

# Use of MATLAB in Teaching the Fundamentals of Random Variables

**Gordana Jovanovic-Dolecek and Alfonso Fernandez-Vazquez**

Department of Electronics  
P.O. Box 51 and 216, 72000  
INAOE, Puebla Mexico  
{gordana,afernan}@inaoep.mx

**Abstract:** The way engineering is being taught has changed in recent years with the introduction of commercial and educational software thus allowing a better understanding of the subject matter and increased teaching efficiency. This paper presents the educational software package designed to work in PCs and provides an attractive way to introduce the fundamentals of random variables. The software can be used as a complement to theoretical classes or alone as a self-study tool. The package was implemented in MATLAB and uses MATLAB<sup>TM</sup> tool “makeshow” which allows the students to create interactive slideshows without building their own graphic interface. The software is easy to use and is in accordance with the established standards, mouse operations, icon menus, etc. The demos include Probability Density Function (PDF), and Cumulative Distribution Function (CDF), normal, lognormal, exponential, Rayleigh and Rician random variables. The package also demonstrates how to generate the desired random sequence, how to examine its characteristics and how to estimate its PDF and CDF.

**Categories and Subject Descriptors:** G.3 [Probability and statistics]: Distribution functions

**General terms:** Experimentation, Theory

**Keywords:** Educational software, MATLAB, Probability theory

## 1. INTRODUCTION

Many of the engineering lecture courses are taught using a traditional method i.e. only with the support of a board or transparencies. The major disadvantages of traditional teaching methods are that students usually have to choose between taking notes and listening to the lecturer. As a result, most of the learning in traditional teaching is individual, since the students in the classrooms are taking notes or listening instead of participating [6]. Consequently, many students in conventional classrooms develop little confidence in their own ability to learn [14].

Demos are one of the important tools in the field of engineering education. They help students realize how theory and systems work in reality [5]. Students learn better, remember longer and are better able to identify the appropriate concepts to solve new problems when they learn by addressing concrete problems and actively participate in exploration and pursuit of knowledge [14].

The availability of personal computers (PCs), their increasing calculation power, and their enhanced graphical possibilities enable teachers to take the advantage of the pedagogical possibilities afforded by new technologies [20]. Computer-aided learning has become extremely popular and its use in classroom can be very helpful by adding more analytical capabilities in all engineering areas [4]. It can be applied in the aspects of teaching, learning,

validation, and research in engineering education [5]. Besides, it is beneficial especially in terms of saving time and efforts for both teaching and learning in the educational process [5].

As a result computer-aided learning has become an important educational research activity in various engineering disciplines and there has been a growing interest in the development of educational software in all areas of study [4] and many computer packages have been developed to assist learning [2][23].

Our experience at the National Institute INAOE [9-11], shows that the development and usage of software tools represent an effective teaching approach and increase students' learning. Such software tools must be reliable, student-friendly and they must also preserve portability among different software platforms. In pursuance of this goal we developed the educational software package to support the theory provided in basic courses in which the fundamentals of random signals are taught.

It is known that random variable is generally considered as one of the most abstract and conceptually difficult areas in the engineering education and teaching of random variables is one of the subjects that requires more time for its understanding. The use of computers gives students the visual and intuitive representation of the

random variables which had traditionally been stated in terms of abstract mathematical description.

The package was implemented in MATLAB. We choose MATLAB because MATLAB along with the accompanying toolboxes is the tool of choice for most educational and research purposes, [1,12,16,18,19]. It provides powerful computation and advanced visualization tools and is also available on a number of hardware platforms.

The advantage of our approach is that the student does not need MATLAB or any other programming language experience.

We used the MATLAB tool *makeShow* which allows the student to create her/his own interactive slideshows without building her/his own graphic interface [15]. There are two windows in each slide. We used the upper windows for specific explanations, like in Figure 1.a, or for graphics, like in Figure 1.b. The bottom window was used for general explanations like in Figure 1.a, or for the explanations of the upper graphics, like in Figure 1.b. The order of the slide and the command buttons are on the right side of each slide. (See Figures 1.a and 1.b).

The student can change the slides automatically using the option *AutoPlay*. In this case the student may use the button *Stop* to stop the presentation and later hit *Continue* to continue it. The student can also choose the option to change slides manually with the mouse, by clicking *Next* to go forward or *Prev*, to go backward. The buttons *Reset* and *Close* are used to reset the demos and exit the program, respectively.

In the next section we briefly describe the package and give some examples of demos.

## 2. DESCRIPTION OF PACKAGE

As shown in Figure 2, the content of the package is given in the Menu. In the following we briefly describe the options shown in the Menu.

The random variable  $X$  with the range of values  $x$  is described by the cumulative distribution function (CDF)

$$F_X(x) = P(X \leq x) \quad (1)$$

where  $P$  is the probability and  $x$  is defined for every  $x$  from  $-\infty$  to  $\infty$ .

The probability density function (PDF) of  $X$  is defined as

$$f_X(x) = \frac{dF_X(x)}{dx} = \frac{P(x < X \leq x + dx)}{dx} \quad (2)$$

The goal of the demo *Density and Distribution* is to show how to apply the relations (1) and (2) to a given random sequence to estimate its CDF and PDF using a histogram and the law of large numbers and also to illustrate the relationship between the PDF and the CDF [7,8].

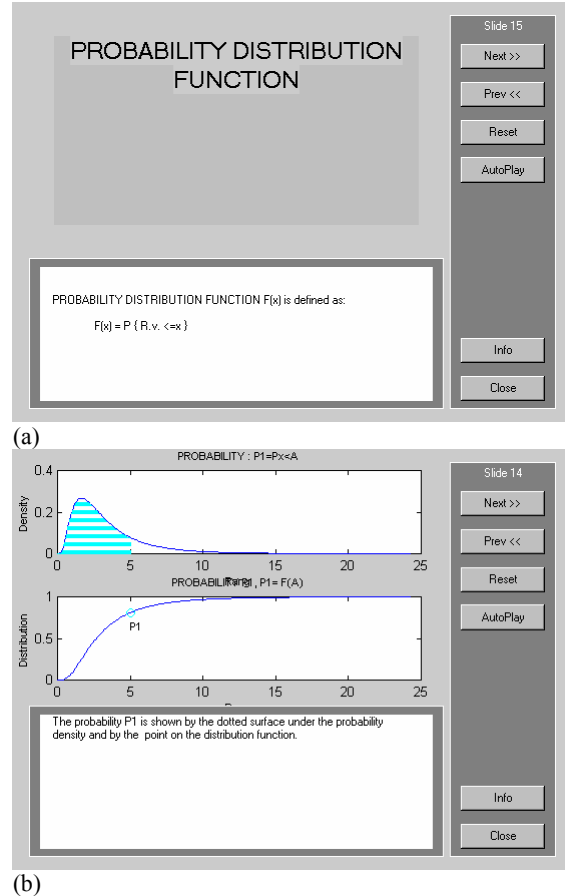


Figure 1. Different slide examples.

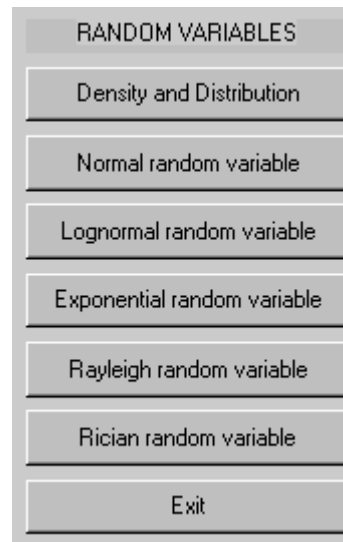


Figure 2. Menu.

The rang of the given random variable is divided into equidistant intervals called cells. The histogram presents how many values of the given random variable belong to each cell. The probability that the random variable is inside the cell is obtained as the ratio of the cell histogram

value and the total number of the values of the random variable  $N$ . (Numerator in Eq. (2)). According to Eq. (2) the PDF estimation is obtained as the ratio of the probability and the length of the cell. Figure 3 shows  $N=500$  samples of the uniform random variable in the interval  $[6,14]$ , its histogram with 10 cells, probability estimation and the corresponding PDF estimation.

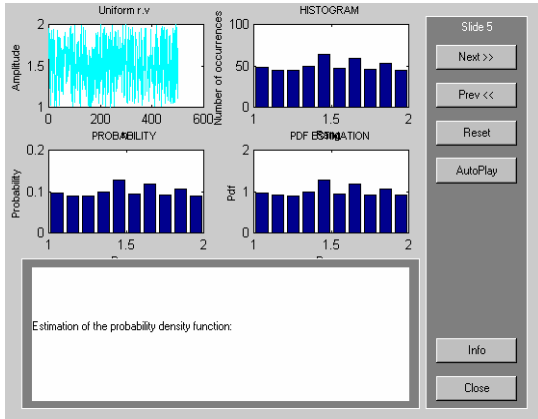


Figure 3. Illustration of the Density and Distribution demo.

The demo is useful for understanding what the probability density function is and what the cumulative distribution function of a given sequence is. This demo also shows the relationship between the histogram, probability for each cells and the PDF estimation. It also illustrates the relationship between the PDF and the distribution.

In this demo the student can change the type of the random variable, the length of the sequence and the number of cells for the histogram, as well as the density and distribution estimation and therefore understand the relationship between those parameters and the quality of the estimation.

The normal (or Gaussian) random variable has played a significant role in the study of random phenomena in nature [7], because many naturally occurring random phenomena are approximately normal. In addition, the sum of a large number of independent random variables, under certain conditions, can be approximated by a normal random variable (Central-limit theorem).

A random variable  $X$  is called normal if its PDF is of the form

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/(2\sigma^2)} \quad (3)$$

where  $m$  is the mean value and  $\sigma^2$  is the variance. The corresponding cumulative distribution is

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-(x-m)^2/(2\sigma^2)} dx \quad (4)$$

The goal of the demo Normal random variable is to illustrate the main properties of the normal random variable, which include:

- The rule of  $3\sigma$  which says that 99% of all values of the normal random variable are in the interval  $m \pm 3\sigma$ .
- How the shape of the PDF changes with the parameters  $m$  and  $\sigma^2$ .
- How the probability that the random variable takes values in the given interval related to the corresponding surface under the PDF and the corresponding points of the cumulative distribution is.
- The fact that the sum of the normal random variable is also normal.
- The fact that the linear transformation of the normal random variable is also normal.
- Central limit theorem. (Under a wide range of conditions the normal random variable can be used to approximate the sum of a large number of independent random variables).

The  $N$  values of the Normal random variable with the variance  $\sigma^2=1$  and the mean value  $m=0$  is generated using the MATLAB function `randn`.

$$n = \text{randn}(1, N) \quad (5)$$

In order to generate normal random variable  $n_1$  with the desired mean value  $m$  and the desired variance  $\sigma^2$ , the following transformation of the normal random variable  $n$ , generated by the MATLAB command `randn` is used [8].

$$n_1 = \sigma * \text{randn}(1, N) + m \quad (6)$$

Figure 4 shows different  $N=500$  length random sequences for the same mean values and for different variances.

Lognormal variable  $X$  can be generated using the exponential transformation of the corresponding normal random variable  $n$  [7,8,17].

$$X = e^n \quad (7)$$

The corresponding MATLAB function is

$$X = \exp(\text{randn}(1, N)). \quad (8)$$

The PDF of the lognormal variable is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_n x} e^{-(\ln x - m_n)^2 / 2\sigma_n^2} \quad (9)$$

where  $m_n$  and  $\sigma_n^2$  are the mean value and the variance of the normal random variable  $n$ . The corresponding CDF is

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma_n} \int_0^x \frac{e^{-(\ln x - m_n)^2 / 2\sigma_n^2}}{x} dx \quad (10)$$

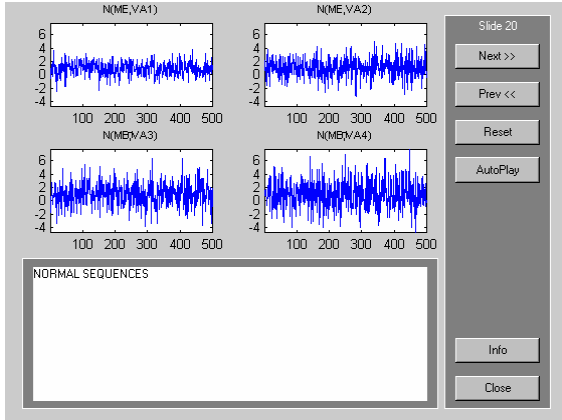


Figure 4. Illustration of the Normal random variable demo.

The goal of the demo **Lognormal variable** is to show how to generate a lognormal sequence using (7) (See Figure 5) and to illustrate how the characteristics of the variable changes with the parameters of the corresponding normal random variable. The student can choose the number of the desired data  $N$  and the desired values for  $m_n$  and  $\sigma_n^2$ . Since the PDF and CDF are estimated for the generated lognormal sequence, the student can also observe how the shapes of the PDF and CDF change with the parameters of the normal variable parameters. The student can also specify the range of values of this random variable and observe how is the probability that the lognormal random variable lies in this interval related to the corresponding PDF and CDF. Using this demo, it can also be verified that the sum of lognormal variables can be approximated by a lognormal variable.

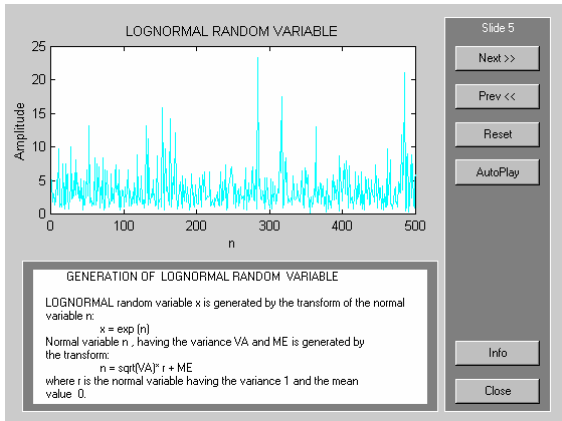


Figure 5. Illustration of the Lognormal random variable demo.

Exponential random variable arises in modeling of the time between occurrence of events, and in the modeling of the lifetime of devices and systems [13,17]. The parameter  $\lambda$  is the rate at which events occur, so that the probability of an event occurring by time  $x$  increases as the rate  $\lambda$  increases.

The exponential variable  $X$  with the parameter  $\lambda$  has PDF

$$f_X(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0 \end{cases} \quad (11)$$

Its CDF is

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases} \quad (12)$$

$N$  values of the uniform random variable  $U$  are generated using MATLAB function `rand`.

$$U = \text{rand}(1, N) \quad (14)$$

Using (13) and (14), the following MATLAB command can be used for the generation of  $N$  values of the exponential random variable.

$$X = -\log(1 - \text{rand}(1, N)) / \lambda \quad (15)$$

An exponential random variable  $X$  with the parameter  $\lambda$ , can be obtained by the following transformation of the uniform random variable  $U$  in the interval  $[0,1]$ ,

$$X = -\frac{1}{\lambda} * \ln(1 - U) \quad (13)$$

The exponential random variable is the only continuous variable with the memoryless property [13,17],

$$\begin{aligned} P\{X > x_0 + x | X > x_0\} &= \frac{P\{(X > x_0 + x) \cap (X > x_0)\}}{P\{X > x_0\}} \\ &= \frac{P\{X > x_0 + x\}}{P\{X > x_0\}} = \frac{e^{-\lambda(x_0+x)}}{e^{-\lambda x_0}} \\ &= e^{-\lambda x} = P\{X > x\} \end{aligned} \quad (16)$$

This result demonstrates that the probability of waiting at least an additional  $x$  second does not depend on how long one has already been waiting [13].

The demo **Exponential random variable** shows how to generate an exponential random sequence using (13) and how to approximate its PDF and CDF. The student also can investigate what is the probability that the exponential random variable is inside of the chosen interval or that it is less than the chosen value. The probability values are depicted as the area under the PDF and also as the corresponding points on the distribution function. The memoryless property (16) is also illustrated in the demo.

As an example, Figure 6 shows the estimations and the exact shapes of the exponential density and distribution. A Rayleigh random variable  $X$  can be generated using the transformation of two normal random variables  $n_1$  and  $n_2$  with  $m_1=m_2=0$  and  $\sigma_1^2 = \sigma_2^2 = \sigma^2$  [8],

$$X = \sqrt{n_1^2 + n_2^2} \quad (17)$$

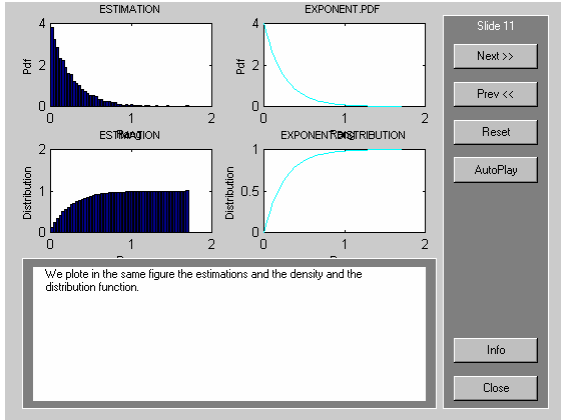


Figure 6. Demo of Exponential random variable.

A MATLAB command that generates  $X$  is

$$X = \text{sqrt}(\sigma * \text{randn}(1, N).^2 + \sigma * \text{randn}(1, N).^2). \quad (18)$$

The random variable  $X$  has the following PDF

$$f_X(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} & x \geq 0 \end{cases} \quad (19)$$

and CDF

$$F_X(x) = 1 - e^{-x^2/2\sigma^2} \quad (20)$$

The demo Rayleigh random variable shows the generation of the Rayleigh random sequence, and the estimation of the density and the distribution functions. Figure 7 illustrates the estimation of the PDF.

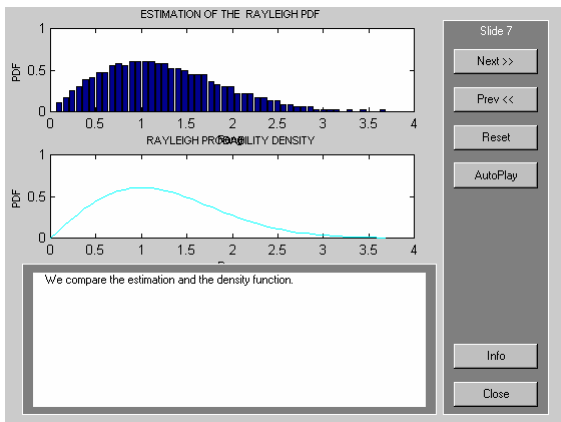


Figure 7. Illustration of the Rayleigh random variable demo.

The student can change the parameters of the corresponding normal variables and observe the consecutive changes of the Rayleigh sequence. It is also possible to investigate the probability that the Raleigh random variable is in the chosen interval or less than the chosen value and relate the probabilities with the PDF and Density.

The last demo Rician random variable describes Rician random variable  $X$ , its generation, and estimation of the density and distribution function. This variable can be generated using the following transformation of the normal variables  $n_1$  and  $n_2$ , [8]

$$X = \sqrt{(A + n_1)^2 + n_2^2} \quad (21)$$

where  $A$  is a constant. The corresponding MATLAB command is

$$X = \text{sqrt}(A + \text{randn}(1, N).^2 + \text{randn}(1, N).^2) \quad (22)$$

This random variable has the PDF

$$f_X(x) = \frac{x}{\sigma^2} e^{-(x^2 + A^2)/2\sigma^2} I_0\left(\frac{x A}{\sigma^2}\right) \quad (23)$$

where  $I_0(x)$  is the modified Bessel function

$$I_0(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{2^{2n} (n!)^2} \quad (24)$$

The demo shows how the shape of the density function changes with the parameters  $A$  and  $\sigma$ . It also shows the limiting cases in which the Rician random variable approaches a normal or a Rayleigh random variable. As an illustration, the last case is shown in Figure 8.

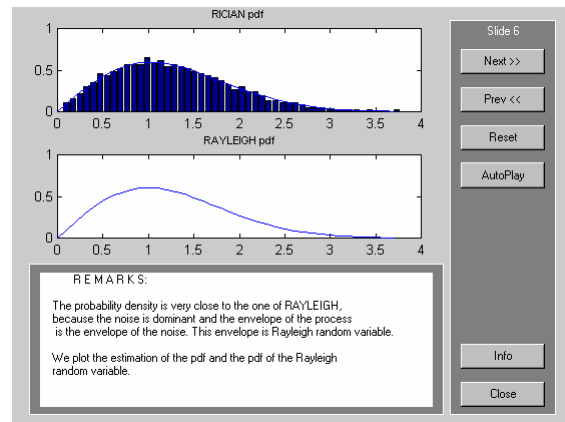


Figure 8. Illustration of Rician random variable demo.

### 3. CONCLUSIONS

We created this educational software package with the idea of supporting a course on the fundamentals of random signals. It encompasses the main aspects of the teaching process: motivation, experimentation, theoretical study, and evaluation.

The software can be used as a complement to theoretical classes or alone as a self-study tool. The package has all features we consider important for any educational software such as personalized and active learning process, repeatability, capacity to motivate students, interactivity, versatility, easy to use etc.



The demos include the estimation of the density and distribution functions in general, as well as for specific random variables such as normal, lognormal, exponential, Rayleigh and Rician. The relationship between the density and distribution and the probability is demonstrated for each variable. Some specific characteristics of these variables are also illustrated.

The package has been conceived as a complementary tool enabling students to get a more intuitive understanding of random variables.

The package is implemented in MATLAB and uses graphical tool makeshow, available in MATLAB. This

tool allows the student to create her/his own interactive slideshows without building a graphic interface. The package is composed of slides that contain the plots of the corresponding signals and a short explanation. Each demo can run automatically or manually, and each slide can be repeated many times. In that way the student can compare the actual result with any previous one.

Our experience at INAOE has shown that such an education tool allows for a better understanding of this important subject and increases teaching efficiency.

## REFERENCES

- [1] Cathey J.J., A MATLAB-based Graphical Technique for Amortization Study of Adjustable Speed Drivers, *IEEE Trans. Education*, 2002, 45(2), 177-181.
- [2] Ebel W.J., and Younan N., Counting on Computers in DSP Education, *IEEE Signal Processing Magazine*, 1995, 12(11), 38-43.
- [3] Fernandes F.A.N. and Rodrigues S., Educational Software for Gas-Solid and Liquid-Solid Separation Equipment, *Computer Applications in Engineering Education*, 2003, 11(4), 226-233.
- [4] Fernandez A.J. and Sanchez J.M. CGRAPHIC: Educational Software for Learning the Foundations of Programming, *Computer Applications in Engineering Education*, 204, 11(4), 167-179.
- [5] Fu T.T., Applications of Computer Simulation in Mechanism Teaching, *Computer Applications in Engineering Education*, 2003, 11(3), 156-165.
- [6] Fuller D.A. and Moreno A.F., Experimenting With a Computer-Mediated Collaborative Interaction Model to Support Engineering Courses, *Computer Applications in Engineering Education*, 2004, 12(3), 175-187.
- [7] Hsu H.P. *Probability Random Variables and Random Processes*. (McGraw-Hill) 1997.
- [8] Jovanovic Dolecek G., *Random Variables and Processes in Telecommunications (in Serbian)*, Svjetlost, Sarajevo, 1987.
- [9] Jovanovic-Dolecek G., RANDEMO: Educational Software for Random Signal Analysis, *Computer Applications in Engineering Education*, 1997 5(2), 93-99.
- [10] Jovanovic-Dolecek G. and Champac V., CGTDEMO- Educational Software for the Central Limit Theorem, *ACM Press, SIGCSE Bulletin*, 2000 32(2), 46-48.
- [11] Jovanovic-Dolecek G., Madrigal J.M., Ibarra-Manzano O., MuDSPDEMO- Demo Package for Multirate Digital Signal Processing, *Computer Applications in Engineering Education*, 2000, 8(2), 132-138.
- [12] Kim A.S., Park C., and Park S.H., Development of web-based Engineering Numerical Software (WENS) Using MATLAB: Applications to Linear Algebra, *Computer Applications in Engineering Education*, 2003, 11(2), 67-75.
- [13] Leon-Garcia A., *Probability and Random Processes for Electrical Engineering*, (Addison-Wesley Publishing Company, Inc), 1994.
- [14] Li S.G. and Lie Q., Interactive Groundwater (IGW): An Innovative Digital Laboratory for Groundwater Education and Research, *Computer Applications in Engineering Education*, 2004, 11(4), 179-203.
- [15] MathWorks *Using MATLAB*, 2007.
- [16] Orsak G. C., Etter D.M., Collaborative Signal Processing Education Using the Internet and MATLAB. *IEEE Signal Processing Magazine*, 1995, 12(6), 23-32.
- [17] Peebles P. *Probability random Variables and Random Signal Principles*, (McGraw-Hill), 2000.
- [18] Pires V.F. and Silva J.F.A., Teaching Nonlinear Modeling, Simulation and Control of Electronic Power Converters Using MATLAB/SIMULINK, *IEEE Trans. Education*, 2002, 45(3), 253-261.
- [19] Prabhu G.S., and Shankar P.M., Simulation of Flat Fading Using MATLAB for Classroom Instructions, *IEEE Trans. Education*, 2002, 45(2), 19-25.
- [20] Riera V.J., Gimenez M.H., and Monsoriu J. A., Contribution of Digital Simulation in Visualizing Physics Processes, *Computer Applications in Engineering Education*, 2002, 10(1), 45-49.
- [21] Yann X.F. and Teng J.G., Interactive Web-Based Package for Computer-Aided Learning of Structural Behavior, *Computer Applications in Engineering Education*, 2003, 10(3), 121-137.