AC 2012-3579: A NOVEL WEB-BASED SUPPORT TOOL FOR LEARNING RANDOM VARIABLES

Dr. Anahita Zarei, University of the Pacific

Dr. Jinzhu Gao, University of the Pacific

Mr. Jason Roy Ortiz

Mr. Alan Joe

A Novel Web Based Support Tool for Learning Random Variables

Abstract

In probability and statistics, a random variable is a function that assigns a number to each outcome of a random experiment. Random variables have various applications in different scientific and engineering fields including health-care, genetics, communication, engineering management, etc. There is an inherent complexity in learning random variables and their distribution models. We have identified two issues that contribute the most to the difficulty students experience: 1) the complexity of the mathematical logic behind the probability theory and 2) lack of motivation to attempt and explore more problems due partly to the static nature of textbook problems. Therefore, our objectives were to enhance students' understanding of random variables and to increase motivation for learning by developing an interactive web-based tool.

We have developed a novel e-learning module to assist students in exploring three types of random variables, namely Poisson, Exponential, and Erlang, that differs from the current available resources. We assessed the effectiveness of the module by measuring the change in cognitive and affective behavior of students. We utilized independent diagnostic measures, a homework assignment, a quiz and a question on the final that exclusively cover the three mentioned families of random variables to assess the changes in the cognitive behavior. The treatment group had a superior performance in all measures. The Mann-Whitney U-test analysis shows that the improved performance on the quiz and the final was statistically significant. We also developed a survey to evaluate the students' affective behavior by measuring their motivation for learning and their perceptions of effectiveness of the module. A majority of students (82%) enjoyed doing the web module problems more than the textbook problems. Students agree (91%) that they would explore the web module problems beyond what they are asked to and 86% feel that web module was more motivating than the textbook.

Introduction

Computers have changed every aspect of our lives including engineering education. The traditional form of teaching can be complemented by the use of interactive multimedia based on computers. These systems can improve student learning by providing a high-quality learning environment and subsequently increase the educator's productivity[1].

Students that enter college today have become accustomed to computer games, digital gadgets, and web browsers that allow them to switch content at will. They have shorter attention spans, lower tolerance for repetition, and are averse to static media such as textbooks[2]. This necessitates providing information in more dynamic, compelling, and interactive ways.

Cyberlearning is a welcome addition both to the classic educational process and for autonomous asynchronous study. It can incorporate dynamic displays, interactivity, online calculations, audio, and video. All these attributes can have a stimulating effect on students. At the same time, well presented material can encourage students to go beyond the existing content, providing a stepping stone for further research[2]. In the realm of engineering and mathematics computer simulations are utilized in a variety of contexts such as power electronics[3], fluid mechanics[4], kinematics and kinetics[5], and probability and statistics[6-9].

Probability theory is one of the well-established topics in mathematics with various applications in all the engineering fields. A probability model is a mathematical representation of a random phenomenon. Many physical systems can be modeled by similar random experiments. Each of these common systems is referred to as family of random variables[10].

In practical applications, certain families of random variables appear over and over again in various experiments. Three of these families are Poisson, Exponential and Erlang. The probability model of each family has the same mathematical form. Therefore, understanding the characteristics of each family will allow applying this knowledge to analyze specific practical applications in each category. It is also important to note that there is an underlying relationship between the three mentioned families of random variables. Understanding the link between Poisson, Exponential, and Erlang can simplify the approach in tackling problems.

There is an inherent complexity in learning random variables and their distribution models. It takes a lot of careful thought and practice to fully comprehend such concepts and their applications. While there is an abundance of textbooks directed at introducing random variables to engineering students, most use the same abstract methods for presenting the material, leaving the students overwhelmed by complexity and unmotivated. There are also a number of on-line resources available for enhancing students' understanding of these concepts. Most of these resources are in form of computer simulations. Computer simulations are based on computer programs that incorporate a mathematical model of a system allowing the user to specify the values of one or more parameters and to examine the resulting values of those parameters[11]. The random variable simulation tools allow users to visualize the Probability Distribution Function (PDF) of random variables by setting their parameters such as seen in references [6-9]. These applets are usually limited to graphical displays of PDFs and mostly serve as statistical calculators.

We have identified two issues that contribute the most to difficulty students experience while learning the subject: 1) complexity of the mathematical logic behind the probability theory and 2) lack of motivation to attempt and explore more problems due partly to the static nature of textbook problems. Therefore, our objectives were to enhance students' understanding of random variables, and create enthusiasm and increase motivation for learning by developing an interactive web-based tool.

Our innovative web-based tool is different than the current available resources in several respects. First, this tool focuses on the three mentioned families of random variables and carefully investigates the relationship between them through "word problems", an element missing in most of other tools. Our resource does not simply give a mathematical formula or a graphical display of PDFs, it rather provides a framework for students to dig deeper into a variety of problems and receive instantaneous feedback to point them in the right direction. Moreover, our tool has a gaming structure that is very appealing to the targeted users. We developed this tool through collaborations with students that have taken probability courses while constantly reaching out for their feedback in order to improve this tool. Lastly, this tool is user friendly and straight forward to use. This is an important consideration since the complexity of some tools often results in the students losing interest in the activity.

This paper will proceed as follows. First we describe the design and different components of the web module and walk the reader through an example. Then we discuss the implementation details and finally describe the evaluation techniques.

Design

The web module consists of three main components: Workout, Shop, and Contest (figure 1). The general idea of the module is that players start by working out in the gym, gain the necessary knowledge and accumulate enough points to shop for items to power-up for the contest. Players then head to the competition where they will face the more complex questions that will determine the final outcome of the game.
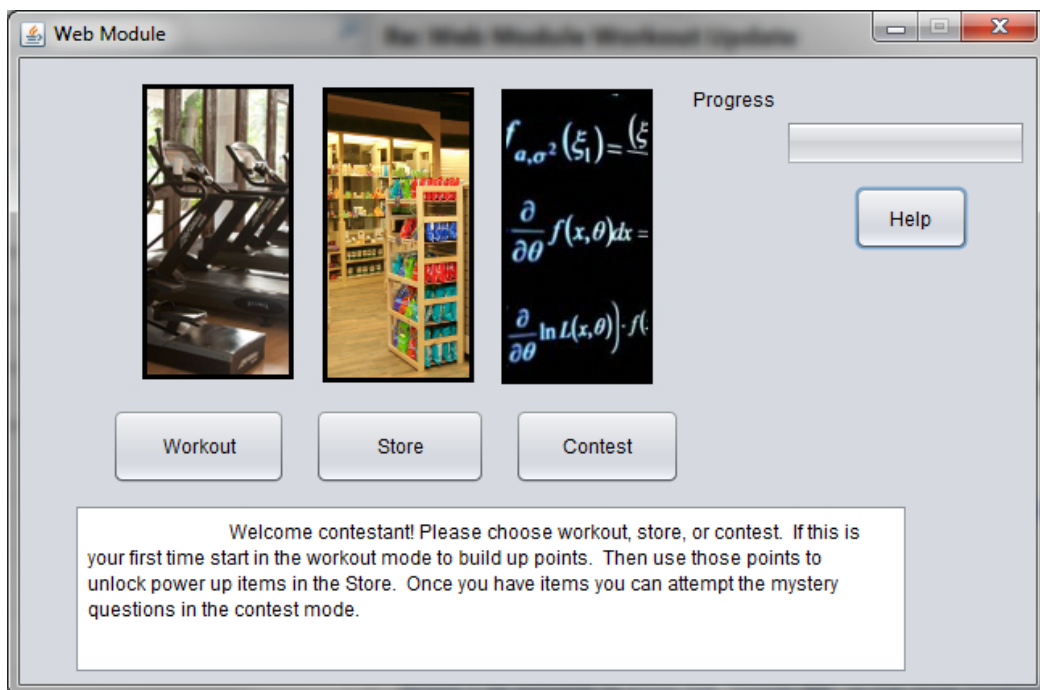

Figure 1. Main Window

The Workout mode provides students with a framework to test their knowledge on Poisson, Exponential, and Erlang random variables. They carry out the mathematical analysis here and explore different problems to observe the link between families of the three random variables. The design of the Workout mode is similar to the well-known TV game, Jeopardy (figure 2). There are five questions in each category of Poisson, Exponential, and Erlang, ranging from 100 to 500 points. The low level problems are single concept introductory questions focusing on the mathematical model of each distribution. The higher level problems not only require the mathematical knowledge, but also require the user to utilize an understating of material to differentiate between the three distributions. These problems were carefully selected to focus on common issues and misconceptions that students usually encounter.
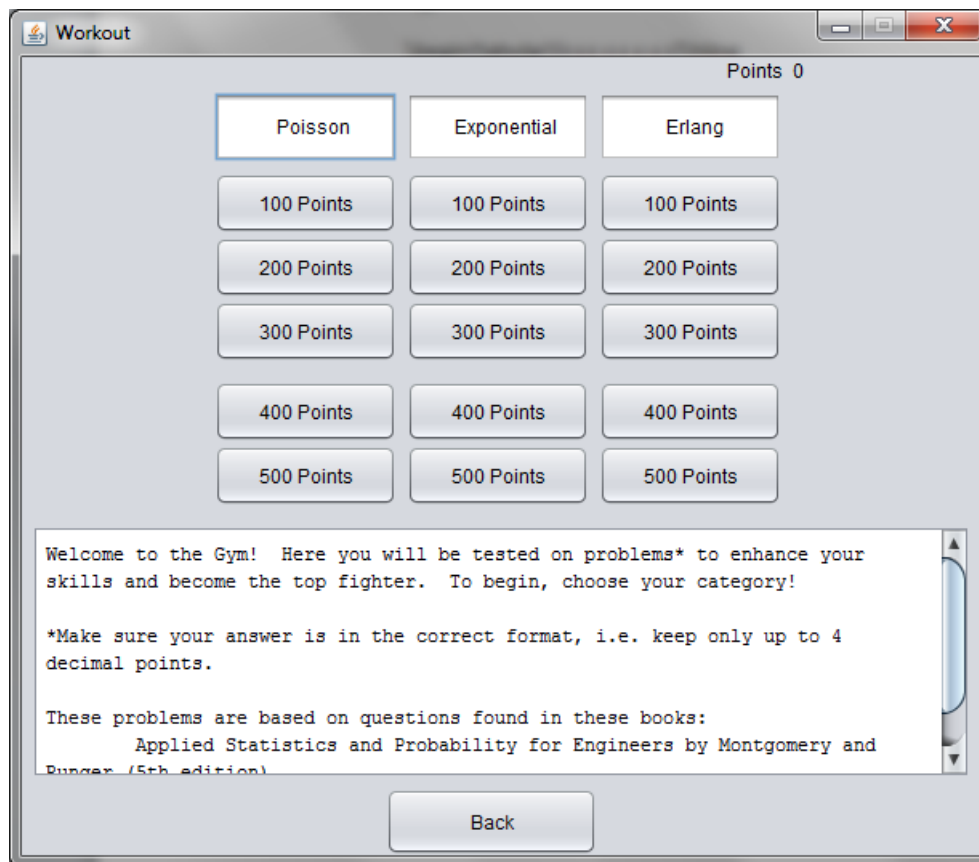

Figure 2. Workout Mode

Shop (figure 3) and Contest (figure 4) modes provide the students with an amusing experience, giving them the opportunity to shop for various items that empower them against a computer opponent. Each item in the shop unlocks at a certain point value (table 1). Users will be able to obtain stronger items as they gain more points through the Workout mode. In this segment, the number of problems solved correlates directly with the amount of credits available to the player. Therefore, the correct answers ultimately result in greater strength that is reflected in the Progress bar (figure 1), providing the player access to a larger selection of moves for the contest.

Power-up Item	Minimum points required
Fruit	500
Energy Drink	1000
Protein Shake	2000
Power Pasta	3500
Secret Power	4200

Table 1. Required points for power-up items in the Shop mode
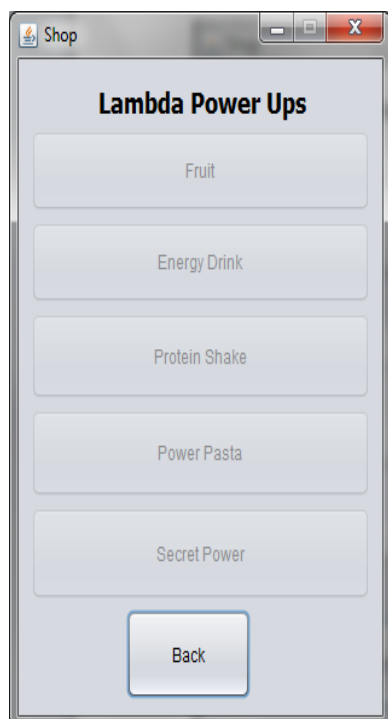

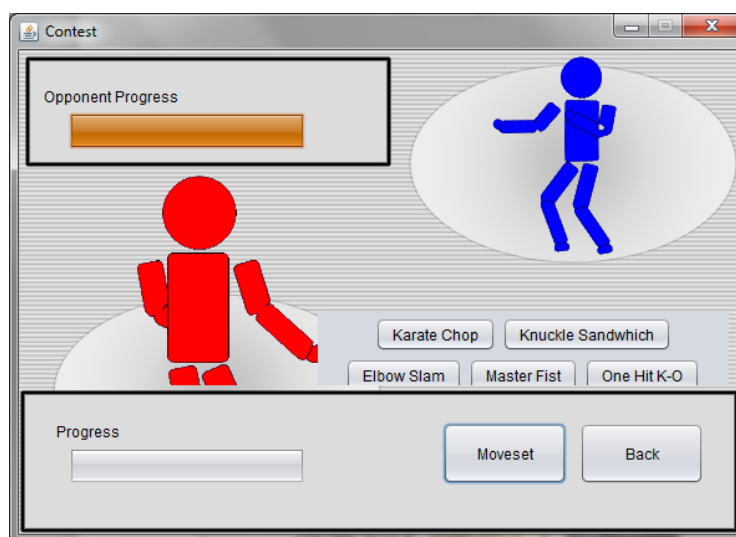Figure 3. Shop Mode   Figure 4. Contest Mode

Contest is the final stage of the game where students' knowledge is put to the test. Here users can battle the computer in a match to see who comes out victorious. Exercises in the Workout mode prepare students for answering more conceptual questions in the Contest mode. With the power-up items students purchase moves to improve their competiveness. Each move has a flash video associated with it depicting the character in action. Upon answering the questions associated with each technique, players can land the move. The success of the move is itself tied to a random outcome, but directly proportional to correctness of their answer to the corresponding question. A correct answer rewards the player with a high chance of landing the move successfully and an up to 16% increase in power. In contrast, a false answer can result in up to 10% decrease in power. Players will have a second chance of answering the question again after reading the feedback and landing the move. Ultimately, the result of the game will be determined by the players' power level in accordance with table 2.

Power Level	Medal
90%-100%	Gold
80%-90%	Silver
70%-80%	Bronze
Less than 70%	Not placed

Table 2. Winning scheme of the web module

Proposed questions in the Contest mode are at a higher difficulty level than the Workout mode. Unlike the gym, the category of the questions is not disclosed to users. In order to successfully solve the problems in this segment and win, students must master the concepts presented in the gym and retain a full grasp of the interweaving relationship between the three mentioned families of random variables.

For clarification, a detailed round of competition is demonstrated through a hypothetical user, Jenny. Figure 1 shows the window displayed upon her successful log-in. The Help button provides a friendly introduction to the module. It describes the features involved with the goal of motivating her to explore all attributes of the module. The Progress bar represents her power (currently at zero level), that can be unlocked in the store to allow her to use more advanced techniques for the contest.

Jenny starts by practicing in the gym. Upon clicking Poisson 100, figure 5 is displayed. If she answers the question incorrectly feedback will be provided to point her to the right direction. Figures 6 and 7 show the first and second feedback windows for Poisson 100. The feedback system provides a series of hints that will appear gradually on the screen without providing the correct answer at once. The system will give her several opportunities for modifying her solution and gaining points.
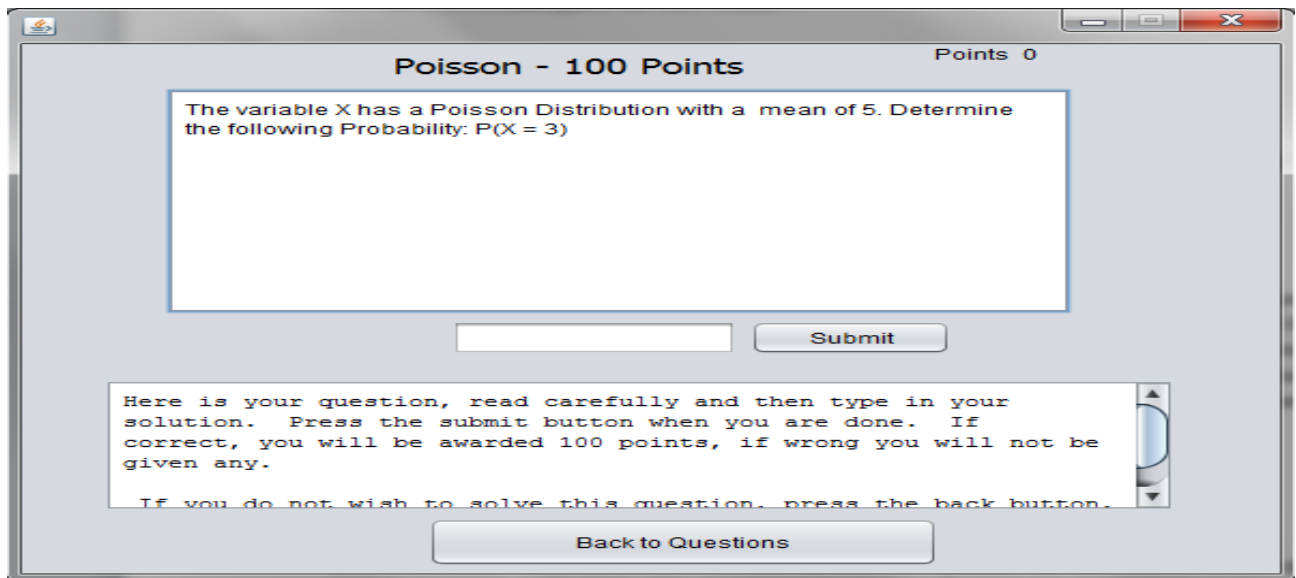

Figure 5. A sample 100-point Poisson question

If she answers the question correctly, 100 points will be awarded. If the question is not answered correctly, she can still proceed to the next question but will not be awarded any points. Let's assume that she has accumulated 500 points in the gym. Now she can enter the store and purchase a fruit item that will increase her power level to 5%. This power level provides access to the Contest mode and allows her to apply the first technique, karate chop. Upon clicking on karate chop icon, she will be presented with a high level question. In the scenario that she does not solve the question correctly the first time, she loses this move to the computer and Progress bar is drained by 2%. After reading the feedback, if she answers the question correctly, she will land the move, the opponent will be visibly affected by the move, and she will gain back 3% of the power. On the other hand, if she decides to quit at this point with only 6% progress, she will not be ranked. Never the less, since her work is saved, she retains the opportunity to return to the game at any time, work on rest of the questions and win a medal.
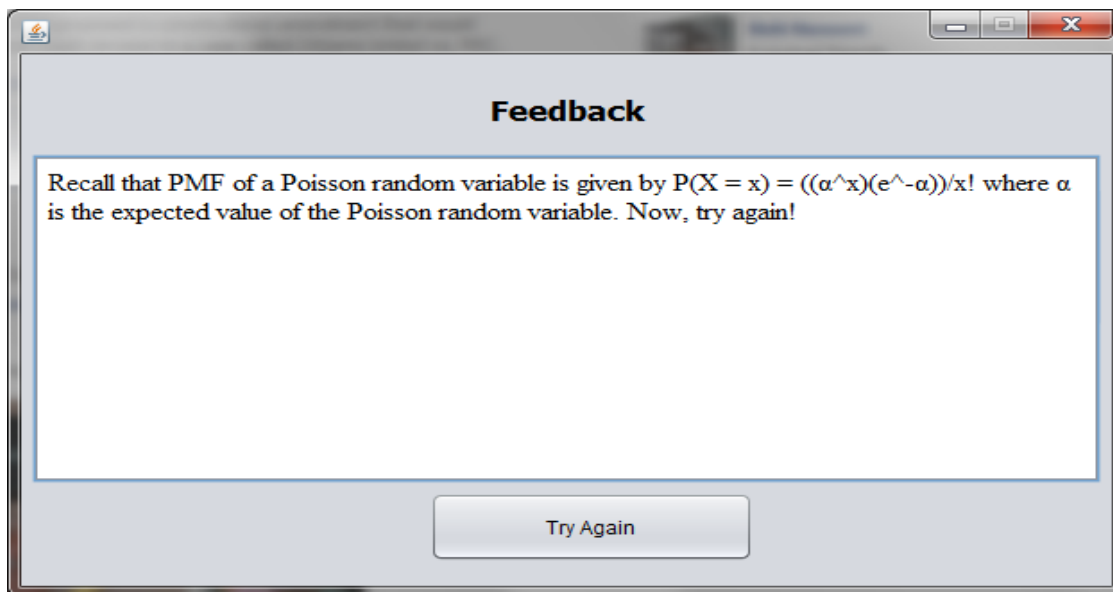

Feedback

Recall that PMF of a Poisson random variable is given by $P(X = x) = ((\alpha^{\wedge}x)(e^{\wedge}-\alpha))/x!$ where α is the expected value of the Poisson random variable. Now, try again!

Try Again

Figure 6. First feedback to the sample 100-point Poisson question

Implementation

We created the animations in the Contest window using the Adobe Flash. A clip associated with each move is played when the user or the opponent lands the move. Java was our main programming language due to its universal platform, network and operating system transparency, and easy-to-use user interface[3].

The solution to a statistical problem usually involves parsing and interpreting a string of complicated mathematical expressions. In addition, to stop students from sharing answers and skipping out on the learning from the module, it is preferred to be able to generate random numbers for the variables in each problem. In order to facilitate Java implementation of these

complicated expressions, we broke the solution into simple expressions. We then implemented those expressions in a functional style to reuse it for multiple problems.
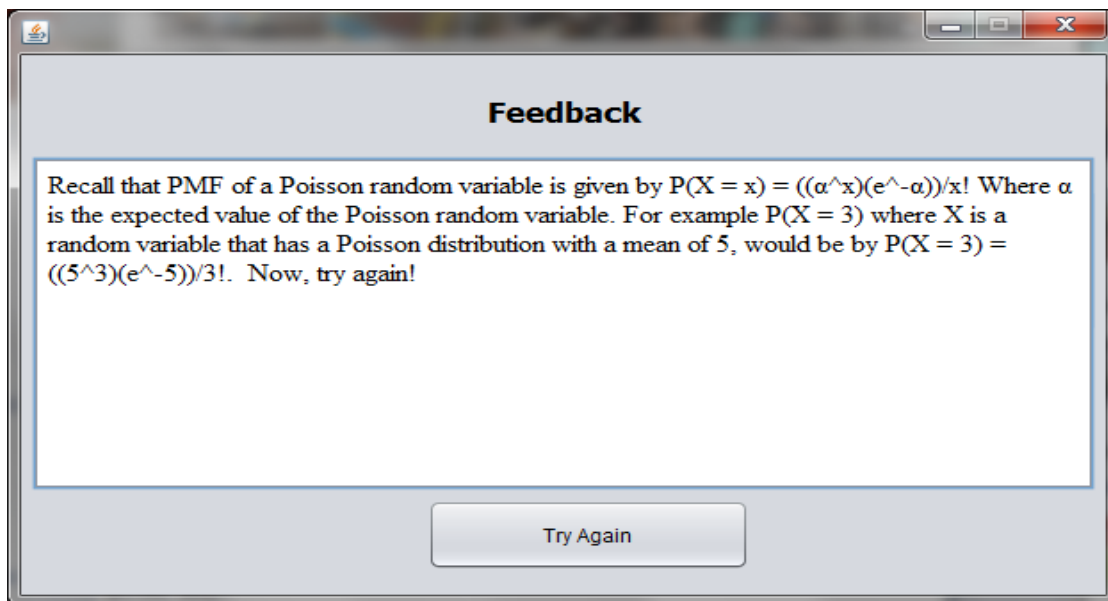

Figure 7. Second feedback to the sample 100-point Poisson question

For instance, calculating an expression such as $P(X \leq 15) = \int\limits_{0}^{15} \dfrac{0.2^2 \, x^3 e^{-0.2x}}{2!} dx$ requires the evaluation of simpler functions such as exponentials, factorials and integrands. This technique will also facilitate accepting random numbers and generating a solution. This feature allows students to return to a problem they have previously completed and practice it again with new parameters.

This same method of development was used for each of the equations for this application. A solution checker was also developed that would compare the user's answer with the solution and accept it as a correct answer if the difference is within a small error margin due to rounding. In order for students to save their progress while working through the problems, a MySQL database was employed to house the information.

Evaluation

In the preparation of our evaluation plan, we considered the guidelines of several program directors in NSF's Division of Undergraduate Education described in [12]. The goal of evaluation plan was to measure the changes in cognitive and affective behavior. Measuring the changes in cognitive behavior is in accordance with the project's first objective of increasing students' understanding of concept of random variables. Evaluating the changes in affective behavior is in line with the project's second objective of further motivation for learning random variables.

We assessed changes in cognitive behavior through the following protocol. We identified a homework assignment, a quiz, and a question on the final that exclusively the three mentioned families of random variables and were given to students in ENGR 250 in 2010 (control group) and 2011 (treatment group). ENGR 250 is a calculus-based course taken typically by engineering graduate students who have never taken a probability class before. This course was taught by the same instructor in both 2010 and 2011. Furthermore, the classes taught were designed to resemble one another as closely as possible in all characteristics relevant to the outcome measures of interest. Such similarities included the material covered, class discourse, and the time frame for assignments, quizzes, and exams. The only disparity was that students who took the course in 2011 worked on the web module prior to working on the homework assignment, the quiz, and the final.

Table 3 shows the quantitative performance of the two groups of students. The treatment group students did better than the control group in all three measures. The p-value from the Mann-Whitney U-test indicates that the superior performance of the treatment group on the quiz and the 10-point question on the final is statistically significant.

Group	High	Low	Median	Mean	Stand. Dev.	P value
Homework						
Control	100	47	86	84	18	
Treatment	100	87	95	94	5	0.28
Quiz						
Control	100	40	81.5	80	20	
Treatment	100	67	97	93	9	0.03
Question on Final						
Control	10	6	7.5	8	1	
Treatment	10	7	9	9	1	0.04

Table 3. Quantitative results measuring students' learning in the two groups

We developed a survey to evaluate the students' affective behavior, i.e. measuring their motivation for learning. This survey gave an opportunity to the students to assess the new module. Table 4 illustrates the survey questions and the ranking scheme.

	1(Strongly Disagree)	2 (Disagree)	3 (Neutral)	4 (Agree)	5(Strongly Agree)
Q1: I enjoyed doing the web module problems more than textbook problems					
Q2: Web module was more motivating than textbook					
Q3: I would explore the web module problems beyond what I am asked					

Table 4. Survey questions for module assessment

22 students participated in the survey. Table 5 illustrates the details of the survey results. Mean values ranged from 4.00 to 4.41, all in the areas of Agree (A) and Strongly Agree (SA). Majority of students (82%) enjoyed doing the web module problems more than the textbook problems. Students agree (91%) that they would explore the web module problems beyond what they are asked to and 86% feel that web module was more motivating than the textbook.

	SD	D	N	A	SA	Mean	Stand. Dev.
Q1	1	1	2	12	6	4.00	1.00
Q2	0	1	1	8	12	4.41	0.78
Q3	0	1	2	11	8	4.18	0.78

Table 5. Students' responses to the survey questions.

These finding are very promising and in accordance with our objectives. We plan to expand the web module to incorporate students' suggestions and needs. The gaming component has been a very attractive and motivating feature of the module. We plan to expand the Shop and the Contest to make the module more exciting. Further study in the subsequent semesters is required to fully explore how the web module will work in other courses and to evaluate it to a greater extent.

Conclusion

We have developed a novel interactive e-learning tool to assist students in exploring three types of random variables namely Poisson, Exponential, and Erlang. Our web-based interactive module enhances students' understanding of these concepts and increases their motivation for learning. The Workout mode addresses our first objective to increase students understanding of the concept of random variables. The Shop and the Contest modes address our second objective to increase students' motivation by incorporating gaming elements to the module.

Our interactive learning module can be beneficial in a variety of ways. It can serve as supplementary material enriching a lecture with a compelling gaming component. It can also be used as a complement to a homework assignment, presenting the material in a multimedia format, or serve as a tinkering environment, enabling exploratory, student-centered learning.

This module provides a flexible framework that can be adjusted to include different problems with various levels of complexity for undergraduate and graduate courses. This effective pedagogical tool can even be altered for use in other courses by simply changing the problems in the Workout and Contest modes.

The analysis of students' response to our survey questions reveals that they enjoyed exploring the module and found it more motivating than their textbook problems to investigate the concept of random variables.

This module is currently available on the University of Pacific website to users with valid university ID numbers. We are planning to make this resource available to the public upon administrative approval.

Bibliography

[1] W. Swart and Bengu, "A computer-aided, total quality approach to manufacturing education in engineering," IEEE Transactions on Education, vol. 39, no. 3, pp. 415–422.

[2] "nsf08204.pdf." [Online]. Available: http://www.nsf.gov/pubs/2008/nsf08204/nsf08204.pdf

[3] C. A. Canesin, F. A. S. Goncalves, and L. P. Sampaio, "Simulation tools for power electronics courses based on java technologies," IEEE Transactions on Education, vol. 53, no. 4, pp. 580–586, Nov. 2010.

[4] Fraser, D.M., R. Pillay, R.L. Tjatindi, and J.M. Case. "Enhancing the learning of fluid mechanics using computer simulations". Journal of Engineering Education 96 (4): 381–88.2007

[5] "ASEE12008_0084_paper.pdf. "[Online] Available: http://www.asee.org/documents/zones/zone1/2008/professional/ASEE12008_0084_paper.pdf

[6] Dinov, J. Sánchez, and N. Christou, "Pedagogical utilization and assessment of the statistic online computational resource in introductory probability and statistics courses."Computers & Education, 50, 284-300, 2008

[7] Online]. Available: http://www.math.csusb.edu/faculty/stanton/m262/intro_prob_models/intro_prob_models.html

[8] [Online]. Available: http://users.ece.gatech.edu/~gtz/java/

[9] [Online]. Available: http://nlvm.usu.edu/en/nav/frames_asid_305_g_3_t_5.html

[10] R. D. Yates and D. Goodman, Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers, 2nd ed. Wiley, May 2004.

[11] Smith, P.R., and D. Pollard, "The Role of Computer Simulations in Engineering Education," Computer Education, Vol. 10, No. 3, pp. 335–340, 1986

[12] R. Pimmel, B. Anderegg, S. Burkett, B. Watford, and C. Della-Piana, "Evaluation of educational development projects," 2006.