitive understanding of these important concepts. Here we relate the root-mean-square (RMS) value of sinusoidal signals that students have studied in circuit analysis courses to the definition of *variance*. Also, the concept of signal-to-noise ratio (SNR) is introduced.

The project requires students to generate N traces of discrete-time sinusoidal signals in which independent uniformly-distributed noise is added to each sample. Then a linear averaging of the traces is computed, which is similar to the processing that is performed in digital oscilloscopes [5]. Students develop a Matlab program that allows them to investigate the effect that the number of traces N has on the SNR and the appearance of the averaged signal. Figure 3 illustrates the noise reduction achieved by averaging $N = 10$ traces. The SNR is improved from 7.9 dB in a single trace to 17.7 dB in the averaged waveform.

```
% Matlab simulation of disk access system for N = 2 tries and probability
% of success P.

P    = input('Enter the value for P: ');
Runs = input('Enter the number of simulation runs: ');

x = rand(2, Runs); % Generate uniform random numbers
x1 = x(1,:); % Extract 1st row for simulation of 1st try
x2 = x(2,:); % Extract 2nd row for simulation of 2nd try

% Identify successes on first trial
s1 = (x1 > (1-P)); % s1 = 1 for success and s1 = 0 for failure
loc1 = 1-s1; % loc1 puts 0's in locations of success

% Now perform 2nd trial, only when 1st trial failed
x2 = loc1 .* x2; % Place 0's in x2 when success on 1st try
s2 = (x2 > (1-P)) ; % Identify success on 2nd try with s2 = 1

success = sum(s1 + s2); % Now s1 and s2 have 1's when successful
fprintf('Number of successes = %d in %d simulations\n', success, Runs);
```

Figure 1: *Matlab script to simulate disk access system that introduces students to matrix/vector capabilities in Matlab.*

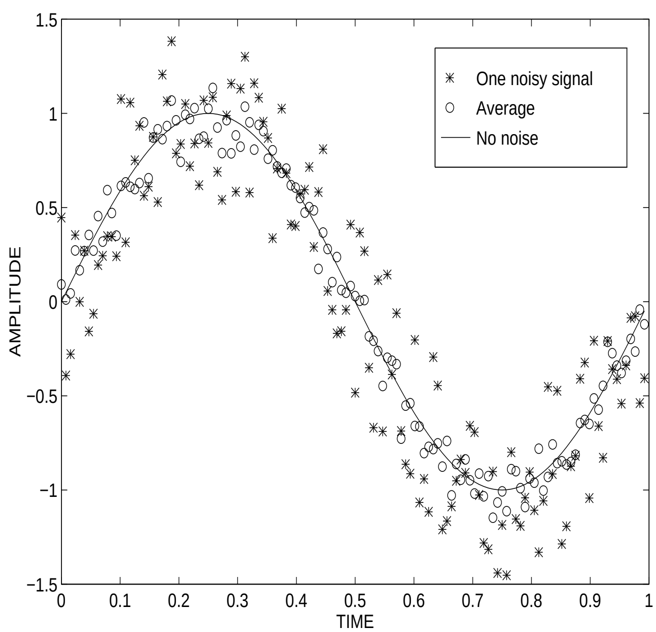

Figure 3: *Effect of averaging 10 waveforms to reduce noise.*

4. PARAMETER ESTIMATION

The concept of least-squares parameter estimation is introduced with a simple example of fitting a straight line to a data set. Then least-squares estimation is applied to linear system parameter identification from measured frequency response data. Explicit solutions are derived for first-order and second-order systems, and these are applied to measured data from RC and RLC circuits. The project assignment follows .

Project Introduction

The first part of this project is to design a Matlab program to obtain the frequency response of a series RLC circuit and add noise to it to simulate laboratory measurements. The program should plot the frequency response, the frequency response plus noise (simulated laboratory measurements), and also estimate the resonant frequency, bandwidth, and quality factor of the filter from the noisy measurements. We ask the student to try various noise distributions and noise power. The program design should allow the user to compare the estimated results with the true values, based on the nominal element values in the circuit. It should also do a least-squares fit between the noisy frequency response and a model for the frequency response.

The series RLC circuit that we will work with is shown in Figure 4. The resistance r is the internal resistance of the inductor L. The transfer function of the

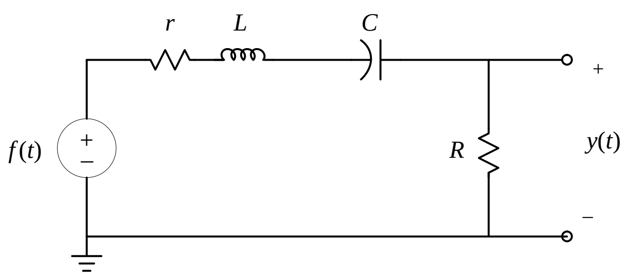

Figure 4: *Series RLC circuit.*

circuit is

$$H(\omega) = \frac{V_0}{V_1} = \frac{R}{(R+r) + j(\omega L - \frac{1}{\omega C})}.$$

The resonant frequency ω_0, the bandwidth B, and the quality factor Q for this circuit are related to the element values according to

$$\omega_0 = \frac{1}{\sqrt{LC}}, \qquad B = \frac{R+r}{L}, \qquad Q = \frac{\sqrt{L/C}}{R+r}. \qquad (1)$$

Least Squares Modeling

From the bell-shaped appearance of the frequency response, we might hypothesize that the circuit can be modeled by a second-order transfer function of the form

$$|H_{model}(\omega)|^2 = \frac{\omega^2}{X + Y\omega^2 + Z\omega^4}. \qquad (2)$$

Equation (2) is a frequency-domain *model* for the system with parameters X, Y, and Z. It is easy to verify that the theoretical transfer function for the RLC circuit has the form in (2), and is

$$|H_{theo}(\omega)|^2 = \frac{\omega^2}{\left(\frac{1}{RC}\right)^2 + \left[\left(\frac{R+r}{R}\right)^2 - \frac{2L}{R^2C}\right]\omega^2 + \left(\frac{L}{R}\right)^2\omega^4}. \qquad (3)$$

Our objective in this section is to develop a procedure for choosing the parameters X, Y, and Z in (1) so that the model $|H_{model}(\omega)|^2$ is a good approximation to the noisy frequency response $|H_{meas}(\omega)|^2$ that would be obtained through laboratory measurements. Then, once X, Y, and Z are known, you should be able to to combine (2), (3), and the equations for ω_o, B, and Q in (1) to derive ω_o, B, and Q from X, Y, and Z.

Solving for the optimum X, Y, and Z directly from (2) is difficult, because $|H_{model}(\omega)|^2$ is a *nonlinear* function of X, Y, and Z. Note, however, that the reciprocal is a linear function of the parameters:

$$G_{model}(\omega) \equiv \frac{1}{|H_{model}(\omega)|^2} = X\omega^{-2} + Y + Z\omega^2. \qquad (4)$$

Thus it is simpler to find X, Y, and Z that make $G_{model}(\omega)$ a "best approximation" to the reciprocal of the measured data, $G_{meas}(\omega) \equiv 1/|H_{meas}(\omega)|^2$.

By "best approximation" we mean that the total squared error between the experimental data $G_{meas}(\omega)$ and the model $G_{model}(\omega)$ is as small as possible. The measured frequency response is only available at N frequency values, and let us denote those (radian) frequency values by $\omega_1, \omega_2, \ldots, \omega_N$. Then total squared error is defined as follows:

$$\begin{aligned} e_n &= G_{model}(\omega_n) - G_{meas}(\omega_n) \qquad (5) \\ E &= \sum_{n=1}^{N} e_n^2 \\ &= \sum_{n=1}^{N} \left[G_{model}(\omega_n) - G_{meas}(\omega_n)\right]^2 \\ &= \sum_{n=1}^{N} \left[X\omega_n^{-2} + Y + Z\omega_n^2 - G_{meas}(\omega_n)\right]^2 \qquad (6) \end{aligned}$$

The parameters X, Y, and Z are chosen to make E in (6) as small as possible.

Solving for the best X, Y, and Z is a multivariable optimization problem. However, it is the simplest type of such problems, since E in (6) is a three-dimensional parabola with a single, global minimum. We must find the X, Y, and Z that correspond to the minimum of the parabola in (6). This is easily done by setting the partial derivatives equal to zero, and solving a set of three linear equations in the three unknowns X, Y, and Z. Matlab will solve the equations very easily. Then we can plot the least-squares fit $|H_{model}(\omega)|^2 = 1/G_{model}(\omega)$ using the values of X, Y, Z obtained.

Project Report Guidelines

Student are required to submit an individual project report, in which they :

- Present the noise-free response, the noisy response, and the least-squares model plotted on the same axes for easy comparison.
- Explain how the resonant frequency, bandwidth, and quality are estimated from the measured data.
- Compare the "simple" method with the least-squares approach and note which is more accurate. (Which approach produces better estimates when only a few noisy frequency response values are available?)

In general, the report should be clearly written and self-contained. A reader should understand exactly what was done by reading the report.

n	Probability $P(n)$
-2	0.1
-1	0.2
0	0.4
$+1$	0.2
$+2$	0.1

Table 1: *Noise amplitudes and their probabilities* $P(n)$.

5. DIGITAL COMMUNICATION RECEIVER

The more traditional mathematical topics of sample space, outcomes, events, the axioms of probability, conditional probability, and Bayes' Theorem are introduced through a digital communication application that is described in this section.

We want to study the design of a simple digital communication system. Let us consider a model for the noisy observations x that are measured by a digital receiver of the form

$$x = s + n,$$

where $s = +1$ or -1 is the signal and n is additive noise. The binary signals are equally likely with $P(s = -1) = P(s = +1) = 0.5$, and the noise takes one of 5 values, with probabilities $P(n)$ as shown in Table 1. The objective is to design a detector that decides which signal s was transmitted based on the noisy observation x. This detector should minimize the probability of making an error.

An outline of our approach in this project is as follows.

- A tree diagram is constructed that shows the possible received values x. The concept of conditional probability is explained using the tree diagram. The probabilities $P(x|s = -1)$, $P(x|s = +1)$, $P(x$ and $s = -1)$, $P(x$ and $s = +1)$, and $P(x)$ are computed for each possible x value.

- The conditional probabilities $P(s = -1|x)$ and $P(s = +1|x)$ are computed for each x, and the decision rule is introduced that chooses the signal $s = -1$ or $s = +1$ according to which of these conditional probabilities is larger.

- The results from this particular application are stated in terms of the general formulas for conditional probability and Bayes' theorem.

- The probability of a bit error is computed for the detector. This probability is related back to the disk access system considered in Section 2, and we explain that the P in that project is determined by the signal and noise properties in the system.

- Students perform the following project relating to this application.
  - Write a simulation program for this system, and compare the simulated bit error rate with the analytical result.
  - Modify the program to simulate a system that transmits each signal value s three consecutive times, and assume the receiver selects the majority value as the transmitted signal. What is the probability of a bit error now? Please compare analytical results with simulation results.

6. OTHER PROJECTS

The digital communication receiver project leads us to the important topics of random variables, probability density functions, expectation, Gaussian distribution, and transformations of random variables. Projects are developed for these topics from application areas that include signal processing, noise in resistors and transistors, quantization errors in analog-to-digital converters, computer-communication networks, image coding and processing, digital systems, and communication systems.

7. COURSE EVALUATION

The course was evaluated two ways. First, students were asked to describe their attitude toward the new approach to the course. Second, formal course evaluation forms were distributed at the end of the course. A sampling of student comments follows.

- "I thought that this class was an interesting and effective way to teach probability. ...When math is presented simultaneously with its application, it is far easier to conceptualize. I think that putting this material in the context of electrical engineering, I have developed a greater understanding of the material."

- "The most important thing I found from this course is that I learned more in less time than in the [mathematics course], and now I understand why this stuff works and can apply it."

- "I personally feel that this class was not as informative as probability in the mathematics department. I feel I learned more about Matlab commands than I did about probability and statistics."

- "I thought the method in which the class was taught was much more beneficial than a classic approach to probability. The class dealt with more realistic application of probability theory, especially in EE situations. Examples often help students understand and relate to the subject better ...The projects proposed real-life systems, applicable to everyday problems."

- "The projects provided useful experience working with Matlab to help us visualize the results. The approach of the course in which we completed several projects throughout the semester was a good idea because it gave us a chance to apply what we learned in class. ...I found the content to be interesting. I learned a lot about probability, but in a way that could specifically be applied to EE."

- "I liked how the course uses real problems and systems to explain probability theory. They made it easier to learn, and were more EE relevant."

- "I feel more relaxed and at-home with the EE instructors, and feel that we already have a rapport which helped make this course so much better."

- "...this course is very well arranged. The theory of probability is illustrated and a lot of practical experience is included. The final project is very good way to let the students to have an opportunity to practice what they learn from this course is very useful."

- "I think the material covered in this course was well chosen. I wish that the projects were broken up into smaller homework assignments. That way they would have been easier to tackle."

- "...it ended up turning into a Matlab course. ...I feel that there should be more emphasis on probability and statistics than this course gave."

The formal course evaluation included two questions regarding the course concepts and computer exercises:

1. I gained an excellent understanding of the concepts in this field.

2. The computer exercises were a valuable part of this course.

The scale was disagree strongly (DS), disagree (D), neutral/mixed (N), agree (A), and agree strongly (AS). The responses from 21 students are given in Table 2.

	Rating				
	DS	D	N	A	AS
Question 1	0%	10%	18%	48%	24%
Question 2	0%	5%	5%	24%	66%

Table 2: *Summary of responses from 21 students to two questions on the course evaluation.*

8. SUMMARY

This paper has described four projects in an undergraduate probability and statistics course. The projects illustrate the applications of probability and statistics to practical engineering projects and provide students with hands-on experience using a computer package. After each of the first five projects is completed, students are examined on the project and required to obtain a computer solution. Most of the applications reinforce and expand on concepts that students have studied earlier. We have purposefully avoided using examples involving dice, cards, and games of chance.

This is our first offering of this elective course using this approach. The students appear to be "enjoying" the course, and many have voiced their appreciation of the course and the use of engineering examples.

9. REFERENCES

[1] Carol Ash, *The Probability Tutoring Book*, IEEE Press, New York, 1993.

[2] Rodger E. Ziemer, *Elements of Engineering Probability*, Prentice Hall, New Jersey, 1997.

[3] The Mathworks, Inc., *The Student Edition of MATLAB*, Natick, Massachusetts, 1992.

[4] Maurice F. Aburdene, *Computer Simulation of Dynamic Systems*, Wm. C. Brown, Dubuque, Iowa, 1988.

[5] Hewlett-Packard, "The Fundamentals of Signal Analysis," Application Note 243.