v. 4, n. 7, 2024

International Journal of Human Sciences Research

APPLICATIONS OF NEW INFORMATION TECHNOLOGIES IN THE TEACHING LEARNING PROCESS OF STATISTICS IN ENGINEERING

Felipe Ignacio Briones Nuñez
https://orcid.org/0000-0001-6335-017X

David Echeverria Burbano

Edgar Ramiro Guerrón Varela
https://orcid.org/0000-0002-0844-1113

Sergio Castillo Páez
https://orcid.org/0000-0002-5402-2462

Jacobo Sebastian Vasquez Rodriguez
https://orcid.org/0009-0009-1827-6869

Juan Carlos Sarango Cuenca
https://orcid.org/0009-0009-3948-128X

Vicente Garcia Mancero
https://orcid.org/0009-0005-7501-1657

INTRODUCTION

In the last decade, we have witnessed drastic changes in all areas of society due to the advancement of technology and the internet. These changes have had a notable impact on industry, commerce, education and knowledge, among other aspects.

As a result of these advances, development companies have launched products that are more adaptable to various disciplines. This has allowed the learning and research processes to be more versatile. In short, technology and the internet have revolutionized our way of interacting with the world, giving us infinite possibilities and making our daily lives easier.

Against this backdrop, the accelerated development of information and knowledge raises the need for new and better strategies in education, taking advantage of global resources and adapting to the current needs of the generation. This new technological world offers great advantages both for the teacher, as a fundamental work tool, and for the student, providing a means of access, simulation, practice and development.

In this sense, in this new educational context, the classroom is no longer the only channel through which these new generations come into contact with knowledge. Therefore, both the teacher's word and the written text are no longer the only exclusive supports of education" (Brunner, 2000).

Therefore, the teacher must adapt to new circumstances and assume new roles and functions. It is not only technologies that produce changes or allow innovation. Furthermore, the challenge lies in the pedagogical stance used as a basis for planning. According to Dias-Barriga & Hernández (2002), this position must be constructivist, since knowledge is not a mere copy of reality, but a human construction that is carried out with previous schemes and previous constructions. This conception of knowledge construction motivates us to put ICT into practice to teach statistics.

Finally, we have developed (see fig 1) and applied an application based on the current syllable of the subject of probability and statistics, corresponding to the first stage of the professional training axis of engineering careers. One of the objectives of this application is to allow students to observe in real time how decisions to choose the value of some parameter affect the results.

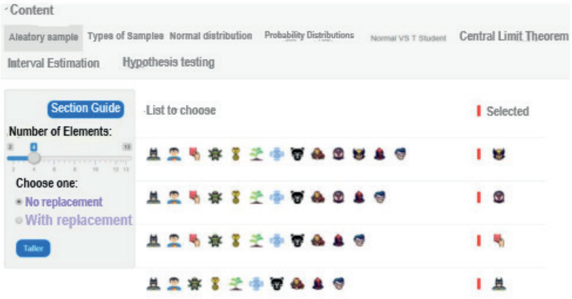

Figure 1: Application Developed in Rstudio for teaching Statistics.

STATE OF THE ART

According to Chanto (2018), the incorporation of Information and Communication Technologies (ICT) in the educational field has introduced an innovative approach to strengthen the teaching and learning process for 21st century students. These technologies, characterized by their flexibility, adapt effectively to the individual needs of students. Chanto emphasizes that the objective of ICT is not to replace traditional education, but to transform it and elevate it to a higher level, thanks to the ease and availability of access to information at any time.

In line with this, Contreras (2007), in his article "Evolution of virtual classrooms in traditional Chilean universities", already anticipates the need to integrate virtual resources in the study programs of university careers. According to Contreras, these resources are essential to complement and enrich learning in the context of traditional education.

Following this trend, at a mathematics conference, Campos and Delgado (2015) proposed that teachers must create specific learning environments for their field of teaching or research. This proposal included the implementation of online teaching sequences, designed to complement the face-to-face, mixed or distance modalities of their respective classes.

In response to these proposals, Acosta (2011) presented a national initiative that sought to integrate the use of virtual tools in ESPE's face-to-face mathematics classes, with a particular focus on algebra. The objective was to strengthen and closely monitor student learning in the classroom through the use of these tools.

Taking this background into account, the previously mentioned works, which are related to the applications of new information technologies in the teaching and learning process of statistics in engineering, are considered as reference points for the present research.

METHOD

According to Pita Fernández and Pértegas Díaz (2002), quantitative research is conceived as a method in which quantitative data on variables are collected and analyzed. On the other hand, qualitative research avoids quantification, and researchers who use it make narrative records of the phenomena studied using techniques such as participant observation and unstructured interviews. The fundamental difference between both methodologies is that the quantitative method focuses on studying the association or relationship between quantified variables, while the qualitative method does so in structural and situational contexts.

This work is developed under a quantitative approach, since quantitative data is collected and analyzed. It starts from the idea of creating an application for learning statistics and the statement of the problem. Subsequently, the independent variable is established, which in this research is the use of the application developed in Shiny, and the dependent variable, which is the improvement in learning the subject. The hypothesis is based on the collection and analysis of data, such as a test of questions that will be applied at the end of the intervention. This test will measure whether the use of the application improved learning, contrasting the results with the control group.

Therefore, an experimental investigation will be carried out with two groups already formed randomly, that is, intact groups. In addition, a post-intervention study will be carried out, which involves carrying out observations after applying an intervention measure in one of the groups.

SAMPLE

The students taking part in the research represent a variety of engineering majors. During the study period, these students were enrolled in the subject of Probability and Statistics. The program for this subject is structured in three partial parts. The application was developed and integrated mainly during the second term, which coincided with the moment in which the students tested and used it. Table #1 illustrates the distribution of students between the control group and the experimental group.

Groups	Student number
Experimental	25
Control	19
Total	44

Table 1:Sample

DATA COLLECTION TECHNIQUES AND INSTRUMENTS

In the context of this research, it was decided to use a questionnaire as a technique and instrument for collecting information and verifying the proposed objectives. This questionnaire, according to the definition of Sampieri (2014), is based on a set of questions that are based on one or more variables to be measured.

Therefore, a questionnaire was developed focused on the contents of the second unit of the ESPE engineering career study program. In order to represent the control and experimental groups, courses A and B from the same degree and semester were selected.

Finally, 20 questions were asked that will be answered by the students of both courses in a period of 1 hour, simultaneously, through the University's Moodle platform. Thus, the results obtained by the control group and the experimental group can be evaluated and contrasted.

WORK METHODOLOGY

Based on the syllabus of the Statistics subject, an application was designed in Shiny that, after being tested, was hosted on a cloud server. From that moment on, virtual tutorials for its use began, as well as laboratory practices and examples. This way, during the course of regular classes, the students of the experimental group had the opportunity to become familiar with the use of the application.

The web application, developed with the Shiny framework based on R (Figure 2), covers content from each of the most representative topics of the three units: Measures of Central Tendency, Binomial Distribution, Types of Samples, Random Sample, Normal Distribution, Probability Distributions, Normal VS T Student, Central Limit Theorem, Interval Estimation, Hypothesis Testing and Linear Regression. Thanks to the interactivity of the platform, students have the possibility of manipulating data entry parameters and observing how these influence the final result.

Figure 2: Sample types exercise developed in Shiny.

Finally, the development of each section in Shiny is designed so that the student is the main protagonist of their learning. The application facilitates the interaction of students with the parameters of the different topics, allowing them to analyze in real time the possible problems to be solved. In this context, the teacher assumes the role of a virtual tutor during laboratory practices, providing activities that help students become familiar with the application and its benefits. This way, students have the opportunity to experiment autonomously with different problems.

Regarding autonomous work at home, both the control and experimental groups have access to the University's virtual classroom. On this platform, the teacher uploads various activities (videos, questionnaires, tasks) that are used in the teaching-learning process.

On the other hand, the students in the experimental group, in addition to using Moodle, began to use the Shiny application at their discretion. This tool gave them the possibility of carrying out activities, regular tasks, exercises and problems that arose during the Statistics course. They also had the opportunity to modify parameters, verify the resulting changes and contrast them with theory.

The work with tutoring and autonomous work lasted three weeks. At the end of this period, the students' knowledge was evaluated using a questionnaire as a data collection tool.

RESEARCH RESULTS

ANALYSIS AND INTERPRETATION OF RESULTS

Table #2 collected the 19 quantitative notes out of 5 from the control group that carried out the evaluation instrument to determine the learning achieved at the end of this experimentation.

Note	Frequency	Relative frequency
3,0	2	0.112
3,0	3	0.158
4,0	2	0.105
5,0	2	0.105
6,0	3	0.158
7,0	3	0.158
10,0	4	0.210
Total	**19**	

Table 2: Frequencies of the control group.

Taking into account that the minimum passing grade is 7.0, the data in table #2 was used to construct the circular diagram shown in figure #3. This diagram reveals that only 37% of the students passed the exam, while the remaining 63% failed to achieve the minimum required score.

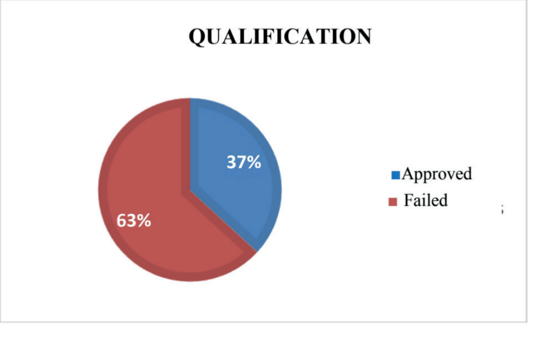

Figure 3: Passed and failed students of the control group.

In Table 3, the quantitative ratings of 25 members of the experimental group were recorded. These ratings were obtained through an evaluation instrument that was used to determine the level of learning achieved at the end of the experimentation.

Note	Frequency	Relative frequency
4,0	1	0.040
5,0	2	0.080
6,0	3	0.120
7,0	7	0.280
8,0	7	0.280
9,0	4	0.160
10,0	1	0.040
Total	25	

Table 3: Frequencies of the experimental group.

Taking into account that the minimum passing grade is 7, the information in Table #3 was used to construct the pie diagram shown in Figure #4. This diagram reveals that 76% of the students passed the exam, while 24% failed to pass.

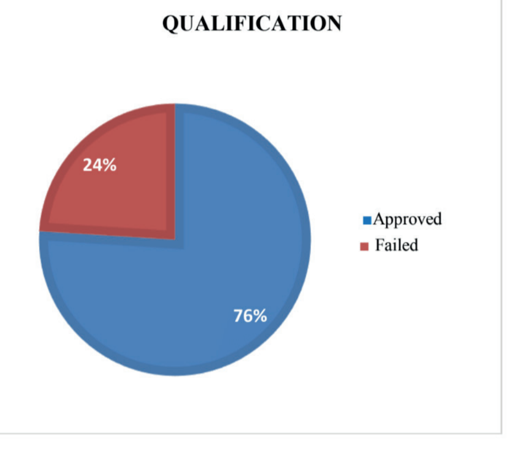

Figure 4: Approved and failed students of the experimental group.

Using various functions in R, statistics were calculated for the experimental group and the control group. Table 4 presents the measures of central tendency for each group.

International Journal of Human Sciences Research ISSN 2764-0558 DOI 10.22533/at.ed.558472423025

For the control group, the mean is 5.789. In contrast, the experimental group, where the statistical application was implemented, has a higher average of 7.32.

	Control Group	Experimental Group
Min	2	4
Max	10	10
Range	8	6
1st quartile	3,5	7,0
Median	6	7
3rd quartile	7	8
Average	5,789	7,320
Standard deviation	2.7402	1.4353
Variance	7.5087	2.0601

Table 4: Position and dispersion measurements.

HYPOTHESIS TESTING HYPOTHESIS CHECKING

With the data obtained in the experimental phase, the hypothesis posed at the beginning of this research was verified. The procedure to follow was:

1. Checking the normality of the two data sets.

2. Test for homoscedasticity of variances

3. Comparison test of two independent means, using the respective t-Student statistic.

The hypotheses to be verified are the following:

- Null Hypothesis

$H0{:}HGE{\leq}HGC$

"The incidence of an implementation of a digital platform to strengthen theoretical and practical knowledge of the subject of statistics at the University of the Armed Forces is lower or equal in the group where the digital platform was applied than in the group where it was not used".

- Alternative Hypothesis

$H1{:}\ HGE{>}HGC$

"The incidence of an implementation of a digital platform to strengthen theoretical and practical knowledge of the statistics subject of the University of the Armed Forces is greater in the group where the digital platform was applied than in the group where it was not." was employed".

NORMALITY TEST

The normality test is an analysis that allows us to determine whether the values of a variable follow a normal distribution. In the context of this research, the dependent variable is academic performance. This is measured based on the grades obtained by the students of the experimental and control groups, who are studying undergraduate courses at the University of the Armed Forces. The objective of this analysis is to verify the normality of the distribution of these ratings.

```
shapiro.test(Control)

##
##  Shapiro-Wilk normality test
##
## data:  Control
## W = 0.90729, p-value = 0.06603
```

```
shapiro.test(Experimental)

##
##  Shapiro-Wilk normality test
##
## data:  Experimental
## W = 0.94928, p-value = 0.2415
```

Figure 5: Shapiro test of the control and experimental group.

As can be seen in Figure 5, the Shapiro-Wilk normality test provides us with a value of p=0.06603 for the control group. This value is higher than the established significance level, α=0.05. This implies that we cannot reject the null hypothesis, H0. In other words, we do not have enough statistical evidence to contradict the null hypothesis. Therefore, we conclude that the control group data follow a normal distribution.

In the normality test carried out for the experimental group, a p-value of 0.2415 was obtained, which is higher than the established significance level (α=0.05). This implies that the null hypothesis (H_0) cannot be rejected. Therefore, it can be inferred that the data of the experimental group follow a normal distribution. This interpretation is based on the results obtained in the statistical analysis.

International Journal of Human Sciences Research ISSN 2764-0558 DOI 10.22533/at.ed.558472423025

HOMOSCEDASTICITY TEST OF TWO MEANS

For the homoscedasticity contrast between two means, we initially examined the distribution of the data using a box-and-whisker plot. Through this visualization, it can be seen that the variances are different. This suggests that the data may not meet the assumption of homoscedasticity, which is crucial for certain statistical analyses. This interpretation is based on visual inspection of the box-and-whisker plot (Figure 6).

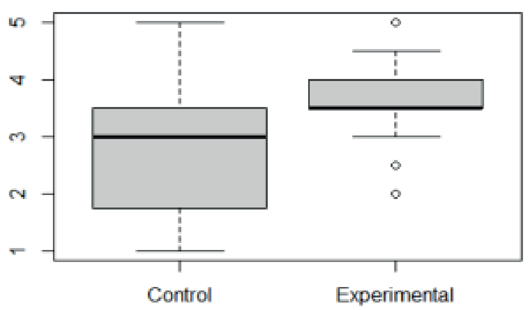

Figure 6: Box and whisker plot for control and experimental group.

To corroborate the presence of heteroskedasticity in our data, we chose to perform a variance test (figure 7). This test is carried out using the var.test function in our analysis. This step is crucial to confirm our initial observations and provide a solid statistical basis for our conclusions.

```
var.test(x=Control,y=Experimental)

##
##  F test to compare two variances
##
## data:  Control and Experimental
## F = 3.645, num df = 18, denom df = 24, p-value = 0.003608
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  1.541374 9.122417
## sample estimates:
## ratio of variances
##          3.645035
```

Figure 7: var.test between Control and Experimental group.

From the variance test, a p-value of 0.00368 was obtained, which is lower than the established significance level (α=0.05). This statistical evidence leads us to reject the null hypothesis and accept the alternative hypothesis. Consequently, we conclude that the variances are significantly different. This result supports our initial observations and confirms the presence of heteroscedasticity in our data.

TWO-MEANS HYPOTHESIS TEST

To carry out the mean contrast test, we begin by verifying the normality of the variables using the Shapiro-Wilk statistic. Subsequently, we carried out the contrast of means, applying the criterion of equality of variances. As a result, we determined that the samples from the control group and the experimental group follow a normal distribution, although with unknown and different variances.

The Student's t-hypothesis test was used to compare the means, specifically for two independent means with unknown and different variances. From this test, the original research hypotheses could be verified.

```
t.test(x = Experimental, y = Control, alternative = "greater", mu = 0,
var.equal = F, conf.level = 0.95)

##
##  Welch Two Sample t-test
##
## data:  Experimental and Control
## t = 2.2147, df = 25.459, p-value = 0.01798
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.1754365       Inf
## sample estimates:
## mean of x mean of y
##  3.660000  2.894737
```

Figure 8: T-test between the control and experimental group.

As can be seen in Figure 8, the test returned a p-value of 0.01798, which is lower than the established significance level (α = 0.05). This implies that the null hypothesis ($H0$) is rejected and the alternative hypothesis ($H1$) is accepted. In other words, the mean of the experimental group is statistically greater than the mean of the control group.

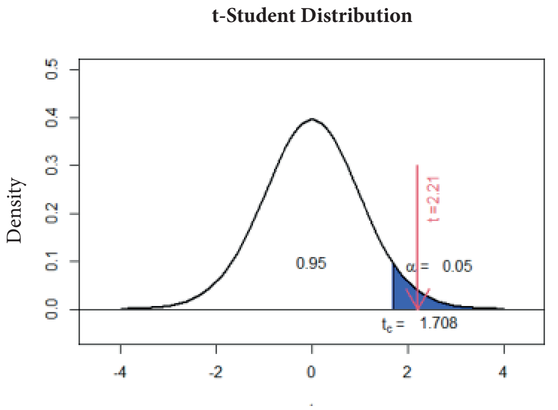

Figure 9: Hypothesis test graph – Comparison of means.

Consequently, "The incidence of an implementation of a digital platform to strengthen theoretical and practical knowledge of the subject of statistics at the University of the Armed Forces is greater in the group where the digital platform was applied than in the group where "It was not used." This observation confirms the hypothesis raised in this research.

CONCLUSIONS

- The main conclusion of this research is that the developed application had a positive influence on student performance. Specifically, those who used the application showed superior performance, which resulted in a greater number of passes in the statistics subject.
- Through the analysis of the data collected, it was evident that the average of the experimental group exceeded that of the control group. This confirms that the understanding of the theory of the statistics subject was better in the group that used the application.
- Shiny, an R tool, proved to be very useful in teaching and learning statistics. The developed application is interactive, allowing students to make changes to the entered data and view the results in real time. This makes it easier to manipulate data and verify results, both in calculating values and generating data or distribution graphs.

REFERENCES

Acosta M. (2011). *Gestion del docente de matemáticas en el aula virtual para el aprendizaje de álgebra de los estudiantes del prepolitécnico de ingeniería en mercadotecnia de la ESPE.* Obtenido de http://repositorio.uta.edu.ec/bitstream/123456789/13297/1/BG-1424.pd

Brunner, J. J. (2000). *Nuevos escenarios de la educación. Revolución tecnológica y Sociedad de la Información. Santiago.* Obtenido de http://www.preal.cl/

Campos, Y., & Delgado, J. (2015). *Entornos de Aprendizaje de la Matemática en la Educación.* Obtenido de http://xiv.ciaem-redumate.org/index.php/xiv_

Chanto, C. (2018). *El Aula Virtual Como Estrategia Para La Enseñanza Y Al Aprendizaje.* Obtenido de http://www.ajer.org/papers/Vol-7-issue-1/L070181 87.pdf.

Contreras, W. (2007). *EVOLUCIÓN DE LAS AULAS VIRTUALES EN LAS UNIVERSIDADES.* Obtenido de https://www.redalyc.org/pdf/979/97916199006.pdf

Díaz Barriga, & Hernández. (2002). *https://www.uv.mx/dgdaie/files/2012/11/CPP-DC-Diaz-Barriga-Estrategias-de-ensenanza.pdf.* Obtenido de https://www.uv.mx/dgdaie/files/2012/11/CPP-DC-Diaz-Barriga-Estrategias-de-ensenanza.pdf

Pita Fernández, & Pértegas Díaz. (2002). *Investigación cuantitativa y cualitativa.* Obtenido de http://www.ecominga.uqam.ca/PDF/BIBLIOGRAPHIE/GUIDE_LECTURE_2/4/2.Pita_Fernandez_y_Pertegas_Diaz.pdf

Sampieri, R. (2014). *Metodología de la investigación (Vol. 6). Mc Graw Hill.* Obtenido de https://www.uca.ac.cr/wp-content/uploads/2017/10/Investigacion.pdf