

Test time assessment on student's performance of statistics subjects by generalized linear models

Avaliação do tempo de teste no desempenho de alunos de disciplinas relacionadas à estatística por meio de modelos lineares generalizados

Evaluación del tiempo de prueba sobre el desempeño de los estudiantes de disciplinas relacionadas con la estadística mediante modelos lineares generalizados

Received: 07/05/2021 | Reviewed: 07/09/2021 | Accept: 07/09/2021 | Published: 07/21/2021

Esttefani Duarte Brum

ORCID: <https://orcid.org/0000-0003-1239-2366>
Municipal Department of Education and Sport, Brazil
E-mail: esttefaniduarte@hotmail.com

Gilberto Rodrigues Liska

ORCID: <https://orcid.org/0000-0002-5108-377X>
Federal University of São Carlos, Brazil
E-mail: gilbertoliska@ufscar.br

Alisson Darós Santos

ORCID: <https://orcid.org/0000-0002-2082-6999>
Federal University of Pampa, Brazil
E-mail: alissonsantos@unipampa.edu.br

Abstract

Can the time it takes a student to complete a test influence his / her performance? To answer this question, the logistic regression model was considered. In its development, evaluation was considered as a way of quantifying the performance of a student reflecting his degree of knowledge in a given content. For this we use records of the initial and final moments when developing an evaluation. The records of time spent were obtained from five different undergraduate classes, with subjects taught by the same teacher, with the same theoretical content, at the same university. The results confirm statistically that each additional minute that the student remains taking the test, implies in greater chances of obtaining good performance, as well as differences of performance between the feminine and masculine genders, although not statistically different, demonstrating that feminine students have greater chances of reaching the average. The model also confirms, according to the odds ratios that during the evaluations the students' performance decreases, having the best score in the first test. Through the references consulted, we understand that the difference in the grades of each student is influenced by several factors, the result of their own experiences.

Keywords: Logistic regression; Time under evaluation; Odds ratio; Educational indicators.

Resumo

O tempo que um aluno leva para concluir um teste pode influenciar no seu desempenho? Para responder a essa questão, foi considerado o modelo de regressão logística. No seu desenvolvimento, a avaliação foi considerada como uma forma de quantificar o desempenho de um aluno refletindo seu grau de conhecimento em determinado conteúdo. Para isso utilizamos registros dos momentos iniciais e finais no desenvolvimento de uma avaliação. Os registros dos tempos despendidos foram obtidos em cinco turmas diferentes de graduação, com disciplinas ministradas pelo mesmo professor, com o mesmo conteúdo teórico, na mesma universidade. Os resultados confirmam estatisticamente que cada minuto a mais que o aluno fica fazendo a prova, implica em maiores chances de obtenção de bom desempenho, bem como diferenças de desempenho entre os gêneros feminino e masculino, embora não sejam estatisticamente diferentes, demonstrando que os alunos do sexo feminino têm maiores chances de atingir a média. O modelo também confirma, de acordo com as razões de chances, que durante as avaliações o desempenho dos alunos diminui, tendo a melhor pontuação no primeiro teste. Pelas referências consultadas, entendemos que a diferença nas notas de cada aluno é influenciada por diversos fatores, fruto de suas próprias experiências.

Palavras-chave: Regressão logística; Tempo em avaliação; Razão de chances; Indicadores educacionais.

Resumen

¿Puede el tiempo que le toma a un estudiante completar una prueba influir en su desempeño? Para responder a esta pregunta se consideró el modelo de regresión logística. En su desarrollo, la evaluación se consideró como una forma de cuantificar el desempeño de un alumno reflejando su grado de conocimiento en un determinado contenido. Para

ello utilizamos registros de los momentos iniciales y finales a la hora de desarrollar una evaluación. Los registros de tiempo dedicado se obtuvieron de cinco clases de pregrado diferentes, con asignaturas impartidas por el mismo profesor, con el mismo contenido teórico, en la misma universidad. Los resultados confirman estadísticamente que cada minuto adicional que la alumna permanece haciendo la prueba, implica en mayores posibilidades de obtener un buen desempeño, así como diferencias de desempeño entre los géneros femenino y masculino, aunque no estadísticamente diferentes, demostrando que las alumnas tienen mayores posibilidades de alcanzar la media. El modelo también confirma, según los odds ratios, que durante las evaluaciones el desempeño de los estudiantes disminuye, teniendo la mejor puntuación en la primera prueba. A través de las referencias consultadas, entendemos que la diferencia en las calificaciones de cada alumno está influenciada por varios factores, resultado de sus propias vivencias.

Palabras clave: Regresión logística; Tiempo en evaluación; Razón de probabilidades; Indicadores educativos.

1. Introduction

Discussions about possible factors that may influence a student's performance while being assessed have been increasing significantly in educational environments, where such concern runs through all levels of education. However, evaluating is a practice that is found in the performance of any educator, being indispensable, because it permeates the different competences and skills of the group of students. The area of Mathematical Education has been concerned with the evaluation and proposition of teaching methods in mathematics that can provide a better understanding of the concepts involved in this science.

One way to quantify a student's performance against an evaluation method is to assign a grade that reflects the student's level of knowledge to a particular problem. There are several ways to assess a student's performance by looking at their grade, which can be summarized by simple calculations of descriptive statistics or, more complex methods such as those used in the national high school exam, the item response theory (IRT) (C. Zanon, Hutz, Yoo, & Hambleton, 2016).

As mentioned, these methods take into account the student's grade for a given question and represent, in fact, a very coherent form of performance evaluation from a statistical point of view. But, perhaps, other variables can influence a student's performance, such as anxiety level, preparation time for an assessment, previous knowledge, health conditions, food, among others.

Among these variables, the time taken for a test can, in several situations, determine success or failure. For a sprinter, race time is obviously the determining factor for victory. In general, for races time can be considered the most important variable in order to quantify an athlete's performance. But in relation to a student who will be subjected to an evaluation exam, can the time spent on the exam influence his performance?

Given the above, the present study used the logistic regression model, since this model describes the behavior between a dependent variable and independent variables (Favéro, Belfiore, Da Silva, & Chan, 2009). In addition to the influence of the test time, it was investigated whether gender is also a variable capable of influencing the achievement of an average grade. We evaluate the model's fit quality and provide odds ratios to measure the effect of test time and other variables on obtaining an average grade.

In this sense, the grades and time spent in the development of evaluations of curriculum components related to statistics, of students from five different courses, both taught at the Federal University of Pampa - UNIPAMPA campus of Itaquí / RS, were evaluated. Of the evaluations, the time spent for their development and the grades obtained were recorded.

Finally, we consider that, as a teacher, to be concerned with these factors and reflect on them, is to be committed to pedagogical practice and professional training itself, permanently in practice and in reflection on practice (Freire, 1997).. In this sense, studies of the type can enable actions with respect to lesson planning, evaluations and conduct in the academic environment.

2. Literature review

This section will mention the evaluation systems and their importance for the pedagogical practice, as well as ways that this happens. Next, statistical methods and some aspects of the influence of time under evaluation and good performance will be discussed.

2.1 Evaluation Systems

The evaluation process is part of the daily life of all educators, which contributes to the diagnosis of student learning and performance, as well as to the reflection of teaching practice. Therefore, for teaching practice, evaluating is not performing a technical activity, as it does not occur in the conceptual void, but is referenced in a theoretical model of the world, science and education, which is done in pedagogical practice (Chuire, 2008).

This idea refers to the evaluation in a global way and continues to permeate the different competences and skills of the group of students, applied at different times, in view of the arguments, interests and positions presented by the students in the classroom, since it is in the exchange of knowledge learning takes place. For Kraemer (2005), the term “evaluation” comes from the Latin, and means value or merit to the object in research, joining the act of evaluating to measuring the knowledge acquired by the individual.

The author also highlights that assessment is a valuable tool in the school system, since it is able to describe the knowledge that students have appropriated, as well as the difficulties encountered in the teaching-learning process (Kraemer, 2005). Once the assessment is consolidated, it is possible to identify evidence, capable of pointing out needs for improvements in teaching and achieved objectives that were significant.

For the student being evaluated, in addition to the study, the correct management of time during the evaluation is a fundamental factor for a good performance in a test. So that the student is careful not to stay too long on a single question and extrapolate the time, or become discouraged at the beginning of the test and give up on testing his reasoning ability. Boff (2010) points out that managing time well is seeking efficiency in activities and consequently increasing productivity.

2.2 Assessment ways

The assessment is the most important tool that the teacher has at his / her disposal, to reach the academic objective: Justify the advance or the permanence of the student in a certain discipline. However, this should take place in different ways and, during classes, with pre-established dates, using different instruments, such as presentation of papers, tests and resolution of exercises.

Mathematical Education proposes to enable all students to be able to investigate and formulate mathematical conjectures, to develop, evaluate, apply and present arguments, to select and use various types of reasoning and proof methods (Viseu, Menezes, Fernandes, Gomes, & Martins, 2017). Thus, even if each teacher adopts a way of evaluating, it is essential that the meanings behind the essence of the process are considered, as a means to contemplate their students and to identify learning or difficulty expressed by the student. Antunes (2002) states that the teacher's task when evaluating requires competence, discernment, balance, besides, of course, technical knowledge.

Valuing the student's production in solving proposed exercises also makes it possible to evaluate and verify his performance in a given discipline, since it comprises an intentional collection of student work, with the aim of telling the story of efforts, progress and performance in a given area (Freire, 1997).

However, when the evaluation permeates the development of activities using materials for consultation, new ways of learning emerge, since learning, in fact, is carried out when the content is put into practice. The works developed with consultation “help the student to solve a case, to choose the appropriate sources to extract information from them” (Althaus &

Godoy, 2008; D. P. Zanon, Althaus, & Bagio, 2018).

2.3 Statistical methods

Considering the importance of Statistics for the understanding of concepts nowadays, promoting effective learning becomes increasingly important, since its applications are numerous. However, there are studies that indicate that students see in Statistics a high level of difficulty in understanding, since they make associations with mathematical learning and conclude, a priori, that they will also not be able to learn Statistics (Jacobini & Wodewotzki, 2001).

It is evident that the study of Statistics requires basic skills for its understanding. Such as the ability to organize data, build and present tables, work with different representations of the data, understanding of concepts, vocabulary and symbols and, in addition, an understanding of probability as a measure of uncertainty (Campos, Jacobini, Wodewotzki, & Ferreira, 2011).

In this sense, Wild e Pfannkuch (1999 apud De Oliveira Mendonça & Lopes, 2011) present some important actions for statistical reasoning, being these:

- Recognition of the need for data: it consists of realizing whether, based on data analysis, decision making is possible.
- Transnumbing: indicates the understanding that can arise when changing the form of representation (graph, table).
- The perception of variability: the adequate collection of data and the judgments elaborated from these require an understanding of the variability that exists and is transmitted in the data.
- The reasoning with models: Any representation of data, such as a graph, a table, a regression line or a summary, is a model representative of reality and allows observing the behavior of the variable under study, which makes it possible to seek its approximation with a statistical model.
- The integration of Statistics with the context: the data that will be observed from the perspective of statistical concepts belong to a context, and knowing it is fundamental for the understanding of the meanings expressed in the data.

As a result of these considerations, the aforementioned investigation took place, which aims to identify whether the time that students use to develop a dissertation test, in disciplines of curricular components related to Statistics, may or may not influence the results obtained by the student. Considering that these make up groups of students from different undergraduate courses, but who take the course with the same teacher and the same theoretical content.

3. Methodology

Given the above, in order to investigate the relationship between test time and meet the objectives of the present work, an experimental research was carried out, as recommended and reported by Köche (2011) and Yin (2001). The descriptions of the dataset, as well as the statistical methodology are given below.

3.1 Dataset description

The data were obtained from evaluations of students who attended curriculum components related to Statistics, both taught by the same teacher. The students belonged to different courses, such as: Agronomy, Food Science and Technology, Cartographic and Surveying Engineering, Interdisciplinary Bachelor of Science and Technology and Degree in Mathematics, from the Federal University of Pampa- UNIPAMPA, campus of Itaqui / RS. 784 marks were registered, distributed in five courses, three years and three types of exam.

Of the evaluations, the initial and delivery moment of each student was recorded. The test time is given by the difference between the delivery and initial time, converted into minutes. In order to establish a test performance threshold, the

score of six (6) points was adopted, since if the student obtains this grade in all assessment systems, he is approved. Thus, the data were analyzed statistically by means of logistic regression, and for the applicability of the proposed model, the categorization of notes is as follows:

- 1: Student grades greater than or equal to six (6) points.
- 0: Student grades less than six (6) points.

It should be noted that the choice of the 6-point threshold was due to the student's interest in obtaining a grade higher than the average practiced at UNIPAMPA for approval in the curricular component and, in the specific case of the Statistics components, this same threshold was applied to all assessments of this component. This threshold could, according to a specific interest, be changed.

3.2 Statistical model

This section presents the context in which the response variable has only two categories, that is, binary or dichotomous nature. The logistic regression model is part of the Generalized Linear Models (GLM) class. The GLM is specified by three components: a random component, which identifies the probability distribution of the dependent variable, which specifies a linear function between the independent variables and a link function, which describes the mathematical relationship between the component systematic and the expected value of the random component (Hosmer, Lemeshow, & Sturdivant, 2013).

Consider the logistical transformation, also called logit (de Menezes, Liska, Cirillo, & Vivanco, 2017), given by

$$\text{logit} = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (1)$$

Analyzing $\pi(x)$ requires a sample x_1, x_2, \dots, x_n , which each variable x varies on the real line and $\pi(x)$ is a monotonic function with values between zero and one. Usually, $\pi(\cdot)$ is referred as the inverse of the function *logit*, the logistic function, which is given by

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \quad (2)$$

on what $\pi(x)$ ranges in $[0; 1]$. In case we have a single independent variable in the model, say x_1 , if $\beta_1 > 0$, π is growing and if $\beta_1 < 0$, π is decreasing. When x tends to infinity, $\pi(x)$ tends to zero when $\beta_1 < 0$ and tends to a while $\beta_1 > 0$. For instance $\beta_1 = 0$, the response variable is independent of the variable x , soon $\pi(x)$ is constant. The case $\beta_0 = 0$ e $\beta_1 = 0$ stands for $\pi(x) = 0.5$.

Once the model is defined, the parameters must be estimated. Be $\beta = (\beta_0, \beta_1)$ the parameter vector related to the conditional probability $P(Y_i = 1|x_i) = \pi(x_i)$, com $\pi(x_i)$ given by

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_{i1})}{1 + \exp(\beta_0 + \beta_1 x_{i1})} \quad (3)$$

The usual method for estimating $\hat{\beta}$ is through Maximum Likelihood method. To do this, be y_i a random variable that follows a Bernoulli distribution, so the contribution of (x_i, y_i) for the likelihood function is given by

$$\pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}$$

for values $y_i = 0$ or $y_i = 1$, for $i \in \{1, \dots, n\}$. As the observations are assumed to be independent, the likelihood function (L), is given by

$$L(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (4)$$

Applying neperian logarithm to the equation (4), it leads to the expression (5), $l(\beta)$,

$$l(\boldsymbol{\beta}) = \ln[L(\boldsymbol{\beta})] = \sum_{i=1}^n [y_i \ln(\pi(x_i)) + (1 - y_i) \ln(1 - \pi(x_i))] \quad (5)$$

By the maximum likelihood method, the estimator $\hat{\boldsymbol{\beta}}$ is the value that maximizes the equation (5). So, it derives $l(\boldsymbol{\beta})$ with respect to β_0 and β_1 and the resulting expressions are equal to zero, resulting in a system of non-linear equations in the parameters and, thus, the use of a numerical iterative procedure is required to find $\hat{\boldsymbol{\beta}}$, such as the Newton-Raphson method (McCulloch, Searle, & Neuhaus, 2009).

3.3 Model selection and goodness of fit criteria

The stepwise method of selection of variables consists of eliminating from the model variables that do not significantly contribute to the expected value of the response variable, in the case of the logistic model, the probability of the occurrence of an event of interest. In this work, the Akaike Information Criterion (AIC) was used to select variables using the stepwise variable selection method, which consists of the following steps:

- (1) Fit the complete model with all independent variables;
- (2) Remove one independent variable at a time, fit the model and calculate the AIC, proposed in Akaike (1974), given

by

$$AIC = -2 l(\hat{\boldsymbol{\beta}}) + 2p \quad (6)$$

where p is the number of model parameters;

- (3) Remove the independent variable that produced the highest AIC from the complete model;
- (4) Fit the model without the independent variable removed in step (3);
- (5) Go back to step (2) and redo the process until there is no independent variable to be removed. After this procedure and with the model obtained, we redo the inverse process of step (1), that is, we will include, if possible, independent variables in the model using the AIC.

Given a set of models fitted to the data, the preferred model is the one with the lowest AIC value (Emiliano, Vivanco, & de Menezes, 2014; Liska, Cirillo, de Menezes, & Bueno Filho, 2019). The goodness of the fit can also be verified by a normal probability plot for standard residuals, providing a graphic indication that the Bernoulli distribution assumption is violated. This graph, also called a simulated envelope plot, contains confidence bands built by resampling and it is considered that an adequate fit has occurred if all model residuals (or at least most of them) are contained in these bands (Cordeiro, 2004; Gilberto, 1995; Moral, Hinde, & Demétrio, 2017).

3.4 Odds ratio

Initially consider the model defined in 3, with $\pi(x)$ the probability of a student obtaining a grade higher than six points. The odds ratio will quantify the effect of a particular independent variable, which may be the student's test time, gender and / or class, on the student's probability of obtaining a grade higher than six points. Thus, the odds (OD) of the student to obtain a score higher than 6 points, for an individual with ($x = 1$) factor (for example, sex, where sex = 1 refers to male student), denoted by OD_1 , is given by

$$OD_1 = \frac{\pi(1)}{1 - \pi(1)} = e^{\beta_0 + \beta_1} \quad (7)$$

while the chance of a grade higher than 6 for an individual with absence ($x = 0$) of the factor, denoted by OD_0 , is

$$OD_0 = \frac{\pi(0)}{1 - \pi(0)} = e^{\beta_0}.$$

Therefore, the odds ratio (OR) of a score higher than 6 among individuals with presence and absence of the factor is

given by

$$OR = \frac{\pi(1)[1-\pi(0)]}{\pi(0)[1-\pi(1)]} = e^{\beta_1}, \quad (8)$$

depending only on parameter β_1 . This is one of the great advantages of logistic regression: the possibility of direct interpretation of the coefficients as measures of association (Giannini, Latorre, & Ferreira, 2012; Sio Jyh Lih & Ismail, 2019). This concept can be extended to the multiple independent variables, except that, in this case, the odds ratio is made between the variable of interest, as mentioned above, and keeping all other explanatory variables fixed, taking equation 8 (Leite, Guimarães, Lima, Barrozo, & Tavares, 2011).

The observed value of the independent variable in the logistic model may represent the value of any quantitative variable, such as, for example, a student's test time. In this case, it makes sense to calculate the odds ratio for an individual to have a score higher than 6 points for each increment that may be of one unit or more, in their test time. The odds ratio of obtaining a grade higher than 6 for an increment c , such that $c = x^* - x$, is given by

$$OR_{(x^*-x)} = \frac{\pi(x^*)[1-\pi(x)]}{\pi(x)[1-\pi(x^*)]} = e^{\beta_1(x^*-x)} \quad (9)$$

Finalizing the methodology, the R software (R Core Team, 2017) and its integrated environment RStudio (RStudio Team, 2015), were used as a tool for the development of the model, which are free software that allow programming for graphics and statistical analysis.

4. Results and Discussion

For better organization of results, these will be presented in three scenarios: influence of time on student performance; addition of the gender covariate regarding the student's performance; and finally the type of proof.

4.1 Time influence on student's performance

The evaluation process has its contributions based on the diagnosis of student performance, where several factors are capable of causing positive and negative impacts, such as those highlighted by Zewude and Ashine (2016) who point out that the time of study, influence of colleagues, quantity of money received from the family, a good life after school years and parents' education are variables capable of influencing academic performance.

For the author, students who study more than 48 hours are 5.85 times more likely to have better academic performance than those who study less hours (Zewude & Ashine, 2016). Another interesting study also points to family income as a significant variable for frequency and good academic performance, and when the monthly income increases by one the probability of the student continuing to attend classes is 1,0001 times greater than the probability of those students whose income monthly is less than one unit (Abdellatif Fadul & Osman Abd Elghafar Mohammed, 2018).

Pondering in this sense when interpreting the coefficient in the logistic regression model, we found statistically that time is capable of influencing the student's performance when performing a test (table 1). Given this statement, there is an apparent interpretation that for each additional minute that the student remains taking the test, the odds ratio of success reaching the average increases by 1.55% (OR_I). In addition, every 10 minutes that the student remains doing the test, it results in an 11.6% increase (OR_{10}) in his chance of reaching the average.

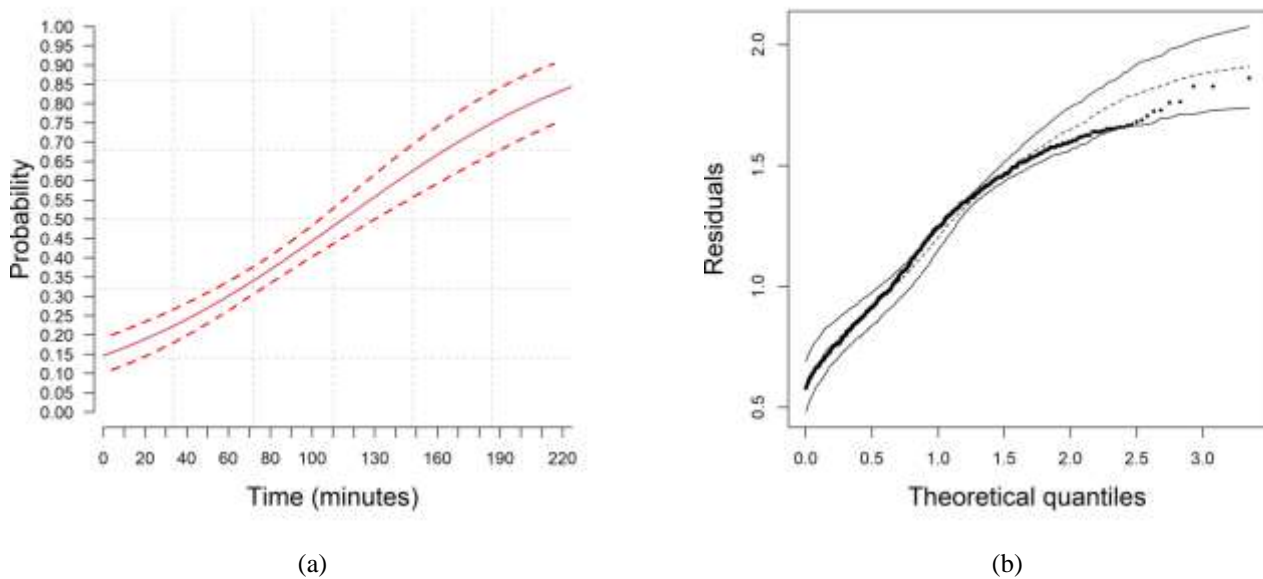
Table 1: Parameters estimates of the logistic regression model, odds ratio (*OR*) and Akaike information criteria (AIC) for student performance in relation to time, according to equation 3.

Parameter	Estimate	Standard error	<i>p</i> -value	<i>OR</i> ₁	<i>OR</i> ₁₀	AIC
β_0	-1.7634	0.1845	<0.0001	1.0155	1.116	964.73
β_1	0.0154	0.0020	<0.0001			

*OR*₁: Odds ratio for each additional minute that the student remains taking the test.
*OR*₁₀: Odds ratio for additional 10 minutes that the student remains taking the test.
 Source: Authors

All the results are supported by satisfactory fit of the logistic regression models, as can be seen in Figures 1 (b), 2 (b) and 3 (b), which show that, with 95% confidence, the Bernoulli distribution assumption for student's performance by obtaining a grade equal to or greater than 6 points is suitable and the proposed models can be used to model the performance of students in relation to the time's proof and other covariates, such as gender and type of proof.

Figure 1: Graph of the fitted model to the time's proof (solid line) with 95% confident intervals for fitted probabilities (dashed lines) (a) and simulated envelope plot of the fitted model (b).



Source: Authors.

Thus, evaluating is not simply performing a technical activity, but a global and continuous process that cuts across the different skills of pedagogical practice, where starting from the provision of clear learning objectives is a means of positively contemplating students, obtaining greater chances of satisfying students. In general with the service of disciplines increasing 6 times in relation to the subjects not taught with such objectivity (Alija, Snopce, & Aliu, 2016).

A similar study interpreted from the regression model coefficient, confirmed an association between voice disorder and loss of ability to work in teachers in São Paulo, pointing that teachers are 2.2 times more likely to have voice disorder when compared to other professionals with less vocal use (Giannini, Latorre, & Ferreira, 2013).

As for the present study, even if there is a direct relationship between the time spent to develop an assessment and the student's performance, it is important to pay attention to the ideal time that implies greater chances of the student reaching a

compatible average for his approval. Thus, and according to the estimates of the parameters of the logistic regression model in Table 1, it is statistically suggested that, for the student to have a 50% chance of reaching the average, it is necessary that he remains 114.8 minutes taking the test. Such a result can be seen by doing:

$$x_{50\%} = \frac{\ln\left(\frac{0.5}{1-0.5}\right) - \hat{\beta}_0}{\hat{\beta}_1} = \frac{\ln\left(\frac{0.5}{1-0.5}\right) + 1.763370}{0.015362} = 114.8 \text{ min.} \quad (10)$$

According to the references, when it comes to time spent by the student in the postgraduate program, whenever it increases, the chances of completing the master's degree also increase by a factor of 2.88% (Pyke & Sheridan, 1993).

For Alija (2015), who also developed a study based on the use of logistic regression and sought to determine some of the satisfaction factors, which positively influence students' perception of some teaching practices, found that more than 90% of students are satisfied or completely satisfied with the teaching offered by their teachers, as the data indicates that 67.5% of the students are completely satisfied, 29.3% are satisfied and 3.2% disagree, with 96.8% of students satisfied versus 3.2% dissatisfied.

4.2 Time and gender influence on student's performance

When analyzing other variables such as the influence of time and gender on the performance of students, it was possible to notice a significant difference between women and men, which indicate that women are more likely to reach the average, because for every minute that female students remain taking the test, their chances increase by 1.96%, while for male students, each additional minute making the assessment is 1.24% (Table 2). For Abedalaziz (2010) these differences occur because female people have advantages in relation to males in numerical capacity, hence the importance of the evaluation taking place through several instruments allowing contemplate the students as much as those presented in section 2.2.

According to the perspective presented by Mazulo (2015 apud Hair, 1998), since it is a multiple linear regression, where the objective is to use the independent variables that have known values, to predict the values of the independent variables, it is possible to affirm with 5% of significance, that the variables β_0 and β_1 are non-zero coefficients having real association with the response variable.

The model proposed by means of logistic regression that relates the time of test and gender is given by:

$$\pi(\text{Time}, \text{Sex}) = \frac{\exp(\beta_0 + \beta_1 \text{Time} + \beta_2 \text{Gender} + \beta_3 \text{Time} \times \text{Gender})}{1 + \exp(\beta_0 + \beta_1 \text{Time} + \beta_2 \text{Gender} + \beta_3 \text{Time} \times \text{Gender})} \quad (11)$$

As the variable Gender is categorical and according to Table 2, the models for the Feminine and Male genders are given by

$$\pi_{\text{Feminine}}(\text{Time}) = \frac{\exp(\beta_0 + \beta_1 \text{Time})}{1 + \exp(\beta_0 + \beta_1 \text{Time})} = \frac{\exp(-2.0864 + 0.0194 \times \text{Time})}{1 + \exp(-2.0864 + 0.0194 \times \text{Time})}$$

$$\pi_{\text{Male}}(\text{Time}) = \frac{\exp(\beta_0 + \beta_2 + (\beta_1 + \beta_3) \times \text{Time})}{1 + \exp(\beta_0 + \beta_2 + (\beta_1 + \beta_3) \times \text{Time})} = \frac{\exp(-1.49675 + 0.01234 \times \text{Time})}{1 + \exp(-1.49675 + 0.01234 \times \text{Time})}$$

Table 2: Parameters estimates of the logistic regression model, odds ratio (*OR*) and Akaike information criteria (AIC) for students' performance in relation to time and gender, according to equation 11.

Parameter	Estimate	Standard error	<i>p</i> -value	<i>OR</i> _{Feminine}	<i>OR</i> _{Male}	AIC
β_0	-2.0864	0.2747	<0.0001			
β_1	0.0194	0.0031	<0.0001	1.0196	1.0124	965.71
β_2	0.5896	0.3749	0.1158			
β_3	-0.0070	0.0041	0.0835			

*OR*_{Feminine}: Odds ratio for each additional minute that the feminine student remains taking the test.

*OR*_{Male}: Odds ratio for each additional minute that the male student remains taking the test.

Source: Authors.

Such statements agree with the statements made by Magalhães and Andrade (2002) which indicate that feminine students obtain better performance than male students. Guney (2009) who also investigates the gender variable, points out that feminine students outperform male students in academic performance, and Dal 'Igna (2007) adds that even though the activities are the same and the knowledge required is also the same for male and female students, the performance of female students is generally better.

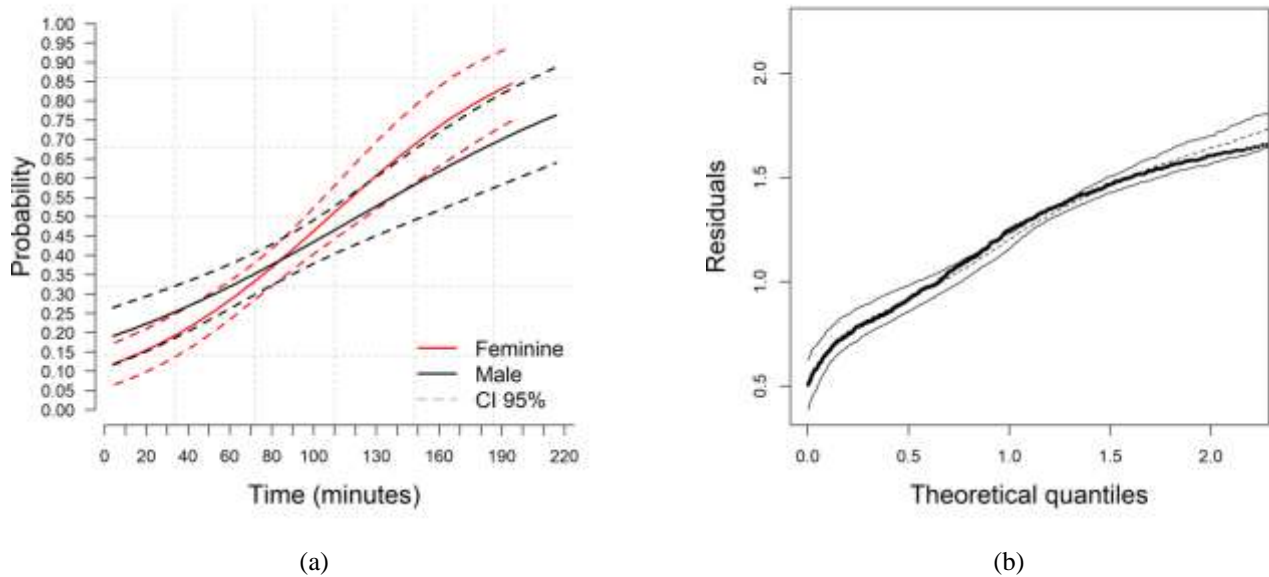
However, the investigation by Eratli Sirin and Sahin (2020) indicates the opposite, pointing out that the chances of male students being successful is 1.427 times greater than the chance of success for female students, which is still justified due to the ease having a good relationship with teachers that make the educational process more active.

In Daguplo (2017) the interest in attending higher education among high school students hears a statistically significant difference between sex and age. Older students are less likely to pursue higher education in studies conducted in the city of Irvine-California-USA. There are also references that point to the decision of young people to go on to higher education without significant influences due to the gender variable (De Dios Jiménez & Salas-Velasco, 2000). However, male students are less likely to persist in higher education (Ferrão & Almeida, 2018).

Analogously to equation 10, the model also explains that, for the male student to have a 50% chance of reaching the average in an evaluation, then he must remain taking the test, at least 121.2 minutes, whereas, for female students, based on the same probability of chances, the time decreases, to 107.7 minutes (Figure 2 (a)).

In this sense, Alija (2015) agrees when analyzing the gender variable, as he affirms that the achievements of female students are approximately 1.022 times greater than the achievement of male students, this perspective was determined when analyzing the probability of male and female students. pass the external test of the Secondary School of Economics. Niu (2020) also considered the gender variable relevant when it comes to academic development, according to the author women are 33% more likely than men to receive a university degree.

Figure 2: Graph of the fitted model (solid lines) to the time and gender test with 95% confident intervals for fitted probabilities (dashed lines) (a) and simulated envelope plot of the fitted model (b).



Source: Authors.

4.3 Variables selection

In addition to the variables previously analyzed and considering the inclusion of the course to which the student belongs and the type of test, the complete model is given by:

$$\begin{aligned} \text{logit} &= \ln\left(\frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})}\right) \\ &= \beta_0 + \beta_1 \text{Time} + \beta_2 \text{Gender} + \beta_{31} \text{course}_1 + \beta_{32} \text{course}_2 + \beta_{33} \text{course}_3 + \beta_{34} \text{course}_4 \\ &\quad + \beta_{35} \text{course}_5 + \beta_{36} \text{course}_6 + \beta_{41} \text{proof}_1 + \beta_{42} \text{proof}_2 \end{aligned} \quad (12)$$

Table 3: Parameters estimates of the complete logistic regression model and Akaike information criteria (AIC), according to equation 12.

Parameter	Estimate	Standard error	p-value	AIC
β_0	-0.80284	0.33969	0.01811	
β_1	0.01598	0.00211	<0.0001	
β_2	0.01429	0.17258	0.93400	
β_{31}	-0.74871	0.32513	0.02129	
β_{32}	-0.68505	0.34377	0.04629	
β_{33}	-0.41212	0.34669	0.23455	920.63
β_{34}	-0.57590	0.35527	0.10501	
β_{35}	-0.53871	0.81493	0.50858	
β_{36}	0.74814	0.87400	0.39199	
β_{41}	-1.14521	0.21162	<0.0001	
β_{42}	-0.93795	0.25121	0.00019	

Source: Authors.

After the procedure for selecting variables using the stepwise method, it was possible to obtain the set of variables in Table 3 that produced the most parsimonious model. Thus, once the fitted logistic model proved to be adequate (Figure 3(b)), expressing a more complete result is fundamental, in general the fitted logistic model showed again that it is possible for time (β_1) to influence students' performance.

In order to express a more practical result, Table 4 compares the estimated probabilities of performance in test 2 (β_{41}) and test 3 (β_{42}), which configures that students show different performance in the tests. The student's odds of getting a grade higher than 6 in the second test in relation to the first test decreases by 71%. Regarding the third test, the odds ratio is 62%. It is observed that, according to the odds ratios, that the student's performance decreases throughout the tests, even though his highest grade is higher in the first test in relation to test 2 and 3 and is lower in the second test.

Table 4: Parameters estimates of the complete logistic regression model (equation 12) with variables selected by the stepwise method, odds ratio (*OR*) and Akaike information criteria (*AIC*).

Parameter	Estimate	Standard error	p-value	<i>OR</i>		<i>AIC</i>
β_0	-1.3071	0.1973	<0.01	-		
β_1	0.0159	0.0020	<0.01	$OR_I=1.016$	$OR_{10}=1.173$	916.35
β_{41}	-1.2501	0.1933	<0.01	1- <i>OR</i> =0.7135		
β_{42}	-0.9652	0.2248	<0.01	1- <i>OR</i> =0.6191		

OR_I: Odds ratio for each additional minute that the student remains taking the test.

OR₁₀: Odds ratio for additional 10 minutes that the student remains taking the test.

1-OR: The student's odds ratio of getting a grade higher than 6 in the *k* test (*k*=2 or 3) in relation to the first (*k*=1) test.

Source: Authors.

Thus, this study makes it possible to state that in test 2 is where students have the lowest performance, with test 1 being the one with the best performance of students. Considering the results obtained and the concepts covered in each evaluation, we infer this performance with the fact that test 2 has more theoretical contents (calculation of probabilities and their distributions), and consequently less concrete to the student, which can add less assimilation of the content. On the other hand, in tests 1 and 3, the content presupposes to illustrate reality through theory (frequency distribution tables, graphs, sampling, confidence intervals, simple linear regression, among others), which can generate incentive for creativity and investigative practice, as it allows the student to interact with the phenomena by applying the studied concepts (Andrade & Massabni, 2011).

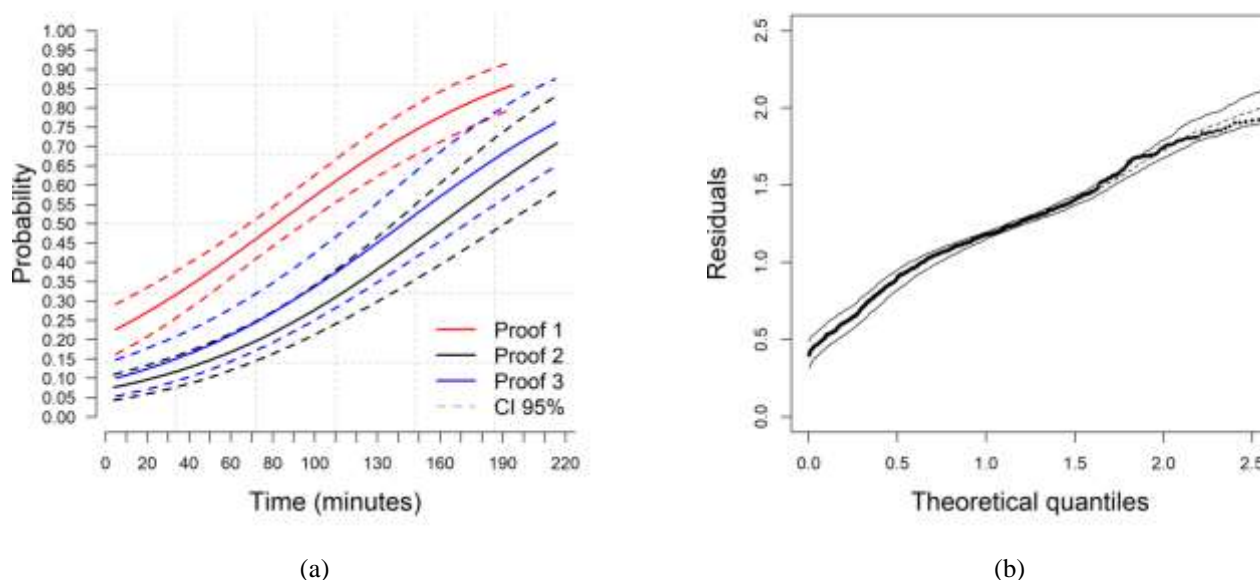
However, this is the moment that allows verifying if the actions for statistical reasoning were understood by the student, such as those highlighted by Wild and Pfannkuch (1999) presented in section 2.4. Formiga (2004) points out that academic success is the student's responsibility, as it depends on him knowing that he has to study, since it is impossible to conceive that in order to be a good student, the person will not devote time to studies, because genius or ability to understand and assimilate scientific content is not learned in a vacuum.

In such a way, we believe that the evaluation for the student comes to guide him, which results in a better performance in test 3 in relation to test 2, since it provides the student with a review of his performance, which can be seen in Figure 3(a).

Studies such as that by Pyke and Sheridan (1993), also present the use of logistic regression to formulate models that describe the retention patterns of master students, it is noticed that when selecting twenty-one of the independent variables with three possible terms of interaction, only four variables were found to be significant and related to student retention at the master's level. Among them are the time in the program, average amount of funding, academic average of graduation and type

of program.

Figure 3: Graph of the fitted model (solid lines) of the duration of the test and type of proof with 95% confident intervals for fitted probabilities (dashed lines) (a) and simulated envelope of the fitted model (b).



Source: Authors.

According to conclusions by Batistela, Rodrigues and Bononi (2009), the chances of a student dropping out of the undergraduate course can be explained by the variables “Residence”, “Family Income” and “Father's Occupation”. Based on the results presented, the authors justify that the chance of a student who lives in a residence outside the city of the course to give up is 2.29 times the chance of a student who lives in the city where he / she studies. Referring to the “Family Income” variable, they highlight that the chance of a student with an income above 5 minimum wages to give up is 2.02 times that of a student with an income below 5 wages. With regard to “Occupation of the Father”, they point out that the chance of a student that the occupation of the father is a businesswoman or politician or liberal professional or civil servant has the chance of giving up 5.29 the chance of a student whose father is military or agribusiness or salaried or standalone.

However, the period of time spent in the program was the most important contributor to the retention model at the master's level, which leads them to affirm that the highest probability of completing the master's degree is represented by students who stay longer in the program and consequently the least probability is with the students who stay the shortest time in the program, it is still pointed out that when assessing the influence of the variable “time”, the findings did not surprise them considering that the average time to complete the master's degree is equivalent to 2.74 years, for those who voluntarily withdraw 1.39 and those who are taken out by the university 1.47 years, which proves a significant difference in the amount of time (Pyke & Sheridan, 1993).

For Peng; Lee; Ingersoll (2002), when testing the hypothesis of whether or not there is a relationship between the likelihood of a child being recommended for corrective reading instructions and his score and gender, it was found that for boys the chances of them being indicated for reading programs corrective rate is 1.9111 times higher than for girls, where the statement came from plotting the variables of each gender group against the variables of readings and scores.

Therefore, the act of evaluating permeates all levels of education, and leads us to rethink and invest in strategies and possibilities to reduce the culture of disapproval or retention not only as a pedagogical method, but also for reflection and

reframing of learning (Lima, Casela, Ribeiro, & Rezende, 2019).

Based on the results obtained, we emphasize the importance of evaluation for the teaching-learning process, as we understand that in the evaluation the student is able to show that he or she dominates the content worked, so that he can use it to understand phenomena and act in society (Vizzotto & Mackedanz, 2018).

It should be noted that, for the results presented here, we start from the principle that the student occupies his time for the effective development of the evaluation, considering that the results cannot be considered if the student remains his time in the evaluation without doing any activity.

5. Conclusion

The logistic regression models fitted satisfactorily to the data that relate the study time to the grade obtained by students in a statistical evaluation. We found that as the student's test time increases, the chances of good performance increase. For each additional minute that the student remains taking the test, the chance of success reaching the average increases by 1.55% and that in approximately 115 minutes of effective work, represents a 50% chance of obtaining a grade higher than 6 points.

We noticed that the references consulted also indicate an apparent influence of the variable time in several aspects related to student life, presenting itself as a factor capable of determining the success or failure of the students.

The results did not show any significant difference between genders in relation to student performance. However, for every minute those female student remains taking the test their chances are slightly higher compared to male students. The model reiterates that of the three assessments proposed during the course, the one that students have the best performance according to the odds ratio is test 1, and throughout the test their performance decreases, with the lowest score in test 2.

However, this study is conclusively limited only when the student uses his time for the effective development of the assessment, in the search for solutions to the proposed questions. Despite this, the study is important to assess the test time between students and assist in proof planning. The study indicates that, disregarding any covariates, a test time of approximately 2 hours in exact subjects, in this case those related to applied statistics, provides a probability of almost 50% that the student has a grade higher than 6 points. We know that many variables possibly influence this process and we present analyzes with two of them (proof type and gender) and we emphasize that the study can be useful as a pilot study in related analyses. Possibly, relating the test time to the student's affinity with the test subject, the amount of hours the student studied to take the proof, the amount of sleep hours the night before the proof, multiple choice and discursive tests, can provide interesting information about the subject and constitute future studies.

All of this makes us believe that in education, there are countless factors that can impact the student's performance, above all it is important that the teacher consider, when assessing, what meanings were behind the essence of the process, as a means for contemplate the students and, identify learning or difficulty expressed by the student.

Acknowledgments

We are grateful to UNIPAMPA and the Post-Graduation (Specialization) in Exact Sciences and Technology for the opportunity to carry out this research.

References

- Abdellatif Fadul, A., & Osman Abd Elghafar Mohammed, K. (2018). Using Logistic Regression to Model Factors Associated With Basic School Drop-Out in Kassala State, Sudan. *Global Journal of Pure and Applied Mathematics*, 14(2), 301–313. <http://www.ripublication.com>
- Abedalaziz, N. (2010). Detecting gender related DIF using logistic Regression and Mantel-Haenszel approaches. *Procedia - Social and Behavioral Sciences*,

7(2), 406–413. 10.1016/j.sbspro.2010.10.055

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Alija, S. (2015). Application of Ordinal Logistic Regression in the Study of Students' Achievement in External Testing. *Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences - Law*, 8(57), 17–22. <http://www.iiste.org/Journals/index.php/MTM/article/view/7679>

Alija, S., Snopce, H., & Aliu, A. (2016). Logistic Regression for Determining Factors Influencing Students' Perception of Course. *The Eurasia Proceedings of Educational and Social Sciences*, 5, 99–106. <https://dergipark.org.tr/en/pub/epess/issue/30752/332916>

Althaus, M. T. M., & Godoy, M. T. (2008). Lessons to pilot classes: interfaces between didactics and supervised training in biology. *Olhar de Professor*, 11(2), 295–311. 10.5212/OlharProfr.v.11i2.295311

Andrade, M. L. F. de, & Massabni, V. G. (2011). Practical activities development: a challenge to science teachers. *Ciência & Educação*, 17(4), 835–854. 10.1590/S1516-73132011000400005

Antunes, C. (2002). *Avaliação da Aprendizagem Escolar* (4a ed.) Vozes, Ed. Vozes.

Batistela, G. C., Rodrigues, S. A., & Bononi, J. T. C. M. (2009). Estudo sobre a evasão escolar usando regressão logística: análise dos alunos do curso de administração da Fundação Educacional de Ituverava. *Tekhnê e Logos*, 1(1), 103–120.

Boff, R. J. (2010). Administração do Tempo. *InterLink*, 1(1), 59–74.

Campos, C. R., Jacobini, O. R., Wodewotzki, M. L. L., & Ferreira, D. H. L. (2011). Educação Estatística no Contexto da Educação Crítica. *BOLEMA*, 24(39), 473–494. <http://www.periodicos.rc.biblioteca.unesp.br/index.php/bolema/article/view/5104>

Chueire, M. S. F. (2008). Concepções sobre a Avaliação Escolar. *Estudos Em Avaliação Educacional*, 19(39), 49. 10.18222/ea193920082469

Cordeiro, G. M. (2004). On Pearson's residuals in generalized linear models. *Statistics & Probability Letters*, 66(3), 213–219. 10.1016/j.spl.2003.09.004

Daguplo, M. S. (2017). Binary Logistic Predictive Model in Determining Students' Intention to Take Higher Education. *Asia Pacific Journal of Multidisciplinary Research*, 5(4), 138–143. <http://www.apjmr.com/wp-content/uploads/2017/12/APJMR-2017.5.4.17.pdf3d351f6db0140be97b35f1e5945a3c8&site=eds-live&scope=site>

Dal'Igna, M. C. (2007). Desempenho escolar de meninos e meninas: há diferença? *Educação Em Revista*, (46), 241–267. 10.1590/S0102-46982007000200010

De Dios Jiménez, J., & Salas-Velasco, M. (2000). Modeling educational choices. A binomial logit model applied to the demand for Higher Education. *Higher Education*, 40(3), 293–311. 10.1023/A:1004098300436

de Menezes, F. S., Liska, G. R., Cirillo, M. A., & Vivanco, M. J. F. (2017). Data classification with binary response through the Boosting algorithm and logistic regression. *Expert Systems with Applications*, 69, 62–73. 10.1016/j.eswa.2016.08.014

De Oliveira Mendonça, L., & Lopes, C. E. (2011). Mathematical Modeling: A learning environment for the implementation of Statistics Education in High School. *Bolema - Mathematics Education Bulletin*, 24(40), 701–724. <http://www.periodicos.rc.biblioteca.unesp.br/index.php/bolema/article/view/5290>

Emiliano, P. C., Vivanco, M. J. F., & de Menezes, F. S. (2014). Information criteria: How do they behave in different models? *Computational Statistics & Data Analysis*, 69, 141–153. 10.1016/j.csda.2013.07.032

Erathl Şirin, Y., & Şahin, M. (2020). Investigation of Factors Affecting the Achievement of University Students with Logistic Regression Analysis: School of Physical Education and Sport Example. *SAGE Open*, 10(1). 10.1177/2158244020902082

Favéro, L. P., Belfiore, P., Da Silva, F. L., & Chan, B. L. (2009). *Análise de dados: modelagem multivariada para a tomada de decisões* (Elsevier, Ed.). Elsevier.

Ferrão, M. E., & Almeida, L. S. (2018). Multilevel modeling of persistence in higher education. *Ensaio*, 26(100), 664–683. 10.1590/S0104-40362018002601610

Formiga, N. S. (2004). Atribuição de causalidade e rendimento acadêmico: Predição e diferença em alunos da universidade pública e privada. *Revista de Psicologia Da UNC*, 2(1), 3–14.

Freire, P. (1997). Pedagogia da Autonomia: Saberes necessários à prática educativa. In P. & Terra (Ed.), *Paz e terra* (58th ed.). Paz & Terra.

Giannini, S. P. P., Latorre, M. do R. D. de O., & Ferreira, L. P. (2012). Distúrbio de voz e estresse no trabalho docente: um estudo caso-controle. *Cadernos de Saúde Pública*, 28(11), 2115–2124. 10.1590/S0102-311X2012001100011

Giannini, S. P. P., Latorre, M. do R. D. de O., & Ferreira, L. P. (2013). Distúrbio de voz relacionado ao trabalho docente: um estudo caso-controle. *Brazilian Journal of Otorhinolaryngology*, 3(1), 129–134.

Gilberto, P. A. (1995). Influence and residuals in restricted generalized linear models. *Journal of Statistical Computation and Simulation*, 51(2–4), 315–331. 10.1080/00949659508811640

Guney, Y. (2009). Exogenous and endogenous factors influencing students' performance in undergraduate accounting modules. *Accounting Education*, 18(1), 51–73. 10.1080/09639280701740142

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression. In *Wiley series in probability and statistics CN - QA278.2 .H67 2013*

(3rd ed.). 10.1002/9781118548387

Jacobini, O. R., & Wodewotzki, M. L. L. (2001). A Modelagem Matemática Aplicada no Ensino de Estatística em Cursos de Graduação. *Bolema - Boletim de Educação Matemática*, 14(15), 47–68.

Köche, J. C. (2011). *Fundamentos de metodologia científica : teoria da ciência e iniciação à pesquisa*. Vozes.

Kraemer, M. E. P. (2005). A avaliação da aprendizagem como processo construtivo de um novo fazer. *Avaliação: Revista Da Avaliação Da Educação Superior*, 10(2), 137–147.

Leite, R. C. M., Guimarães, E. C., Lima, E. A. P. de, Barrozo, M. A. de S., & Tavares, M. (2011). Utilização de regressão logística simples na verificação da qualidade do ar atmosférico de Uberlândia. *Engenharia Sanitaria e Ambiental*, 16(2), 175–180. 10.1590/S1413-41522011000200011

Lima, N. da C. M., Casela, A. L. M., Ribeiro, L. V. F., & Rezende, W. (2019). Associação do índice de atitudes e práticas pedagógicas ao desempenho dos estudantes na avaliação em larga escala do estado do Espírito Santo. *Educação Em Revista*, 35. 10.1590/0102-4698198087

Liska, G. R., Cirillo, M. Â., de Menezes, F. S., & Bueno Filho, J. S. de S. (2019). Machine learning based on extended generalized linear model applied in mixture experiments. *Communications in Statistics - Simulation and Computation*, 1–15. 10.1080/03610918.2019.1697821

Magalhães, F. A. C., & Andrade, J. Xi. (2006). Exame vestibular, características demográficas e desempenho na universidade: em busca de fatores preditivos. *Congresso USP de Controladoria e Contabilidade*, 1–11.

Mazulo, E. S. (2015). Análise da proficiência em matemática por meio de regressão linear múltipla. *Revista Intersaberes*, 10(21), 613–625. 10.22169/revint.v10i21.664

McCulloch, C. E., Searle, S. R., & Neuhaus, J. M. (2009). *Generalized, Linear, and Mixed Models*. John Wiley Sonc Inc.

Moral, R. A., Hinde, J., & Demétrio, C. G. B. (2017). Half-Normal Plots and Overdispersed Models in R : The hnp Package . *Journal of Statistical Software*, 81(10). 10.18637/jss.v081.i10

Niu, L. (2020). A review of the application of logistic regression in educational research: common issues, implications, and suggestions. *Educational Review*, 72(1), 41–67. 10.1080/00131911.2018.1483892

Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research*, 96(1), 3–14. 10.1080/00220670209598786

Pyke, S. W. ., & Sheridan, P. M. (1993). Logistic Regression Analysis of Graduate Student Retention. *The Canadian Journal of Higher Education*, 23(2), 44–64. <https://journals.sfu.ca/cjhe/index.php/cjhe/article/view/183161>

RStudio Team. (2015). *RStudio: Integrated Development for R*. <http://www.rstudio.com/>

Sio Jyh Lih, J., & Ismail, R. Bin. (2019). Binary Logistic Regression Analysis of Teacher Self-Efficacy Factors Influencing Literacy and Numeracy. *World Journal of Education*, 9(1), 209. 10.5430/wje.v9n1p209

R Core Team (2017). *R: A language and environment for statistical computing*. <http://www.r-project.org/>

Viseu, F., Menezes, L., Fernandes, J. A., Gomes, A., & Martins, P. M. (2017). Conceções de Professores do Ensino Básico sobre a Prova Matemática: influência da experiência profissional. *Bolema: Boletim de Educação Matemática*, 31(57), 430–453. 10.1590/1980-4415v31n57a21

Vizzotto, P. A., & Mackedanz, L. F. (2018). Validação De Instrumento De Avaliação Da Alfabetização Científica Para Egressos Do Ensino Médio No Contexto Da Física Do Trânsito. *Educação Em Revista*, 34(0). 10.1590/0102-4698202974

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–248. 10.1111/j.1751-5823.1999.tb00442.x

Yin, R. K. (2001). *Estudo de caso: planejamento e métodos* (2nd ed.). Bookman.

Zanon, C., Hutz, C. S., Yoo, H., & Hambleton, R. K. (2016). An application of item response theory to psychological test development. *Psicologia: Reflexao e Crítica*, 29(1). 10.1186/s41155-016-0040-x

Zanon, D. P., Althaus, M. T. M., & Bagio, V. A. (2018). *Didática na docência universitária em Saúde: metodologias ativas e avaliação*. <https://www.editoraapparis.com.br/produto/737-didtica-na-docncia-universitria-em-sade-metodologias-ativas-e-avaliao>

Zewude, B. T., & Ashine, K. M. (2016). Binary Logistic Regression Analysis in Assessment and Identifying Factors That Influence Students ' Academic Achievement : The Case of College of Natural and Computational. *Journal of Education and Practice*, 7(25), 1–6. <https://eric.ed.gov/?id=EJ1115855>