Improving Methodology To Teaching The Theory Of Random Processes With Wolfram Mathematica

Natalia M. Mezhennaya[1]

[1](Department of Applied Mathematics, Bauman Moscow State Technical University, Moscow, Russia)
Corresponding Author: Natalia M. Mezhennaya

Abstract: *The paper presents an approach to teaching the mathematical formalism to engineering majors by using interactive computer algebra system. The theory of random processes is a core element of mathematics training in an engineering university. The course is designed for students with a previous background in the probability theory and mathematical statistics, and with basic competence in some computer mathematical package. In our case it was Wolfram Mathematica. We provide examples to explain and illustrate our concept and discuss advantages of the method for students and teacher. These materials can also be used in a lecture course not supported by interactive computer systems or by students with insufficient mathematical programming skills. The examples we give belong to the basic concepts of the theory of random processes. We compare two groups of students to assess the results: Group I was taught in a "traditional" format without using interactive computer technologies, Group II had additional lessons in computer class and was involved in the modelling process. Statistically significant differences between groups were established (at 10% significance value). Another outcome was a positive attitude of students towards the proposed method. The presented samples are small, which makes further investigation desirable.*

Keywords: *computer algebra system, engineering education, interactive training, teaching methods, theory of random processes, Wolfram Mathematica.*

I.   Introduction

The modern science and technology and market demands force universities to focus on the students' personal growth, creative application of the acquired competences, and fast adaptation to the professional requirements. Students must be able to learn independently, find relevant information in various sources, analyze, systematize, reason, and draw adequate conclusions [1].

The probability theory, mathematical statistics, and theory of random processes are the key components of fundamental mathematical training in an engineering university [2]. To solve the nature of random processes, students require the integrated knowledge of various mathematical disciplines: calculus, algebra, differential equations etc. Interdisciplinary links are especially important in the context of the students' future careers in engineering [3].

The new educational technologies aim to systematize the students' knowledge, increase their motivation and professional adaptation, encourage achievement [4]. At present interactive computer systems are in high demand in education, especially within the framework of the mathematical disciplines utilized for engineering and scientific calculations [5]. The modern level of science and technology often involves numerical analysis of large amount of data (for example, Big data), so the analytical solution of the problem may be unfeasible or time-consuming. Therefore, the proficiency in computer data analysis and modelling becomes especially significant within the framework of systematic approach [6, 7] Soft skills in mathematical programming help solve many computational problems with high accuracy, speed and efficiency.

Spreadsheets are very common in statistical calculations. So it is one of the main analytical tools included into non-mathematical curriculum [8, 9]. However, these programs are not very convenient for modelling and analyzing phenomena defined by formulas. Consequently, interactive computer mathematical packages, such as Maple, Mathematica, MatLab, MathCAD are very convenient for use by students majoring in engineering [10-12]. Such systems include built-in matrix and complex arithmetic, support work with algebraic polynomials, enable numerical integration of differential and difference equations, plot various types of graphs, three-dimensional surfaces, and contain built-in procedures for modeling and analyzing random variables. Mathematical packages employ the conventional method of displaying mathematical objects and intuitive operating environment allowing the user to state problems and obtain solutions in the usual mathematical form without resorting to coding. The user-friendly interface of such systems also contributes to their popularity. Another feature of this software is that it stimulates the mathematical way of thinking.

Computer algebra systems allow the teacher to create templates for typical problems, where students can vary the initial data and obtain different intermediate and final results as analytical dependencies or graphic images [13]. Moreover, such templates do not require from the students deep understanding of these systems [10].

This paper presents certain methodological aspects of training engineering majors to understand and analyze mathematical formalism in the theory of random processes. Our experience shows that students very often learn basic mathematical formulas without understanding their meanings. Engineering students may also find it difficult to quickly analyze the properties of one- and two-dimensional functions analytically, losing the train of thought during lectures or practical lessons. To solve this discrepancy we suggest using computer algebra systems (CAS) in lectures and practical lessons to model and visualize the phenomena under consideration. We present a concept of teaching the basic notation for the theory of random processes using an innovative method. We consider basic properties of Wiener random process and discuss basic concepts of covariance and correlation functions, one- and two-dimensional distributions of the process and computer methods to explain and visualize them. Our examples are demonstrative and fairly simple. A similar research dealing with spreadsheets is presented in [14]. In the current paper we use the CAS Wolfram Mathematica which is accompanied by a large pool of examples and training materials available online. Also there exists a large amount of books with extra information and examples [15, 16].

For the sake of convenience, we present mainly monochrome images and do not explain and provide some of options used, because they can be neglected without affecting the content.

II. The Problem Of Understanding The Meaning Of Mathematical Formalism

The main idea of the approach is to illustrate the mathematical formalism by examples using interactive computer methods. The course under consideration is "The theory of random processes" for the third-year engineering majors. The course is designed for students who have previously studied the probability theory and mathematical statistics and have basic skills in any computer algebra system (CAS). However, our students may have used different mathematical packages out of the great number that are currently available. And some students have used none at all. That means that the examples should be suitable even for those students who have had no experience with Mathematica.

2.1. Preliminaries

The fundamental concept in the engineering application of the probability theory is normal distribution First of all, we give the definition of probability density function of bivariate normal law [17] with mean vector (m_1, m_2) and covariance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ of the form:

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{\tilde{x}^2}{\sigma_1^2} + \frac{\tilde{y}^2}{\sigma_2^2} - \frac{2r\tilde{x}\tilde{y}}{\sigma_1\sigma_2}\right)\right\}, \tag{1}$$

where $\tilde{x} = x - m_1$, $\tilde{y} = y - m_2$. Marginal probability density function of the first component is:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{\tilde{x}^2}{2\sigma_1^2}\right\}. \tag{2}$$

So we derive conditional probability density function [17] of the second component when the value of the first component is known:

$$f(y \mid x) = \frac{f(x,y)}{f(x)} = \frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} \exp\left\{-\frac{1}{2\sigma_2^2(1-r^2)}\left(\tilde{y} - \tilde{x}r\frac{\sigma_2}{\sigma_1}\right)^2\right\}. \tag{3}$$

The formulas (1)-(3) are most likely to be seen as rather complicated by most students. It is difficult to apply them to the analysis of the normal random processes properties. In engineering, distribution parameters may also be quite complex. So it is very effective to show the Mathematica built-in functions which produce, for example, formula (1). Function MultinormalDistribution[μ,Σ] represents a multivariate normal distribution with mean vector μ and covariance matrix Σ. Usual statistical functions [15, 16] are applicable as well, for example CDF, PDF, Mean etc. As the students will have already covered this material, we don't need to explain it in detail.

Let us proceed to the analysis of one of the most common normal random processes. We consider Wiener random process [18] $w_t, t \geq 0$, which is a continuous-time and continuous-state random process specified by the following three conditions:
1) $w_0 = 0$;

---

2) the increments of the process are independent, i. e. for $n \geq 1$ and $0 < t_1 < t_2 < \ldots < t_n$ random variables $w_{t_1}, w_{t_2} - w_{t_1}, w_{t_3} - w_{t_2}, \ldots, w_{t_n} - w_{t_{n-1}}$ are mutually independent;

3) the increments are distributed by normal law, i. e. $w_t - w_s$ has normal distribution with zero mean and variance $\sigma^2(t-s)$ for $t > s$.

Further we assume that $\sigma = 1$.

2.2. Example 1. Modelling A Path

First of all, let us model a path of the process and generate its plot (fig. 1):

proc=WienerProcess[];
sample1=RandomFunction[proc,{0, 5, 0.01}];
ListLinePlot[sample1]

Figure 1. Path Of Wiener Process w_t **with** $\sigma = 1$

For training purposes, it is essential that the students should be involved into modeling (if possible) [2] and asked to replot the process path several times to consider different cases. We note again that this does not require special knowledge of the CAS Mathematica [10].

2.3. Example 2. One- And Two-Dimensional Distributions Of w_t

Next we look at one-dimensional distribution of w_t for $\sigma = 1$. Covariance and correlation functions of w_t [18] are:

$$K(s,t) = \min\{s,t\}, \ R(s,t) = \frac{\min\{s,t\}}{\sqrt{st}} \qquad (4)$$

respectively, and one-dimensional probability distribution function is given by the formula:

$$f_w(x;t) = \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{x^2}{2t}\right\}, \ t > 0, \ x \in \square. \qquad (5)$$

Functions (4) and (5) can be derived in Mathematica:

R[s_,t_]:=CorrelationFunction[proc,s,t];
K[s_,t_]:=CovarianceFunction[proc,s,t];
PDF[NormalDistribution[0,Sqrt[K[t,t]]],x]
It is reasonable to plot the graphs:
Plot3D[K[s, t], {s, 0, 5}, {t, 0, 5}]
Plot3D[R[s, t], {s, 0, 5}, {t, 0, 5}]
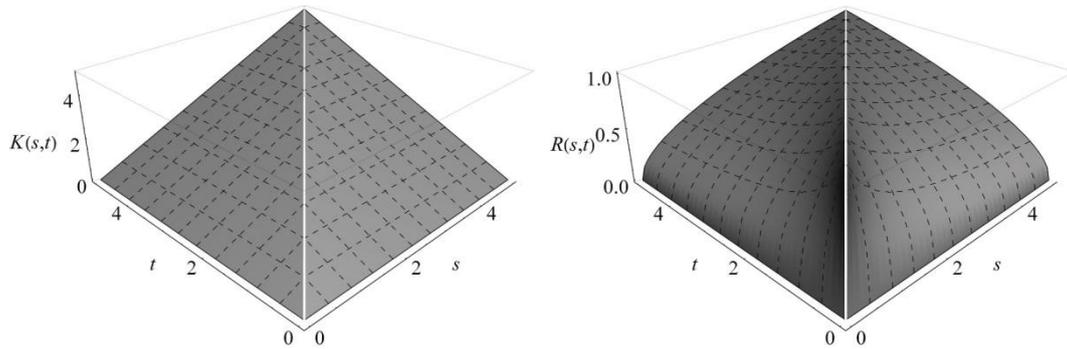Plot[Evaluate@Table[PDF[NormalDistribution[0, Sqrt[K[t, t]]], x], {t, {1, 3, 5}}], {x, -3, 3}]

Figure 2. Plot of covariance $K(s,t)$ **(left) and correlation** $R(s,t)$ **(right) functions of Wiener process** w_t **with** $\sigma = 1$
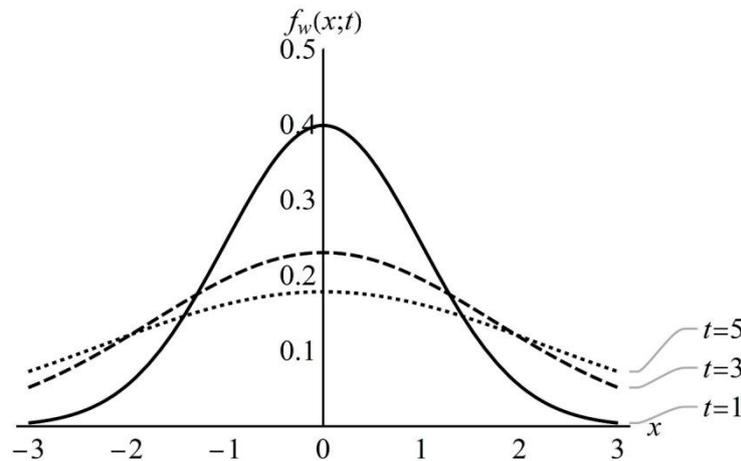

Figure 3. Plot of one-dimensional probability density function (5) of Wiener process w_t **with** $\sigma = 1$ **for different values of** t

We obtain graphs shown in fig. 2 (for covariance and correlation functions) and fig. 3 (for probability density function). The last plot (fig. 3) demonstrates that the process volatility increases with time. At this stage it is advisable to return to the plot of the path (fig. 1) and explain the impact of that fact.

Let us proceed to the properties of two-dimensional distributions [18]. For fixed $s < t$, the random vector (w_s, w_t) has normal distribution with the zero mean $(m_1, m_2) = (0,0)$ and covariance matrix $\begin{pmatrix} s & s \\ s & t \end{pmatrix}$ with two-dimensional density (1). So the probability density function is

$$f_w(x, y; s, t) = \frac{1}{2\pi\sqrt{s(t-s)}} \exp\left\{ -\frac{1}{2s(t-s)}\left(tx^2 + sy^2 - 2sxy \right) \right\}. \tag{6}$$

From (4) we conclude that $R(s,t) \to 0$ if s is fixed and $t \to \infty$. Plots of two-dimensional and conditional densities help us to explain this. Let $s = 1$. We use Mathematica built-in functions for covariance function of Wiener process to plot the graph and animate it.

Dist1[s_,t_]: = MultinormalDistribution[{0, 0}, {{K[s, s], K[s, t]}, {K[s, t], K[t, t]}}];
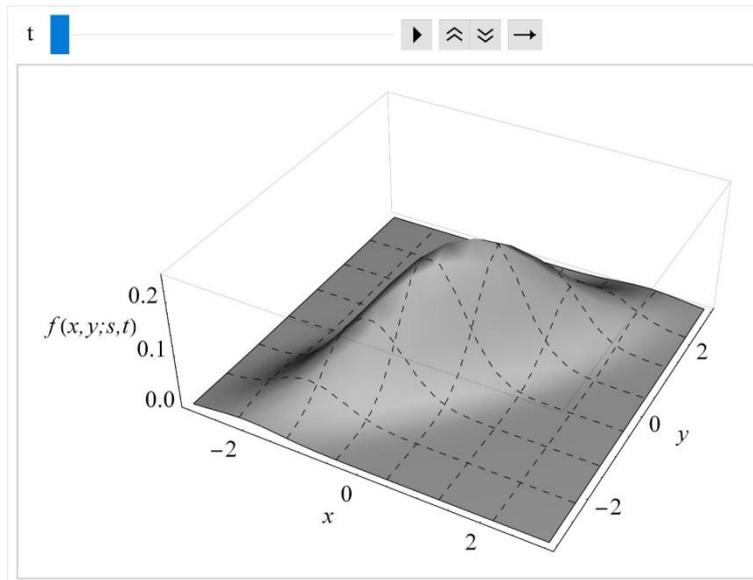Animate[Plot3D[PDF[Dist1[1, t], {x, y}], {x, -3, 3}, {y, -3, 3}],{t, 1.5, 5.0}]

Figure 4. Plot of two-dimensional probability density function (6) of Wiener process w_t **with** $\sigma = 1$ **for** $s = 1$ **and different values of** t **with animation controls**

The result is presented in fig. 4. By varying the plot parameters and initial values we can obtain a great amount of similar plots. This illustrative material can be used in class even without interactive computer support [13].

2.4.   Example 3. Conditional Distributions

Further we consider the conditional distribution of w_t if value of the process w_s is known $(0 < s < t)$. Substituting (5) and (6) into (3) we obtain the conditional density of w_t if $w_s = x_0$:

$$f_w(y;t \mid x_0;s) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{-\frac{(y-x_0)^2}{2(t-s)}\right\}. \tag{7}$$

For further practice the students can be assigned with formula (7) in Mathematica. Let $s = 1$ and find value x_0:

```
x0=Select[sample1["Path"],#[[1]]==1&][[1,2]]
```

For the path in Fig. 1 $w_1 = x_0 = -1.88$. Now we plot functions (5) and (7) to compare.
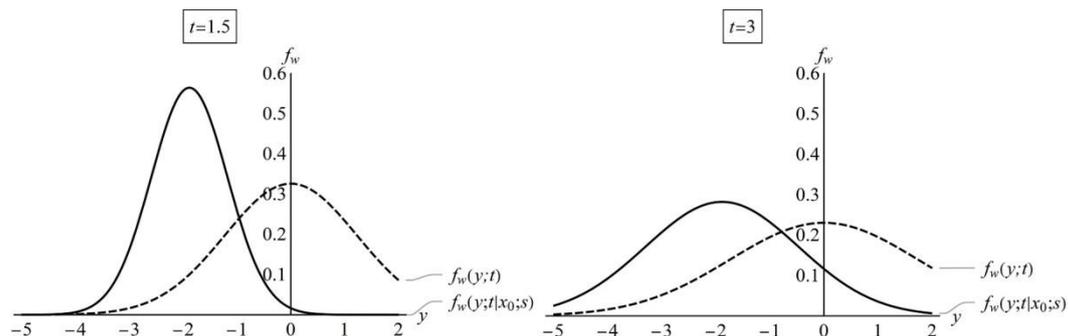

Figure 5. Conditional (solid line) and absolute (dashed line) probability density functions of Wiener process w_t **with** $\sigma = 1$ **and expected value** $x_0 = -1{,}88$ **for** $t = 1.5$ **(left) and** $t = 3$ **(right) and** $s = 1$

As expected (Fig. 5), the greater is the difference $t - s$, the greater is the conditional distribution volatility of process w_t. We also notice that the distribution center moves to x_0 and dispersion of the conditional distribution is less than dispersion of the absolute distribution. Finally, we combine the plot of path and the expected value x_0:

```
Show[ListLinePlot[sample1], Plot[x0, {t, 1, 5}]]
```
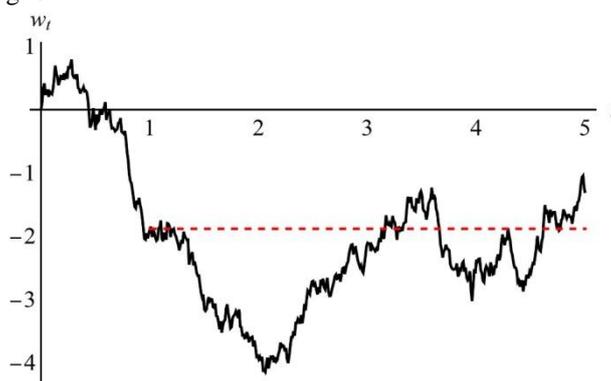
The result is presented in Fig. 6.

Figure 6. Plot of the path of Wiener process w_t **with** $\sigma = 1$ **and expected value** x_0 **(dashed line) of the process** w_t **for** $t > 1$**.**

2.5.   Example 4. Modelling The Distribution Of Running Maximum

Let $m_t = \sup\limits_{0 \le s \le t} w_s$ be the supremum value of the process till time t ("running supremum" or "running maximum" [18]). Fig. 7 can be used to demonstrate the running maximum concept to the students. Additionally, the students can remodel the path using the following template:

```
sample2 = RandomFunction[proc, {0, 5, 0.01}]["Path"];
L = Length[sample5];
max[l_] := Last[SortBy[l, Last]][[2]];
sample3 = Table[{sample2[[i, 1]], Last[SortBy[Take[sample2, i], Last]][[2]]}, {i, 1, L}];
ListLinePlot[{sample2, sample3}]
```
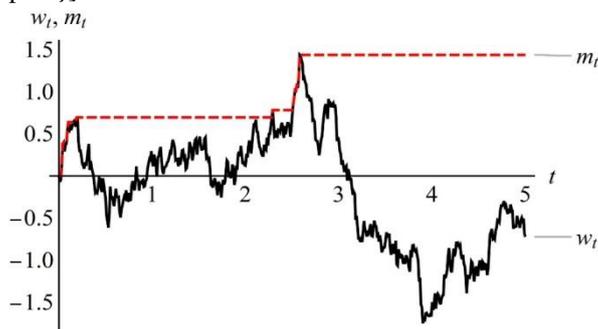

Figure 7. Plot of paths of Wiener process w_t with $\sigma = 1$ and running maximum m_t .

We create the function max[l_] := Last[SortBy[l, Last]][[2]] which generates maximum value for the list l if l represents the path of w_t . We review the way it works in practical lessons.

Next we consider the distribution of m_t . The cumulative distribution function and probability density functions of m_t [18] are:

$$F_m(x;t) = \mathbf{P}\{|w_t| < x\} = \frac{2}{\sqrt{2\pi t}} \int_0^x \exp\left\{-\frac{u^2}{2t}\right\} du, \tag{8}$$

$$f_m(x;t) = \left(F_m(x;t)\right)'_x = \frac{2}{\sqrt{2\pi t}} \exp\left\{-\frac{x^2}{2t}\right\}. \tag{9}$$

Let us see the plots (Fig. 8) of functions (8) and (9) for different values of t .
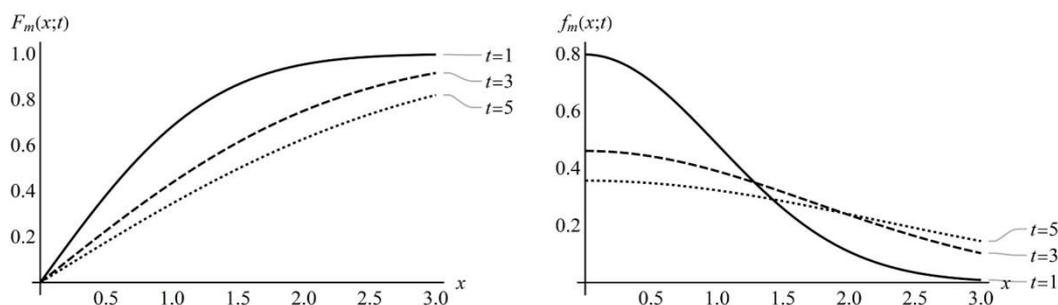
Figure 8. Plots of cumulative distribution function $F_m(x;t)$ **(8) (left) and probability density function**

$f_m(x;t)$ **(9) (right) for** t **=1, 3, 5**

The next step is to model the distribution of m_t for any fixed value of t. We generate a group of paths and estimate empirical distribution of the running maximum m_t. Our key teaching objective is to analyze the distributions and compare them. Let $t = 2$.

```
sample4 = Table[max[RandomFunction[WienerProcess[],{0,t,0.01}]["Path"]], {i, 1, 200}];
SampleDist = EmpiricalDistribution[sample4];
Fmax[x_, t_] := 2CDF[NormalDistribution[0, Sqrt[t]], x]-1;
fmax[x_, t_] := D[Fmax[x, t], x];
Plot[{Fmax[x, t], CDF[SampleDist, x]}, {x, 0, 5}]
Show[Histogram[sample4, Automatic, "PDF"], Plot[Evaluate@fmax[x, t], {x, 0, 5}]]
```
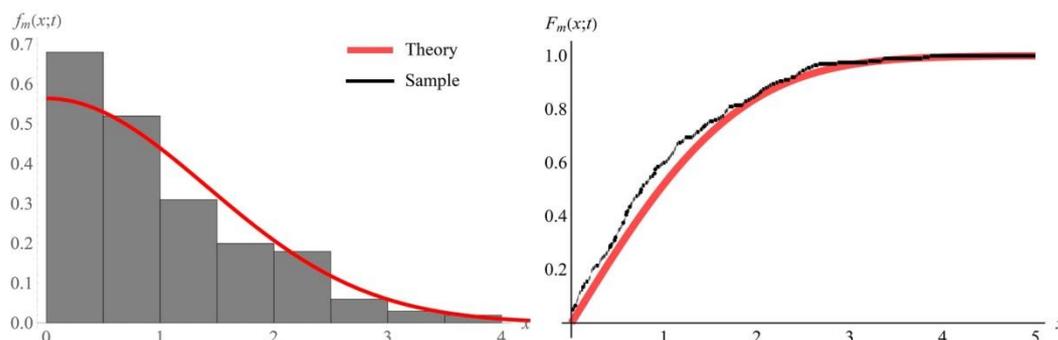

Figure 9. Plot of cumulative distribution functions of m_t **for** $t = 2$**: red curve – theoretical function (8);**

black curve – empirical function.

Fig. 9 demonstrates that for 200 paths the empirical and theoretical distributions are very close. During practical lessons students are asked to vary t and the number of paths to explore the properties of the distributions.

III. Results

We tested our teaching approach on two groups of the 3ʳᵈ year students taking a course in the theory of random processes. Group I consisted of 49 students and was taught in a "traditional" way with lectures and practical lessons without the CAS Mathematica. Group II consisted of 53 students and had additional classes in the computer modelling process. After that all students took a test consisting of 5 questions on "Random properties of Wiener process". Each test question was supplied with 3 answer options only one of which was correct. We calculated the number of correct answers for each student. A pass grade required minimum 3 correct answers. Results for the two groups are presented in Table 1. Percentage of correct answers is presented in Fig. 10.

Table 1. Number of correct answers for two groups and value of chi-square test statistic with p-value.

Number of correct answers	Group I	Group II	Sum	Chi-square statistic	p-value
<=2	17	7	24	7.10	0.069
3	14	17	31		

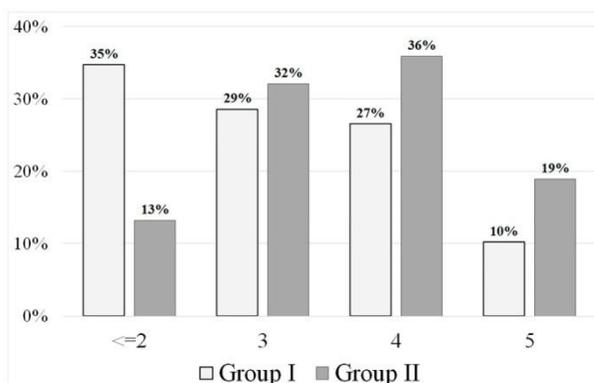| 4	13	19	32		
5	5	10	15		
Sum	49	53	102	-	-

Figure 10. Percentage of correct answers for two groups.

Fig. 10 shows that students of Group II demonstrate better results in comparison with students of Group I. We compare the distributions of the number of correct answers in two groups by the chi-square test to establish significant differences. The results are presented in table 1. We conclude that computer visualization improves understanding and reproduction of the learning material at 10%-significance level. The presented sample is fairly small, so further investigation is necessary.

We additionally questioned students of Group II to understand their opinion. We asked two questions:

1) Have you previously used any computer algebra system?

2) Do computer lessons improve your understanding of the study material?

We received 67% of positive answers for the first question, and 78% for the second. We can conclude that the students' attitude to the teaching approach under consideration is arbitrary positive.

It is worth noting that some questions and suggestions of our students enabled us to revise and significantly improve the teaching material.

IV. Conclusion

It was shown that presented in paper approach to teaching the Theory of Random Processes with support of CAS Mathematica helps significantly improve education outcomes for engineering students comparatively to the students who didn't use those materials. Students attitude towards the method was also positive. We give detailed examples of used interactive material which can be a basis for other researchers to construct analogous didactic complexes. The presented samples are small, so the further investigation should be mentioned.

References

[1]. O. N. Sezonova, S. A. Galchenko, and V. N. Khodirevskaya, The efficiency of higher education institutions as a basis for forming competent personnel for region economy, *European Journal of Contemporary Education, 18 (4),* 2016, 464-471. DOI: 0.13187/ejced.2016.18.464

[2]. E. A. Vlasova, N. M. Mezhennaya, V. S. Popov, and O. V. Pugachev, Methodological aspects of the discipline "Probability theory" in a technical university, *Modern High Technologies, 17 (11),* 2017, 96-103. (In Russian) DOI: 10.17513/snt.36852

[3]. A. Bendell, J. Disney, and C. McCollin (1999). The future role of statistics in quality engineering and management, *Journal of the Royal Statistical Society. Series D (The Statistician), 48 (3),* 1999, 299-326. DOI: 10.1111/1467-9884.00190

[4]. B. Kramarski and C. Hirsch (2003). Using computer algebra systems in mathematical classrooms, *Journal of Computer Assisted Learning, 19,* 2003, 35-45. DOI: 10.1046/j.0266-4909.2003.00004.x.

[5]. [5]. M. O. J. Thomas, Teachers using computers in the mathematics classroom: A longitudinal study. *New Zealand Mathematics Magazine*, *43 (3),* 2006, 6-16.

[6]. M. Kendal and K. Stacey, Teachers in transition: Moving towards CAS-supported classrooms, *Zentralblatt für Didaktik der Mathematik*, *34 (5),* 2002, 196-203. doi: 10.1007/BF02655822.

[7]. L. R. Wiest, The role of computers in mathematics teaching and learning, *Computers in the Schools, 17 (1-2),* 2008, 41-55. DOI: 10.1300/J025v17n01_05.

[8]. J. B. Soper and M. P. Lee, Spreadsheets in teaching statistics, *Journal of the Royal Statistical Society. Series D (The Statistician), 34 (3),* 1985, 317-321. DOI: 10.2307/2987658

[9]. R. R. Nuriakhmetov, Perspective approaches to teaching of statistics for nonmathematical students, *Novosibirsk State Pedagogical University Bulletin*, 3, 2012, 57-64. (In Russian).

[10]. S. Garner, The CAS classroom, *Australian Senior Mathematics Journal, 18 (2),* 2004, 28-42.

[11]. F. A. Ikhsanova, Engagement of staff to mathematical solution of applied problems with the help of the computer system of mathematical Mathematica, *Fundamental Research, 12-1,* 2011, 36-40. (In Russian).

[12]. O. A. Ivanov, V. V. Ivanova, and A. A. Saltan, Discrete mathematics course supported by CAS Mathematica, *International Journal of Mathematical Education in Science and Technology, 48 (6),* 2017, 953-963. DOI: 10.1080/0020739X.2017.1319979

[13]. E. A. Vlasova, N. M. Mezhennaya, V. S. Popov, and O. V. Pugachev, The use of mathematical packages in the framework of methodological support of probabilistic disciplines in a technical university, *Bulletin of Moscow Region State University. Series: Physics and Mathematics, 4,* 2017b, 114-128. (In Russian). DOI: 10.18384/2310-7251-2017-4-114-128

[14]. C. R. Fisher, A pedagogic demonstration of attenuation of correlation due to measurement error, *Spreadsheets in Education*, *7 (1),* 2014, article 4, from http://epublications.bond.edu.au/ejsie/vol7/iss1/4

[15]. S. Wolfram, *An elementary introduction to the Wolfram language.* 2nd ed. (Wolfram Media, Inc. Available from: http://www.wolfram.com/language/elementary-introduction/2nd-ed, 2015)

[16]. M. L. Abell and J. P. Braselton, *Mathematica by example*. *4th ed* (San Diego (SD): Elsevier Academic Press, 2008).

[17]. D. Stirzaker, *Elementary probability*. *2nd ed.* (Cambridge(UK); New York (NY): Cambridge University Press, 2003).

[18]. N. V. Krylov, *Introduction to the theory of random processes* (Providence: American mathematical society, 2002).