http://doi.org/10.14393/ER-v30a2023-46

Spreadsheets: a didactic tool for statistical analysis in the composition of a population[1]

Beatriz Marrón[2]
Ricardo Camina[3]
Verónica San Román[4]

ABSTRACT

This paper presents an experience carried out by teachers of the Statistics area of the Mathematics Department of the Universidad Nacional del Sur. It was carried out with second year students of the Bachelor's Degree in Biological Sciences, who are taking the subject Biostatistics and are aged between 19 and 21 years old. We share a teaching proposal using a didactic device that promotes the use of Information and Communication Technologies (ICT) by students. It was designed by the department and is based on spreadsheets containing the necessary tools for data processing. In this experience, the concept of sample composition of qualitative variables, dominance and diversity indexes and graphic representation by means of ternary diagrams were addressed. This didactic instrument involved students in intellectual activities related to reflection, exploration capacity, investigative interest, critical thinking, and problem-solving skills.

KEYWORDS: University education. Statistical Education. ICT.

[1] English version by the authors: Beatriz Marrón. *E-mail*: beatriz.marron@uns.edu.ar.

[2] Doctora en Matemática. Departamento de Matemática. Universidad. Universidad Nacional del Sur, Bahía Blanca, Buenos Aires, Aregntina. Orcid: https://orcid.org/0000-0002-5447-8515. *E-mail*: beatriz.marron@uns.edu.ar.
[3] Licenciado en Biología. Departamento de Matemática.Universidad. Universidad Nacional del Sur, Bahía Blanca, Buenos Aires, Aregntina. *E-mail*: rcamina@criba.edu.ar.
[4] Licenciada en Educación Matemática. Departamento de Matemática. Universidad Nacional del Sur, Bahía Blanca, Buenos Aires, Aregntina. Orcid: https://orcid.org/0000-0001-7598-0830. *E-mail*: vsanroman@gmail.com.

Folhas de cálculo: uma ferramenta pedagógica para a análise estatística da composição de uma população

RESUMO

Este trabalho apresenta uma experiência realizada por professores da área de Estatística do Departamento de Matemática da Universidad Nacional del Sur. Foi realizada com alunos do segundo ano da Licenciatura em Ciências Biológicas, que estão a frequentar a disciplina de Bioestatística e têm idades compreendidas entre os 19 e os 21 anos. Partilhamos uma proposta de ensino que utiliza um dispositivo didáctico que incentiva a utilização das Tecnologias de Informação e Comunicação (TIC) pelos alunos. Foi concebido pelo departamento e baseia-se em folhas de cálculo que contêm as ferramentas necessárias para o tratamento de dados. Nesta experiência, foi abordado o conceito de composição de amostras de variáveis qualitativas, índices de dominância e diversidade e representação gráfica através de diagramas ternários. Este instrumento didáctico envolveu os alunos em actividades intelectuais relacionadas com a reflexão, a capacidade de exploração, o interesse pela investigação, o pensamento crítico e a capacidade de resolução de problemas.

PALAVRAS-CHAVE: Ensino universitário. Educação estatística. TIC.

Las planillas de cálculo: una herramienta didáctica para el análisis estadístico en la composición de una población.

RESUMEN

En este trabajo se presenta una experiencia llevada a cabo por docentes del área de Estadística perteneciente al Departamento de Matemática de la Universidad Nacional del Sur. La misma se realizó con estudiantes de segundo año de la Licenciatura en Ciencias Biológicas, quienes están cursando la materia Bioestadística y tienen edades comprendidas entre los 19 y 21 años. Compartimos una propuesta de enseñanza donde se utilizó un dispositivo didáctico que fomenta el uso de las Tecnologías de la Información y la Comunicación (TIC) por parte de los estudiantes. El mismo fue diseñado por la cátedra y se basa en hojas de cálculo que contienen las herramientas necesarias para el procesamiento de datos. En esta experiencia se abordó el concepto de composición de muestras de variables

cualitativas, índices de dominancia y diversidad y representación gráfica mediante diagramas ternarios. Este instrumento didáctico involucró al estudiantado en actividades intelectuales relacionadas con la reflexión, la capacidad de exploración, el interés investigativo, el pensamiento crítico, y la habilidad para resolver problemas.
PALABRAS CLAVE: Enseñanza universitaria. Educación Estadística. TIC.

* * *

Introduction

The teaching of statistics in biology courses is becoming increasingly important as the use of statistical methods has evolved both in research and in everyday practice. In this sense, it is crucial to incorporate contextualized activities in the teaching of science that involve controversial situations, which are able to connect the current world with the daily life of the students. This will ensure a good understanding of the nature of science as a determining factor in improving levels of argumentation and decision making.

Taking into account a constructivist epistemological perspective, scientific knowledge is constructed by the scientific community, which seeks to measure and build models of the natural world. Therefore, natural science consists of mental constructions that aim to explain sensory experience and measurements. From this position, it is desirable that the teaching of statistics in the formation courses of future biologists allows, besides other things, that students develop their own conceptualizations, analyze the behavior of qualitative or quantitative variables involved in a biological model and identify the areas of their interest in order to transform and resignify them.

With this proposal, our objective is to give a new meaning to the contents linked to qualitative variables, by recognizing the categories in which the experimental units can be grouped, incorporating them into a

specific category of information known as “compositional data”. Through its analysis, other key concepts emerge, such as “amalgamation”, which arises when some of the categories of interest have a reduced representation, and the evaluation of the complexity or simplicity of the structure of a sample through the analysis of the diversity and dominance indexes, which summarize the distribution of the proportions of the sampled categories.

Therefore, the main task of teachers will be to create a collaborative environment in which students actively participate in their own learning, offering themselves as a guide and support during the teaching-learning process.

The precedents indicate that, despite the growing importance of statistics in biology, little work has been done on teaching it in an integrated way. Horgan et al. (1999) provide anecdotal evidence of student learning through continuing education workshops in statistics for biological research scientists. Metz (2008, p.325) sustain that:

> Students in an introductory biology course showed statistically significant gains in their understanding of statistics when given the means and opportunity to practice data gathering and analysis of biological experiments. Learning gains were measured using a new instrument designed specifically to measure the students’ knowledge of statistics in the biology classroom. The work provides a model for strengthening the statistical analysis skills of biology majors.

As Garcia Perez (2000), cited in Hernandez et al. (2021, p.2), states,

> didactic models are a powerful intellectual tool to guide learning. Models make it possible to address educational problems in all disciplines, including statistics, contributing to establish links between theoretical analysis and practical intervention. However, this connection is not always visible in

> statistical education, where theory and practice usually seem to be separated.

Also, considering the complexity of biological systems, associated with experimental variability, the construction of mathematical models and the development of statistics are required to understand the structure of ecosystems and to analyze experimental data in a quantitative and qualitative way. As Mcleod (2023) argues, integrating modeling into the teaching of statistics encourages the promotion of critical thinking, problem solving and the integration of abstract concepts with real life situations.

To summarize, the development of statistics has had great intervention in the biological sciences and the synergy between the two has enriched the achievements of both sciences. It is a highly interdisciplinary relationship, which is reflected in the stochastic approaches and statistical applications that are required to reach real explanations in the field of biology. Hence, if we are able to overcome the traditional approach, the door opens to a wide range of possibilities, both in a face-to-face environment and in the virtual space, where the potential of Information and Communication Technologies (ICT) can be fully exploited for the development of integrative knowledge from a collaborative research perspective.

Methodological strategies

Methodological strategies are a set of techniques, resources, activities and didactic tools that are intentionally selected and organized to promote meaningful learning. As Davini (2008, p.178) indicates:

> Teachers will make decisions using one of the methods or will freely carry out a combination of their inputs, elaborating their own teaching strategy. Their particular strategy will represent an appropriate methodical sequence according to the characteristics of their students, the

teaching context (educational level, specialty), the content (knowledge, skills, values), as well as the educational intentions and learning objectives.

Under a constructivist approach of learning, the incorporation of digital tools applied to the teaching of statistics offers a great variety of methodological strategies. Therefore, it is essential to take special care in the selection of these tools for the teaching of mathematics [statistics] in the training of undergraduate students in the biological sciences (Pérez et al., 2006).

In this context, some of the methodological strategies available are related to:

- the use of statistical software that allows students to make graphs and then data analysis to reinforce the theoretical knowledge learned in class.
- the incorporation of gamification that allows students to learn in a playful way, applying statistical concepts in real and challenging situations.
- the use of virtual platforms that allow students to interact and collaborate in solving statistical problems, encouraging team work and providing effective feedback.
- the promotion of multimedia educational resources, such as videos, interactive presentations and simulators, that allow students to learn in a visual and participatory way.

It should be noted that the above mentioned pedagogical strategies can be applied inside and outside the classroom, allowing for greater flexibility in the pace and manner in which each student learns.

Didactic resources

The course Biostatistics is part of the basic training of the future graduate in Biology. In general, at the time it is taught, students do not have a deep knowledge of the basic discipline, biology, to use it effectively.

Therefore, the selection of didactic resources that strengthen the teaching-learning process and allow the articulation of the contents is very important. As Vargas Murillo (2017, p.69) argues:

> The functions of didactic resources must take into account the target group, so that the resource is actually useful. Among the functions of didactic resources are: a) to provide information, b) to fulfill an objective, c) to guide the teaching and learning process, d) to contextualize students, e) to facilitate communication between teachers and students, f) to bring ideas closer to the senses, g) to motivate students.

In view of the above, the selection of didactic resources in the proposed teaching sequence is based on two fundamental concepts: to consolidate the learning of knowledge and to mediate the encounter of the students with the reality through the concrete material.

With this objective in mind, it was decided to use a metaphorical representation of the diversity of the population by incorporating packets of different colored candied chocolate. By manipulating this concrete material, students will be able to perform a variety of activities, such as classifying the candies by color, calculating the frequency in each case, and determining the proportion of candies of a specific color to the total. These activities will allow students to understand basic statistical concepts such as relative frequency and proportion, thus laying a solid foundation for the development of more advanced concepts. In addition, it will provide an insight into how statistical data are explained in biological research.

In this context, it is pertinent to incorporate spreadsheets as an additional didactic resource. As is pointed out by Henao Alvarez (1996, p.103):

> Although originally developed as tools for work in the world of projection and financial analysis, spreadsheets have great potential in the field of teaching and learning.

The professor in charge of the course created spreadsheets using the functions already available in the selected program. In order to preserve their value as teaching tools, the cells have been protected so that, in this way, it is only possible to modify the cells used for data entry. These tools facilitate data analysis in a simple and transparent manner, which encourages the development of fundamental experimental designs. In addition, they promote the active participation of students in intellectual activities, strengthening their skills of reflection, exploration, research, critical thinking and problem solving. In this way, a significant integration of ICT in the getting, development and subsequent analysis of the results is achieved.

Experience contextualization

The selected reference institution is the Universidad Nacional del Sur (UNS), a public, autonomous and autarchic university, which provides higher education for professionals and is located in the city of Bahía Blanca, province of Buenos Aires, Argentina. Since its creation, the UNS has adopted the organization in Departments, and not in faculties, for the development of its teaching and research activities. In this environment, interdisciplinarity and contextualization emerge as elements that are highly valued in the framework of teaching and learning processes. These approaches provide teachers with an integral perspective that enriches their ability to make informed decisions.

This teaching-learning proposal was developed by professors belonging to the Department of Mathematics in the area of Statistics. This initiative has been applied in the Biostatistics course, which corresponds to the second

year of the curriculum of the Bachelor in Biological Sciences of the UNS. Although the contents addressed are related to descriptive statistics and are within the first unit of the subject program, they will be relevant in subsequent thematic units.

Methodology

The planning of this teaching sequence is based on a constructivist perspective of learning where reality is an inherent internal construction of the individual. As Serrano Gonzalez-Tejero and Pons Parra (2011, p.11) claim:

> constructivism, in essence, states that knowledge is not the result of a mere copy of the preexisting reality, but of a dynamic and interactive process through which external information is interpreted and reinterpreted by the mind. In this process, the mind progressively builds up explanatory models, each time more complex and powerful, so that we know the reality through the models we build ad hoc to explain it.

On the other hand, in the framework of the constructivist approach of learning, it is relevant to highlight the role of ICTs, which are integrated into the educational process to enhance the experience of students. Incorporating ICTs provides students with the valuable opportunity to "experiment" in a statistical way, which significantly enriches their perceptual field and the mental operations involved in the processes of construction, structuring and analysis of information.

As mentioned above, didactic materials are key elements in the construction of knowledge of each student. Their design and implementation should strongly promote meaningful learning through the interaction between teacher and student, student and activity and student with each other, establishing the basis of a highly enriching way for the development of the pedagogical act (Schwartzman and Odetti, 2013).

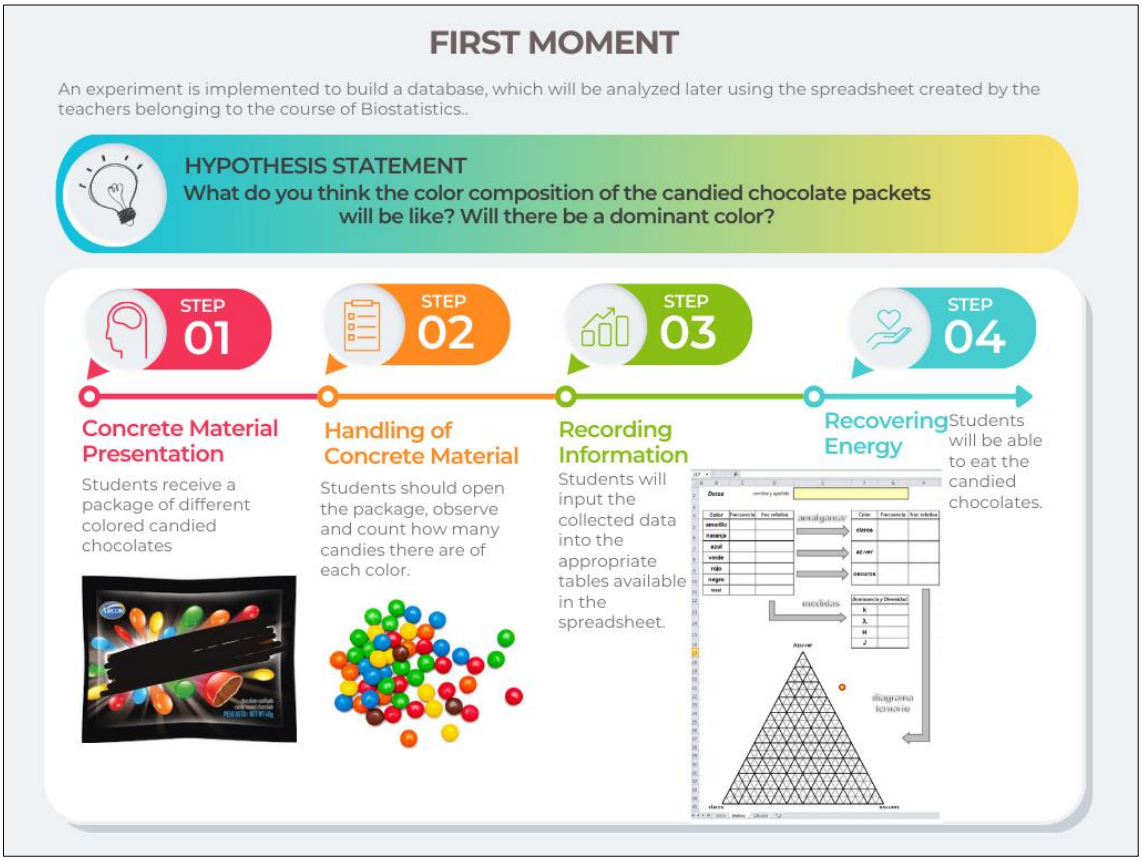
Walk through the didactic proposal

In this experience, the concept of sample composition of qualitative variables, dominance and diversity indexes, and graphic representation through ternary diagrams was addressed. This is a topic of fundamental importance in biology since it contributes to the analysis of the distribution and abundance of the different species in a specific area, making it possible to analyze and predict changes in ecosystems and to investigate the interactions between species and biodiversity in general.

The proposed teaching sequence is divided into two moments. In the first moment, the problem to be investigated is posed and the research hypotheses are formulated. In addition, the relevant variables to address the problem are determined, and an experiment is carried out to collect the necessary data to be used in the spreadsheet. At this point, the students will work with concrete material, in our case packages of colored candied chocolates of a certain brand, which in general have a total of 40 units. In each case, they will determine which colors appear and in what quantity, i.e. the frequency of occurrence of the color. Then, they will record the results of the experiment in the spreadsheet, which will provide them with the result of an amalgam, the number of mutually exclusive categories (k) and the corresponding Simpson's dominance (λ), Shannon's diversity (H) and relative diversity (J) indices.

The following table (FIGURE 1) shows the path of the first stage of the teaching sequence:

FIGURE 1: First Moment - Teaching Sequence.



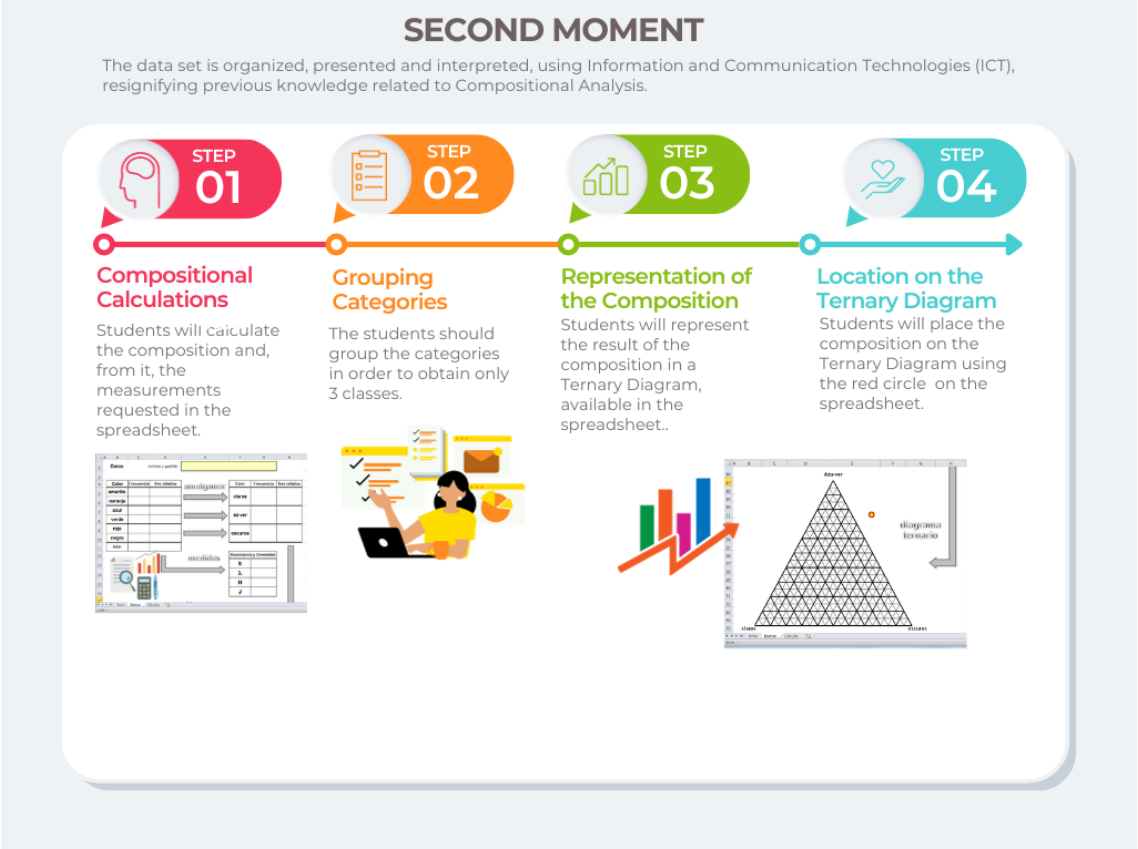
Source: own elaboration using *Canva*.

In the second stage, the data set is organized, presented and interpreted using ICTs. This approach makes it possible to give a new meaning to the theoretical foundations linked to the analysis of sample composition, encouraging the generation of contextualized and more solid conclusions. Thus, the way in which the spreadsheets simplify the calculations allows students to perform, without difficulty, amalgams with different numbers of categories and to analyze, for example, the sensitivity of the indexes to this change, comparing it in each of them. On the other hand, the graphical representation through a ternary diagram allows students to compare both the composition of the sample obtained with the composition of maximum diversity located in the center of the diagram, and to identify in

the diagram another sample that has the same dominance as the one obtained in their pack of candied chocolates.

The following table (FIGURE 2) shows the path of the second moment of the teaching sequence:

FIGURE 2: Second Moment - Teaching Sequence.



Source: own elaboration using *Canva*.

Finally, an evaluation of the contents was carried out using mainly two instruments: a written work, similar to a research report, which addresses in detail what was experienced, and a short answer written test, designed to evaluate the theoretical and practical knowledge and understanding of the subject.

The results obtained were part of a feedback process, both for students and teachers, which allowed establishing the necessary modifications to favor the teaching-learning process.

It should be noted that the predominant role of the teachers in this experience was as coordinators. As stated by Santoyo (1981), the coordinator does not teach but rather promotes learning without assuming the role of leader or director; trying at all times to ensure that there is no independence but rather interdependence among the actors in the educational activity.

Analysis of results

Based on the premise that content is not independent of the form in which it is presented (Edwards, 1997), when comparing the assimilation of the different topics through the written tests, prior to the experience, and the results obtained through the evaluation instruments selected after the implementation of this proposal, we could observe that in the routine problems and exercises, only one concept, property or capacity is reinforced. However, in this new context, both the contextualized data collection and the use of the didactic and computerized means proposed, have given rise to an opportunity to reinterpret previously worked concepts, as for example:

- ✓ Classification of variables.
- ✓ Definition of amalgamation, subcomposition and closure of categories.
- ✓ Computation of dominance and diversity indexes.
- ✓ Representation of compositional data through the Ternary Diagram.

The students presented their individual work to the teaching staff, who analyzed the development of the activities proposed in the work protocol. Subsequently, a comparative analysis of the evolution of the contents was carried out, using as an instrument a checklist before and after the implementation of the experience.

The results of the study revealed that the level of comprehension of the procedural content associated with the computation of indexes remained stable. However, a significant increase was observed in the percentage of

comprehension of the other content addressed after the implementation of the experience. It is important to highlight that the graphic recording of the content experienced a significant improvement after the practical work mediated by the spreadsheets.

The following table (FIGURE 3) shows the comparison of the percentage distributions of the appropriate responses of the students in the proposed evaluation instances, before and after the experience was carried out.

FIGURE 3: Distribution (in %) of correct answers achieved by the students before and after the development of the experience

Contents	EXPERIENCE	
	Before	After
Dominance and Diversity Index Calculation.	90	90
Development of amalgamation, closure and subcomposition operations	65	75
Representation of the composition in a Ternary Diagram	75	90
Development of appropriate answers, taking into account the theoretical framework	60	70

Source: own elaboration using *Canva*.

This instance of comparison was very important because, based on the answers given by the students, we obtained valuable elements to diagnose the degree of mastery in the analysis of the composition of samples and to redefine the approach of our teaching practices.

Advantages observed

- Resignification of the knowledge previously addressed in the classroom.
- Collaborative learning teamwork among the participants of the teaching-learning process.

- Fast and accurate acquisition of results in both analytical and graphical format.
- Efficiency in the development of activities involving data, promoting a more realistic perspective in the management of information.
- Agility and optimization in the processing and analysis of information from the use of digital resources available in the virtual classroom, such as spreadsheets.
- Potential to relate in a transversal way the development of new contents, such as Homogeneity Tests.

Difficulties observed

- Lack of familiarity with the work methodology and computer tools presented, which generated an initial resistance in the development of the proposal.
- The limited time availability of some of the students for the development of activities with these characteristics due to the intense obligatory workload in other subjects.

Conclusions

The journey through this teaching sequence has effectively integrated theory and practice, thereby enhancing the intellectual abilities of each student and transcending mere memorization skills. By using the tools provided by the spreadsheet and working with concrete material, the students acquired a solid understanding of essential concepts, such as the management of compositional data, diversity, dominance or species-specific richness, among others. This knowledge will be of great importance in their future work as biologists.

It should be noted that the use of concrete material, in this case candied chocolates, as a teaching tool is only a first step. As their education progresses, it is essential that students also develop a deeper

understanding of the underlying biological principles. However, by using creative and engaging teaching methods, we can awaken their curiosity and motivation to learn, laying the foundation for further development of their knowledge and advancing their learning of biology and statistics.

Finally, through the analysis of the results obtained, we can conclude that this pedagogical approach has enabled students in various skills related to critical and analytical thinking, including the ability to decompose information or problems for better understanding, evaluating the relevance of the data presented, interpreting information by identifying patterns and trends, deriving logical conclusions based on evidence, making informed decisions, and effectively communicating thoughts and arguments.

As a proposal for the future, the challenge is clear: to advance in this complex reality crossed by new technologies favoring collaborative learning, using telematic supports tending to connect and to enrich learning. Therefore, it is essential to continue designing work methodologies so that the study of Descriptive Statistics becomes the mathematical scaffolding that will lead us to the study of future topics pertaining to this subject, such as Homogeneity Tests, through real "virtual research laboratories".

References

DAVINI, M. C. Métodos de enseñanza. Didáctica general para maestros y profesores. (1ª. ed.). Buenos Aires: Santillana. 2008.

EDWARDS, V. La forma del conocimiento en el aula. En Rockwell. La escuela cotidiana. Méjico. F.C.E. 1997.

HENAO ALVAREZ, O. Las hojas de cálculo como herramienta didáctica. Informática Educativa v. 9, n.2, pp. 103-121,1996, UNIANDES – LIDIE, Bogotá, Colombia. Disponible en: https://hdl.handle.net/10495/3941.

HERNÁNDEZ, R. V.; RAMOS RODRÍGUEZ, E.; MARIÑO, L. F. Avance en la caracterización de modelos didácticos para mejorar el razonamiento estadístico en estudiantes universitarios. Bio-grafía, [S. l.], 2021. Disponible en: https://revistas.pedagogica.edu.co/index.php/bio-grafia/article/view/16258. Accedido el 10 jul. 2023.

HORGAN, G. W., ELSTON, D. A., FRANKLIN, M. F., GLASBEY, C. A., HUNTER, E. A., TALBOT, M., KEMPTON, R. A., MCNICOL, J. W., AND WRIGHT, F. Teaching statistics to biological research scientists. Statistician 48, p. 393– 400, 1999.

MCLEOD, S. Constructivism Learning Theory & Philosophy of Education. SimplyPsychology. Website: https://www.simplypsychology.org/. Actualizado el 15 de junio de 2023

METZ, A. M. Teaching Statistics in Biology: Using Inquiry-based Learning to Strengthen Understanding of Statistical Analysis in Biology Laboratory Courses. CBE—Life Sciences Education, v. 7, 2008, p. 317–326. Disponible en: https://www.researchgate.net/publication/23232000_Teaching_Statistics_in_Biology_Using_Inquiry-based_Learning_to_Strengthen_Understanding_of_Statistical_Analysis_in_Biology_Laboratory_Courses.

PÉREZ, J. E.; PÉREZ, I.; OJEDA G. La enseñanza de las ciencias biológicas en la Universidad. SABLE. Revista Multidisciplinaria del Consejo de Investigación de la Universidad de Oriente, v. 18, n. 2, , p. 234-240. 2006. Disponible en: https://www.redalyc.org/articulo.oa?id=427739430018

SANTOYO, R. Algunas reflexiones sobre la coordinación en los grupos de aprendizaje. Perfiles Educativos, p.3-19, 1981.

SCHWARTZMAN, G.; ODETTI, V. Los materiales didácticos en la educación en línea: sentidos, perspectivas y experiencias. En Conferencia Internacional ICDE 2011. III Foro Internacional de Educación Superior en Entornos Virtuales. 2011. Disponible en: http://www.pent.org.ar/publicaciones/materiales-didacticos-educacion-linea-sentidos-perspectivas-experiencias

SERRANO GONZÁLEZ-TEJERO, J.M.; PONS PARRA, R.M. El Constructivismo hoy: enfoques constructivistas en educación. REDIE, Ensenada, v. 13, n. 1, p. 1-27, enero 2011. Disponible en http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1607-40412011000100001&lng=es&nrm=iso. Accedido en 14 jun. 2023.

VARGAS MURILLO, G. Recursos educativos didácticos en el proceso enseñanza aprendizaje. Cuadernos Hospital de Clínicas, 58(1), p. 68-74, 2017. Disponible en http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S1652-67762017000100011&lng=es&tlng=e. Accedido el 10 jul. 2023.

Received in June 2023.

Approved in October 2023.