Reception: **May 11, 2023** | Acceptance: **April 15, 2024** | Publication: **October 1, 2024**

Econometric Determination of Factors Affecting the Performance of Dissemination Activities. The Case of the JC/CUC DAIA Science Club

—

Verónica De Jesús Romo
veronica.dejesus@ujat.mx

Alejandra Sofía Martín Hernández
alejandrasomh26@gmail.com

Jesús Antonio Pérez Vázquez
jesus.perez961111v@gmail.com

Academic Division of Basic Sciences of the Universidad Juárez Autónoma de Tabasco, Cunduacán, Tabasco

To quote this article:

De Jesús Romo, V., Martín Hernández, A. S., & Pérez Vázquez, J. A. Determinación Econométrica de los Factores que Inciden en el Desempeño de Actividades de Divulgación. El Caso del Club de Ciencias JC/CUC DAIA. *Espacio I+D, Innovación más Desarrollo, 13*(38). https://doi.org/10.31644/IMASD.38.2024.a06

— Abstract—

Scientific dissemination programs or projects, like any activity that pursues an objective, can be evaluated. The objective of this research is to evaluate the scientific dissemination activities carried out by the members of a Science Club during the most critical phase of confinement derived from the COVID-19 pandemic. These activities took place in 20 different communities, 8 towns in the state of Tabasco and 4 towns in the state of Chiapas. The evaluation research was developed by applying methodology for program evaluation and a general linear statistical model to determine the factors that influence the performance of scientific dissemination activities. The statistical model that justifies the conclusions of this research was chosen according to the criteria of correct statistical specification and statistical significance of the variables. Using a correctly specified model, we identified that the factors that had a positive and statistically significant effect on the disseminator's performance were: sex, the number of experiments successfully carried out during the activity, the level of interest of the attendees perceived by the disseminator and the number of close relatives of the disseminator with university studies in science.

Keywords:
Program evaluation, econometric modeling, determinants factors, dissemination

RESEARCH BACKGROUND

The frame of reference that we identified in the specialized literature shows that the generation of evaluations for scientific dissemination programs or projects is a relatively understudied area. Below, we briefly review some of the research in which evaluations of outreach-related programs are presented.

The work of Barahona et al. (2020) aims to investigate the impact of a set of math communication activities in public spaces on strengthening the social fabric and improving the perception of security. The attendees answered a questionnaire composed of 10 Likert-type[1] reagents and 5 categorical questions regarding their socio-demographic profile. With the data obtained, the authors performed a descriptive statistical analysis and an inferential analysis that consisted of 2 statistical methods: analysis of main components and model of structural equations. They evaluated the validity and reliability of the information collection instrument through three indicators: Cronbach's alpha, Guttman's lambda, and the Inter-class Correlation Coefficient. Through their data analysis, the authors concluded that as the project was developed, information was collected on an increase in the perception of safety and higher levels of neighborhood integration in the community, in addition, they showed evidence in favor of the communication of mathematics in public spaces contributed to improving the perception of safety and strengthening the social fabric of the Chamilpa community.

Another research that aimed to *evaluate a workshop* was presented by Pulido (2017). The workshop whose information was analyzed had as its theme the *preservation of air quality in closed places* and was developed with 128 university students from the degrees of Nursing and Business Sciences of the Universidad de Papaloapan, Oaxaca. The author conducted four Likert-type questionnaires as a means of evaluation for each construct[2] and as a measure of reliability and validity conducted a reagent discrimination analysis.[3]

Statistical analysis of student responses before and after the workshop, the author concluded that, in the nursing group, the workshop had a positive impact on most of the constructs assessed. However, in the business sciences group, the changes before and after the workshop were not significant.

Gallardo's thesis (2014) was another research work in which a study was carried out to evaluate the constitutive elements of the online educa-

1 Type of ordinal scale intended to measure the attitude of respondents on a given topic. Developed in 1932 by Rensis Likert.
2 Theoretical construction to understand a given problem.
3 Also known as item discrimination analysis.

tional process from the perspective of the students of the bachelor's degree in nursing of the SUAyED-ENEO (Open University and Distance Education System-National School of Nursing and Obstetrics) of the UNAM. The study consists of a descriptive, cross-sectional, and observational analysis carried out with information from 119 students. The statistical methodology used was a one-factor analysis of variance. The working hypothesis was that the online learning process (which would act as the dependent variable) is based on: age, sex, number of jobs, hours of study per week, basic computer skills, and previous courses online or at the student's headquarters. In the study, the author used Pearson's correlation coefficient between some variables that he considered relevant to understand the student's learning process. Finally, the author concluded that only the *student's home* variable was significant for the study.

In the research of Sánchez (2008), statistical validation of the entrance exam to the online course for reading comprehension in English was carried out and applied to 213 students. The objective of this research was to validate a construct related to a hypothetical model through mathematical models. The author applied a data collection instrument to observe and identify the factors that determine efficient learning in text comprehension. The statistical methodologies used in the research were: representative mathematical models of item response theory (IRT) and Cronbach's alpha internal consistency analysis, causal models, and multivariate statistical techniques such as factor analysis, path analysis, multiple regression analysis, and structural equation modeling. The author evaluated the assumptions of normality, homoscedasticity, and linearity of the applied methodology to ensure convergence towards a feasible solution. Finally, according to the previous analysis and due to the quality of goodness of fit, the author concluded that there was insufficient evidence to affirm that the model does not collect the variability in the data.

Finally, Álvaro et al. (1990), in their work *Hacia un modelo causal del rendimiento académico*, evaluate some models that influence academic performance. The research was developed around two objectives: to arrive at an explanatory model of academic performance and to choose the most appropriate analysis technique to test that model.

Consequently, they used statistical methods for data analysis such as principal component analysis, maximum likelihood factor analysis, path analysis, and analysis of LISREL (Linear Structural Relations) models, in addition to applying goodness-of-fit measures such as χ^2, Goodness-of-fit Index GFI, among others. Through the use of different exploratory and confirmatory analyses, Álvaro et al. (1990) justified a reduction from 89 variables to 14. Thus, the model that they determined best explained the phenomenon of interest was the one in which the values of the adjustment indicators (χ^2 and GFI) met the required limits and the multiple correlations squared of the two variables used were very high.

From their statistical analysis, the authors highlighted 3 conclusions: i) the best predictor for performance is aptitudes; ii) through a general aptitude (composed of a verbal factor, another numerical factor, and logical reasoning) performance in mathematics can be better predicted than in language; iii) the cultural level of the parents has a causal relationship with aptitudes, that is, a high cultural level in the family is conducive to greater aptitude development, consequently, the expected performance in basic instrumental areas such as language and mathematics will be higher.

CONCEPTUAL FRAMEWORK OF THE PROGRAM EVALUATION PROCESS

Evaluation can be defined as the process of systematically collecting information on the activities, characteristics, and results of a program (set of steps that are carried out to achieve an objective) to reduce uncertainty, improve effectiveness, and decisions regarding the achievement of objectives (Jean-Michel & Benot, 2017).

On the other hand, according to the glossary of the main terms on evaluation and results-based management of the Organization for Economic Cooperation and Development (OECD, 2010), the evaluation of a project, program, or policy in progress or concluded is the systematic and objective appreciation of its design, its implementation, and its results. Therefore, we can define evaluation as the process of systematically collecting information on the activities, characteristics, and results of programs, to reduce uncertainty, and improve effectiveness, and decision-making.

Note that an evaluation not only analyzes whether or not the program is effective but also provides information to determine if the program is the most appropriate way to achieve its objectives and if there are other elements to consider.

When carrying out an evaluation process for a program, there are key points that we must keep clear, such as the purpose of the evaluation, the time at which the program will be evaluated, the model with which it will be evaluated, the instrumentation that will assist said evaluation, the institution or professionals in charge of carrying out the evaluation and the reference framework within which the evaluation of the program will be carried out. The points mentioned above are intended to outline a feasible, methodical, objective, transparent, and verifiable evaluation process.

On the other hand, Jean-Michel and Benot (2017) describe a program in terms of needs, design, inputs, and outputs, short- and long-term results. In addition, an assessment program can be represented as a sequence of four phases:

I. **Context analysis**: involves gathering information about what constitutes the problem, who it affects, and how they perceive it, to determine their needs. It relies on descriptive and inferential statistical tools to point out issues that need to be addressed. Actions:

- Describe the social, economic, and institutional context in which the program will be implemented.
- Identify needs, determine their scope, and define the target population (cross-sectional, longitudinal, or panel data study).
- Make a distinction between descriptive statistics and inferential statistics, to identify patterns in the sample.
- Distinguish between univariate, bivariate, and multivariate analyses, depending on the number of variables examined.
- Visualize the status of the population, if the identified needs were met.

II. **Ex-ante evaluation**: It tries to assess aspects that allow us to fine-tune the decisions around the implementation of the program. When a program or project has the evaluation, it influences the improvement of the decision-making on its implementation, in the identification of areas of improvement that, if not observed and corrected, could generate unnecessary costs and inefficiencies in the implementation stage.

At this stage, it is critical to determine the goals and objectives of the program before conducting an evaluation. Alternative strategies for addressing program objectives must be compared based on all relevant dimensions (technological, institutional, environmental, financial, social, and economic). The methods can be:

- Financial Assessment
- Budget impact analysis
- Cost-benefit analysis.
- Cost-effectiveness analysis.
- Multi-criteria decision analysis.

III. **Implementation**: This stage is responsible for designing a monitoring system to help project leaders or managers implement the program.

The construction of a well-documented data management system is essential, for which indicators can be used to measure inputs, outputs, or relate resources to services-products:

- Media indicators (operating expenses, donations received, number of agents).
- Performance indicators (number of beneficiaries or users).
- Management and accounting indicators (operating expenses per user, number of agents per user).

These indicators can be used to report progress and alert program managers to issues, and can also be used subsequently for evaluation purposes.

IV. **Ex-post evaluation**: It seeks to fine-tune elements of the evaluated program, which can be grouped according to the particular type of evaluation in question.

This stage measures what has happened as a direct result of the execution of the program. Consequently, effectiveness has to do with the level of outcomes and whether or not the intervention was successful in reaching the desired goal. This phase also identifies the main factors behind success or failure. Commonly used assessment techniques are:

- Random case follow-up.
- Benchmarking analysis.
- Quasi-experiment.

It is important to mention that we must not forget that:

1. The choice of the method to be used depends mainly on the context of the analysis. For example, random assignment is not always possible in legal, technical, or ethical terms.
2. The choice of the time frame in which to conduct the evaluation is a difficulty since the information needed to assess program outcomes is sometimes available only several years after program completion.

Generally, the results are classified as:

- Short term: if they are immediate effects on the individuals' state.
- Long-term: environmental, social, and economic changes

PHASES OF PROGRAM EVALUATION. THE CASE OF WORKSHOPS AS A MEANS OF SCIENTIFIC DISSEMINATION

The following is an application of the four phases of evaluation: context analysis, evaluation ex ante, implementation, and evaluation ex post, to a student science outreach program.

The *context* of the scientific dissemination activities carried out by a group of students who make up the Science Clubs (CUC's) of the Universidad Juárez Autónoma de Tabasco [UJAT] has been developed for approximately 15 years, through workshops aimed at audiences of different educational levels in university spaces, educational campuses, museums, or public spaces in the state of Tabasco. These activities represent a work of university social retribution, since, in many cases, the members of the clubs perform their social service in scientific dissemination activities to promote and foster scientific culture in society.

The *natural* evaluation of a scientific dissemination program represented by the activities of the CUC's, consists of measuring the impact that these activities have on the scholastic or integral performance of the beneficiaries of the program, that is, of the audiences, or rather, on the young university students who provide these activities. Therefore, it is possible to measure the effect that science outreach activities have on the people who receive the workshops and on the lecturers themselves. It is also possible to carry out an evaluation to identify the factors that influence the performance of the lecturers.

Before implementing a program, the direction of the desired outcome should be defined in a general way, such as demonstrating that outreach activities are a pillar of the educational institution's substantive activities. Or, a specific objective, such as increasing general knowledge on a specific topic.

Let's suppose that the program's *objective* is to disseminate knowledge about a particular subject in a didactic and entertaining way for children.

As part of the *ex-ante* evaluation, it is advisable to think about whether:

- Are the selected outreach strategies consistent with the overall program objective?
- Are outreach strategies suitable for children?
- Will the activity generate new scientific knowledge in the public?
- Do the strategies cover all program objectives?
- Are there any programs with the same or similar objectives?

To achieve a more efficient program, it is advisable to analyze:

- Are the necessary resources available to develop the outreach activity?
- Are existing resources adequate?
- Is the outreach program profitable for the academic development of the participants?
- Will the cost of the outreach program be commensurate with the effectiveness?

When *implementing* the program, it is important to follow each phase of the previously designed strategy, as well as to be careful when collecting the information that will be used to construct the indicators.

Subsequently, as part of the *ex-post* evaluation, we must measure the effects that the program had on the higher education institution's students or audiences, using the method(s) that best suit the program and the type of evaluation. Therefore, it is necessary to identify in the outreach program what the short-term and long-term results are:

Short-Term Results

- Were any effects identified among the outreach students (lecturers) on their academic performance?
- Did the students achieve educational growth?
- Did the students achieve personal growth?
- Were the resources sufficient for the program?
- Did the program meet its goals?

Long-Term Results:

- Is science dissemination a substantive activity for the institution?
- Did the program manage to produce and/or increase knowledge?
- Did the activity contribute to improving the public's relationship with the topics covered?
- Did the program awaken in the students the vocation of science communicator?

Finally, the responsible authorities will make the most appropriate decisions (instrument, strengthen, continue; as the case may be), based on the results of the program evaluation. That is why, from an institutional perspective, evaluations are valuable as an instrument for proper decision-making.

EVALUATION OF OUTREACH EFFORTS

The JC/CUC DAIA Science Club

The Science Club "Youth for Science" of the DAIA of the UJAT began its activities in 2006, and its main activity is to carry out scientific dissemination activities. The club is typically composed of students from degrees in Architecture, Civil Engineering, Electrical and Electronic Engineering, Chemical Engineering, and Electrical Mechanical Engineering who fulfill their social service, professional internship, or volunteering. The activities carried out at the club are dissemination workshops, participation in scientific events, and research projects, among others.

Since the Science Club was founded, it has made significant achievements in outreach. However, they did not have the opportunity to assess the impact of their activities on the population or the club participants themselves. To meet this need, we carried out a first evaluation exercise, to identify the factors that influence the good performance of the members of the Science Club in science outreach activities. The evaluation presented was part of a work carried out as a professional internship and subsequently presented by two of the authors as a thesis work to obtain the degree of Bachelor in Actuary.[4]

The framework for the evaluation was the workshop *¡Más fuerte qué Hercules!*, which consisted of the presentation of three experiments related to the surface tension of water, aimed at a child audience.

Due to sanitary restrictions, the workshop was conducted by each member of the club in the community in which they live, with a child audience. This resulted in the presentation of the workshop in 20 different communities, which belong to 8 municipalities in the state of Tabasco and 4 municipalities in the state of Chiapas. Figure 1 shows some of the club members conducting the ¡Más fuerte que Hércules! Workshop.

Figure 1. ¡Más fuerte que Hércules! workshop

4 Actuarial professionals are those who apply mathematical, statistical, economic, and computational methods to the calculation of financial risks arising from uncertainty when they are covered by a contract, such as insurance, bonds, pensions, social security, labor liabilities, and credit, investment, and derivative financial instruments. However, their skills can be oriented to other tasks such as the evaluation of different types of programs.

METHODOLOGY

Instrument design and database creation

Statistical information to model the performance of Science Club members in the "¡Más fuerte que Hércules!" outreach workshop was collected through a Google Forms questionnaire consisting of 3 sections. The first contained 4 questions dedicated to the collection of general data, such as gender, age, or grade point average in their program of study. The second was focused on measuring the degree to which the lecturers perceived the performance of their presentation, it was integrated with 18 questions, 8 of them were Likert-type with 4 values (Likert-type surveys are psychometric instruments where the respondent must indicate their agreement or disagreement on a statement or item, which is done through an ordinal and unidimensional scale (Matas 2018), which sought to extract information about how the lecturers perceived the audience and their performance.[5] The third and last section consisted of 6 questions, which sought to collect data on the academic profile of the family members of the surveyed Science Club members.

Table 1 below shows the 28 variables generated to study the performance of the workshop participants in their presentations, of which 27 were obtained from the questionnaire, and were treated as independent variables. The remaining variable, labeled CaliTaller, is the score obtained as a qualification in the workshop and was treated as the dependent variable in our modeling. This rating was awarded by academics with experience in disclosure based on the evidence, photos, or videos reported. Finally, the study was conducted with information from 23 lecturers and members of the club.

5 The reader may note that we do not present the results of Cronbach's alpha, which allows us to quantify the level of reliability of a measurement scale for the unobservable magnitude constructed from the n observed variables. This is because it is desirable, to create a reliable scale of a quality that is not directly observable, that the items be highly correlated with each other. In our case, the n observable variables we use are used to model the variability of the response variable that is observable and has a measurement scale. Therefore, we do not use Cronbach's alpha as a measure of the reliability of measurement scales.

Table 1
28 variables were generated to study the lecturer's performance in their presentations.

Variable	Variable name	Description
Y	**Workshop grade**	*Grade obtained at the end of the workshop awarded by competent personalities with experience in dissemination.*
		Section 1: Personal information
	Lecturer's name	*Lecturer's full name*
X_1	**Gender**	*Lecturer's gender*
X_2	**Age**	*Lecturer's age at the time of answering this survey.*
X_3	**PromCali**	*Lecturer's grade point average*
		Section 2: Perception of the lecturer's presentation performance.
X_4	**PE**	*Lecturer's current study program.*
X_5	**DomTema**	*Level of mastery that the lecturer felt about the workshop subject.*
X_6	**Autoper**	*Adjective that best described the lecturer's perception of himself/herself at the time of the workshop.*
X_7	**Seg**	*Level of security felt by the lecturer at the time of the workshop.*
X_8	**Nervio**	*Level of nervousness felt by the lecturer while conducting the workshop.*
X_9	**DisTaller**	*Level of fun or enjoyment that the lecturer felt while doing the workshop.*
X_{10}	**Ninv**	*Number of children invited to the workshop by the lecturer.*
X_{11}	**Nasis**	*Number of children attending the workshop.*
X_{12}	**Niños**	*Indicates if the lecturer frequently spends time with children.*
X_{13}	**TiemEst**	*Minutes that the lecturer spent studying before the workshop.*
X_{14}	**EnsExp**	*Number of times the lecturer rehearsed the experiments before conducting the workshop.*
X_{15}	**TiemEns**	*Minutes that the workshop leader spent rehearsing the experiments before the presentation.*
X_{16}	**ExpExitosos**	*Number of experiments that the lecturer successfully performed during the workshop.*
X_{17}	**InteresAsist**	*Level of interest of the attendees during the workshop as perceived by the lecturer.*
X_{18}	**GustoAsist**	*Level of satisfaction of the attendees with the workshop according to the lecturer's perception.*
X_{19}	**NivSatis**	*Level of satisfaction that the lecturer obtained from the attendees' response while conducting the workshop.*
X_{20}	**TiemEvi**	*Minutes that the workshop leader invested in preparing the workshop evidence.*
X_{21}	**Tsufi**	*Indicates whether the workshop leader felt he/she had enough time to organize the workshop.*
		Section 3: Academic profile of the lecturers' family members
X_{22}	**Beca**	*Indicates whether the lecturer had a scholarship during the period in which the workshop was held.*
X_{23}	**EstPadre**	*Degree obtained by the lecturer's father or guardian.*
X_{24}	**EstMadre**	*Degree obtained by the lecturer's mother or guardian.*
X_{25}	**FamCDuras**	*Number of the leader's close relatives who have university studies related to the hard sciences.*
X_{26}	**FamCSyH**	*Number of the lecturer's close relatives who have university studies related to the social sciences and humanities.*
X_{27}	**FamCEyA**	*Number of the lecturer's close relatives who have university studies related to administrative economic sciences.*

STATISTICAL METHODOLOGY

In this research, we apply *the theory of probabilistic reduction* (Spanos, 1986) developed within the framework of the probabilistic approach of econometrics. This consists of rigorously evaluating the assumptions about the vector of observable variables, to obtain a simplified and acceptable probabilistic structure. This method consists of i) defining the experiment design that relates a theoretical model to the data in a probabilistic scheme through the specification of the statistical model; ii) the verification of the statistical assumptions underlying the specification; and iii) the respecification of the model to establish a correctly specified model to contrast the hypotheses to establish statistically reliable conclusions in the light of the data.

Specifically, the study was performed using a linear regression model (general linear model), given by the equation, $y_i = \beta_o + \boldsymbol{\beta}'\boldsymbol{x}_i + u_i,\ i \in \boldsymbol{n}$, whose probabilistic structure is presented in Table 2 using two distinct approaches.

Table 2
Probabilistic structure with two different approaches

	Supuestos del modelo de regresión lineal	
	Enfoque probabilístico	**Enfoque tradicional**
1 normalcy	$(y_i \mid X_i = x_i) \sim N(\cdot,\cdot)$	$u_i \sim N(\cdot,\cdot)$
2 linearity	$E(y_i \mid X_i = x_i) = \beta_o + \boldsymbol{\beta}'x_i$	$E(u_i \mid X_i = x_i) = 0$
3 homoscedasticity	$Var(y_i \mid X_i = x_i) = \sigma^2$	$Var(u_i \mid X_i = x_i) = \sigma^2$
4 constant parameters	$\beta_o, \boldsymbol{\beta}', \sigma^2$	$\beta_o, \boldsymbol{\beta}', \sigma^2$
5 independence	$\{(y_i \mid \boldsymbol{X}_i = \boldsymbol{x}_i), i \in I\}$	$\{(u_i u_s \mid \boldsymbol{X}_i = \boldsymbol{x}_i) = 0,\ i \neq s, i \in I\}$

The probabilistic approach of the multiple linear regression model, which was based on De Jesús (2016), is intended to highlight why it is necessary to satisfy each assumption in Table 2 with the modeled data and the implications of not complying with them.

Probabilistic approach to multiple linear regression modeling

With the following economic relationship of interest:

$$Y_i = f\left(X_{1,i}, X_{2,i}, \dots, X_{k,i}\right), \qquad i = 1, 2, \dots, n \qquad \dots (1)$$

where Y_i denotes the dependent variable, and $X_j\ j = 1, 2, \dots, k$ denotes the j-th independent variable. The following shows how the statistical model of multiple linear regression

$$y_i = \alpha + \boldsymbol{\beta}'\boldsymbol{x}_i + u_i, u_i \sim N(0, \sigma^2), \dots(2)$$

with parameter vector $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}, \widetilde{\boldsymbol{\Sigma}})$ and under the assumptions in Table 2, is a parameterization of the joint density of all observable variables $X_{1,i}, X_{2,i}, \dots, X_{k,i}$ under the following assumptions: normal distribution, independence, and identical distribution.

By the assumption of independence and identical distribution of $\boldsymbol{X}_i$, the assumption of normal distribution of $\boldsymbol{X}_i$, and exogeneity of the variables $\boldsymbol{X}_{i,j}$ for $j = 1,2, \dots,k$, we know that Y_i given the values of the random vector $\boldsymbol{X}_i = \boldsymbol{x}_i$ is distributed as a normal random variable,

$$Y_i \mid \boldsymbol{X}_i = \boldsymbol{x}_i \sim N_m\left(\alpha + \boldsymbol{\beta}'\boldsymbol{X}, \ \widetilde{\boldsymbol{\Sigma}}\right) \quad \dots(3)$$

where $\alpha = \mu_Y - \boldsymbol{\beta}'\boldsymbol{\mu}_X, \boldsymbol{\beta} = \boldsymbol{\Sigma}_{XX}^{-1}\boldsymbol{\Sigma}_{XY}$ and $\widetilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{YY} - \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{XX}^{-1}\boldsymbol{\Sigma}_{XY}$.

This result shows that there is a linear relationship between Y_i and $\boldsymbol{X}_i$, of the following type

$$Y_i = E(y_i \mid \boldsymbol{X}_i = \boldsymbol{x}_i) + u_i, i = 1, 2, \dots, n,$$

where the error term $u_i = y_i - E(y_i \mid \boldsymbol{X}_i = \boldsymbol{x}_i)$ is not autonomous, its probabilistic structure is completely determined by (3). The assumptions of the statistical model can be expressed in terms of u_i, as in Table 2.

As usual, to determine the most probable values of the parameters of the statistical model, $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}, \widetilde{\boldsymbol{\Sigma}})$, , when the random process $\{\boldsymbol{X}_i\}_{i=1}^n$ has been observed, we maximize the logarithm of the likelihood function concerning $\boldsymbol{\theta}$. But since the likelihood function is the joint density of the observed process $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n)$ conditional on $\boldsymbol{\theta}$ then:

$$\mathrm{L}(\boldsymbol{y}, \boldsymbol{x} \mid \boldsymbol{\theta}) = \prod_{i=1}^{n} D(y_i \mid \boldsymbol{x}_i; \psi_1)$$

where $D(y_i \mid \boldsymbol{x}_i; \psi_1)$ is the multivariate normal density given by (3). Therefore, the probabilistic properties of the maximum likelihood estimators, of any test statistic, and goodness-of-fit measure, will be completely determined by (3).

If the linear regression model fails to meet any of the probabilistic assumptions in Table 2 *vis-à-vis the data*, then the density $D(y_i \mid \boldsymbol{x}_i; \psi_1)$ will be misspecified and will invalidate the probabilistic properties of any statistic derived from it. This not only implies that the statistical inference, goodness-of-fit measures, and forecasts made from the statistical model are unreliable, but also that the entire model will be in question as a *process generating the observed data*.

Note that assumptions 1-3 in Table 2 depend on the normal distribution assumption of X_i. However, it is also one of the most difficult assumptions to meet. According to Hoover et al. (2009); and Hoover (2012), the hypothesis of multivariate normality of economic data is not a characteristic that we expect to be met, it is rather a hypothesis that allows us to ensure that both *unusual* events, which are adequately described by the normal distribution, and *unusual* events that tend to fall outside the range of the normal distribution, have been considered. Such extraordinary events are often the cause of skewness or excess kurtosis in the distribution of the data and, therefore, of the rejection of the normality assumption of u_i. In other cases, inadequate modeling of such events may be the cause of autocorrelation among errors, bias in estimators, and inaccuracy in inferences. The assumption of normality also depends on the linearity, of variables and parameters, of the statistical model. Therefore, if it is detected that the normality assumption is not met, the functional form of the model will also be questioned.

In the probabilistic approach to econometrics, the error term should capture all the factors influencing the phenomenon that were not considered by the empirical model, of course, these factors should be many more than those considered in the model. So, if such factors are independent, the errors could be distributed approximately as normal random variables by the central limit theorem.

The independence assumption of $\boldsymbol{X}_i,\ i = 1, 2, \ldots, n,$, in the observable process of economic variables, has also been questioned (assumption 4 of Table 2) it is very frequent to observe that the process $\{\boldsymbol{X}_i\}_{i=1}^{n}$ shows some kind of dependence especially in the analysis of time series, where the heterogeneity of these variables induces that both the expectation and the variance and covariance matrix of the observable process are a function of time.

Thus, the only correct strategy to achieve valid and reliable inferences is to adopt a statistical model whose probabilistic assumptions are valid *vis-à-vis* the data before making any inference. Therefore, before testing the hypotheses about the phenomenon of interest with the model (2), it is necessary to *verify* that the statistical model satisfies the complete list of probabilistic assumptions underlying the chosen specification with the sample data $\boldsymbol{x}$. Such verification ensures the reliability of any inference based on the model. Note that when the model is **incorrectly** specified, in the sense that any of the model assumptions were rejected, then the distribution $D(\cdot\,;\boldsymbol{\theta})$ will be misspecified for sample x and will invalidate the distribution of estimators, test statistics, and any statistics obtained from it.

Model (2) should be *re-specified*, choosing a new specification that takes into account regularities in the data not explained by an incorrectly specified model. Having done this, once again it must be evaluated that the data, $\boldsymbol{x}$, does not reject the assumptions of this new specification. This procedure should be repeated until a specification is identified that satisfies all assumptions with the data, from which reliable inferences can be made.

RESULTS

After analyzing different statistical methods and models with the data from the 23 lecturers for the 27 variables obtained, we determined, by the criteria of correct statistical specification and significance of the variables, that the general linear specification given by the following equation captured well the variability of the variable Y_i which measures the lecturers' performance.

$$\hat{Y}_i = 6.5522 + 0.6934x_{1i} + 0.3925x_{16i} + 0.3376x_{17i} + 0.2241x_{25i}$$

where the statistically significant variables were:
x_1: Lecturer's gender
x_{16}: Number of successful experiments during the lecturer's presentation.
x_{17}: Level of interest perceived by the lecturer from the attendees.
x_{25}: Lecturer's relatives with university studies in hard sciences (Mathematics, Physics, Chemistry, even Engineering).

We observed that the effect of all the variables included in the model is positive on the lecturer's performance.

Before interpreting the estimated coefficients and evaluating the significance of the variables in the model, we proceed with the evaluation of the correct specification of the estimated model. That is, to verify whether it complies with the assumptions of the general linear model.

Verification of correct specification

The correct specification of the model was determined using statistical tests to evaluate the normal distribution of the estimated errors, the *hetero* and *hetero-X* tests that evaluate whether the errors have constant variance, and, finally, the *reset2* test that evaluates whether the linear relationship does not omit relevant variables.[6]

At the significance level α=0.05, i.e., with 95% confidence we cannot reject the null hypothesis that the errors follow a normal distribution. In other words, there is sufficient statistical evidence not to reject the null hypothesis that the data follow a normal distribution (*p-value*=0.7605).

To evaluate whether the estimated model errors are homoscedastic, i.e., that the variance of the estimated errors is constant, the following tests were performed: *hetero* and *hetero-X*. Once both tests were performed, we obtained in the *hetero* test a *p-value=0.309* and in the *hetero-X* test a *p-value=0.435*. In both cases, we can conclude that there is sufficient statistical evidence not to reject the null hypothesis that the estimated errors of the regression model have constant variance.

Finally, the *reset test23* was performed to evaluate whether the linear model of the nonlinear combinations of the explanatory variables could explain the response variable. The null hypothesis of the test is that the model is well-specified. The *p-value* of the test was *p=0.291*. With which we can conclude that there is sufficient statistical evidence not to reject the null hypothesis, that is, the linear statistical model is correctly specified.

The graphs in the following Figure 2 shows in the first panel, the values that the variable y (performance) and the estimated values $\hat{y}$ (estimated performance), the model residuals graph, and the residual histogram. It can be seen, as with the statistical tests, that the fit using the linear model is statistically adequate.

In summary, we observe that our estimated model meets the statistical and probabilistic assumptions of the general linear model, so it is statistically valid and solid to perform inferential analyses such as that of the individual and joint significance of the explanatory variables.

6 All estimations and statistical evaluations were performed in Stata software. Conclusions from statistical tests of misspecification are taken under the decision criterion, according to which if the p-value is greater than the significance level α=.05, the null hypothesis that the probabilistic assumption is satisfied with the model data is not rejected. The p-value is calculated as $P(X \geq x)$ where X denotes a random variable that has a probability distribution equal to that followed by the test statistic under the null hypothesis and x denotes the observed value of the test statistic obtained. Note that, depending on the assumption evaluated, the test statistic follows a specific probability distribution.

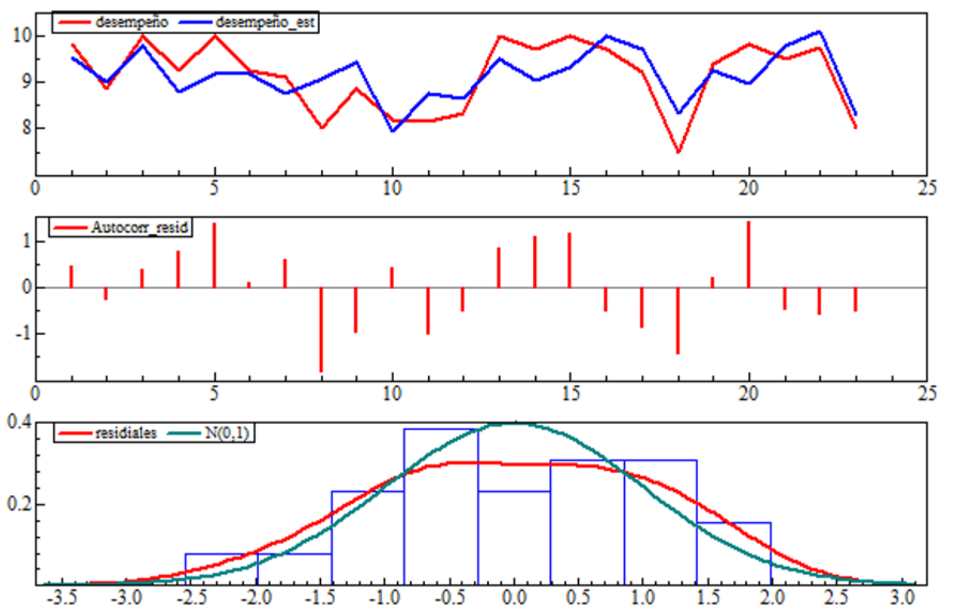

Figure 2. Graphic analysis of correct specification

To verify the significance of the regression parameters, we performed the joint significance test and the individual significance tests in the Stata software.

In the joint significance test, a test statistic was obtained whose statistic follows Fisher's *F* distribution, which took the value of *F=4.70* and a *p-value=0.009*,[7] these data allow us to conclude with 95% confidence that there is sufficient statistical evidence to reject the null hypothesis that the parameters estimated in the model are equal to zero.

On the other hand, the results of the individual significance test of each variable in the regression show that the t-test statistics are outside the non-rejection region of the null hypothesis, [-1.73, 1.73], at 95% confidence and the *p-values* are less than 0.05. Therefore, there is sufficient statistical evidence to reject the null hypothesis that the estimated parameters are equal to zero, that is, the variables x_1, x_{16}, x_{17} and x_{25} are significant for the model.

Therefore, the estimated model indicates that: the measure of performance of the lecturer (rating obtained from each lecturer) has a variation of

7 In the case of the global hypothesis test, the test statistic is calculated as follows

$$F = \frac{(Suma\ de\ cuadrados\ de\ la\ regresión)/(k-1)}{Suma\ de\ cuadrados\ de\ los\ erroes/(n-k)}$$

where *n* is the number of observations and *k* is the number of parameters in the regression equation. In this case, the *p-value* is calculated as *P(X≥x)* where *X* denotes a random variable having a probability distribution *F* with degrees of freedom *(k-1,n-k)* and *x* denotes the value of the test statistic obtained. The interval of values of the t distribution with 18 degrees of freedom for which the null hypothesis, H_0:β_i=0, is not rejected at 95% confidence is defined as all those values between the quantile 0.025 and the quantile 0.975 of that distribution, i.e., the interval [-1.73, 1.73].

0.34 when the perception of the lecturers about the interest of the attendees improves, in addition, a variation of 0.22 in the rating of the workshop's performance for each lecturer's close relative who has university studies related to hard sciences. A variation of 0.39 is also observed on the lecturer's performance measure for each successful experiment performed during the presentation. Finally, there is a difference in favor of the lecturers, that is, there is statistical evidence that performance improves by 0.69 when the workshop is run by a woman.

FINAL THOUGHTS

The research we present is a first statistical exercise to identify the variables that influence good performance in Science Club outreach activities. Likewise, the review of the specialized literature on dissemination allows us to recognize that research constitutes a frame of reference to generate information on potential factors and factors that, according to our model, have a statistically significant impact on the performance of dissemination activities.

Likewise, we emphasize that the variables that do not appear in the statistical model were not statistically significant in our model. This does not mean that they generally do not influence performance in outreach activities. Rather, it means that, before ruling out its possible influence on other evaluation exercises, it is necessary to pay special attention to measuring and quantifying such influence.

In this sense, variables such as enjoying the presentation of the workshop, the level of nervousness, the mastery of the topic, or the time spent rehearsing the experiments; which in principle seemed to us that they should exert some influence, were not statistically significant in the model. We assume that this fact constitutes an area of opportunity to capture its influence in future evaluation exercises.

Additionally, we could refine the method of information collection and the expansion of our exercise to other outreach groups or consider more than one outreach activity, which would give us the possibility of implementing other statistical modeling methodologies.

The evaluation of the performance of the activities of the JC/CUC DAIA Science Club is evidence of an initiative to recognize and enhance the factors that would improve the performance of the members of the club in scientific outreach work and to reassess the benefits of outreach as a substantive activity within universities.

REFERENCES

Álvaro Pag, M., et al. (1990). *Hacia un modelo causal del rendimiento académico*. CIDE.

Barahona, I., et al. (2020). Matemáticas, espacios públicos e integración vecinal. El caso de Cuernavaca (México). *JCOM – América Latina, 3*(2).

De Jesús Romo, Verónica. (2016). ¿Ciencia económica o el arte de hacer economía? Metodología científica y replicaciones en economía. *Investigación económica, 75*(296), 73-110.

Gallardo Santamaria, L. (2014). *Evaluación en el proceso educativo en línea desde la perspectiva de los alumnos de Licenciatura en Enfermería Sistema Abierto y Educación a Distancia*. [Tesis de Maestría]. UNAM. https://repositorio.unam.mx/contenidos/evaluacion-del-proceso-educativo-en-linea-desde-la-perspectiva-de-los-alumnos-de-licenciatura-en-enfermeria-sistema-a-121061?c=rVEPLZ&d=false&q=*:*&i=3&v=1&t=search_0&as=0

Hoover, K. D., Johansen, S., y Juselius, K. (2008). Allowing the data to speak freely: The macroeconometrics of the cointegrated vector autoregression. *The American Economic Review, 98*(2), 251-255.

Hoover, K. D. y Juselius, K. (2012). Experiments, Passive Observation and Scenario Analysis: Trygve Haavelmo and the Cointegrated Vector Autoregression. *Univ. of Copenhagen Dept. of Economics Discussion Paper*, (pp. 12-16).

Jean-Michel, J., y Benot Le, M. (2017). *Statistical Tools for Program Evaluation. Methods and Applications to Economic Policy, Public Health, and Education*. Springer International Publishing.

Martín Hernández, A., y Pérez Vázquez, J. (2022). *Determinación econométrica de los factores que influyen en el desempeño en actividades de divulgación científica de los integrantes del club de ciencias "JC/CUC DAIA, UJAT"*. [Tesis de Licenciatura] UJAT.

Organización para la Cooperación y el Desarrollo Económico [OCDE]. (2010). *Glosario de los principales términos sobre evaluación y gestión basada en resultados*. Francia. https://www.comminit.com/content/glosario-de-los-principales-términos-sobre-evaluación-y-gestión-basada-en-resultados

Pulido Criollo, F. (2017). *Evaluación de un taller de educación para mejorar la calidad del aire en dos grupos universitarios*. [Tesis de Maestría]. UNAM.

Sánchez Delgado, D. (2008). *Validación y evaluación estadística del examen de colocación al curso en línea para la comprensión de lectura en inglés*. [Tesis de Licenciatura]. UNAM.

Spanos, A. (1986). *Statistical Foundations of Econometric Modelling*. University Press, Cambridge.