IMPLEMENTATION OF THE LEARNING EVALUATION MODEL
PRESENTED BY J.G. KALBFLEISCH

L. Castillo Cabay¹, N. Sanchez-Choez²
¹ Universidad Central del Ecuador (ECUADOR)
² Escuela Politécnica Nacional (ECUADOR)

Abstract

The learning at school level is evaluated periodically and if required, the student must repeat it. It is assumed that initially, no one meets the evaluation, but as you acquire the knowledge you must pass the evaluation, otherwise, the exam should be repeated. Thus, the subsequent results are not independent. For this purpose, the learning model presented by Kalbfleisch in his book of Probability and Statistical Inference was applied. This enabled that two groups were compared and exposed under different conditions in which the learning (A) and the need for improvement (B) parameters were analyzed, by maximum verisimilitude estimation. For the data processing, it was used the statistical software R. The obtained results determinate that there are effects of the learning processes: \( A < B \) and \( A < 1 \); In particular group B shows better effect due to the fact that A is smaller and the verisimilitude contours, which was 15% for A, show a significant difference between groups.

Keywords: Learning evaluation, maximum verisimilitude estimation, statistical inference.

1 INTRODUCTION

According to the Ministry of Education of Ecuador, an educational quality standard describes expected accomplishments, indicating educational goals. In particular, when referring to the students, it is looked for to evaluate that they fulfill a series of skills of the curricular area, which is reflected in their performances. Thus, we seek common references that students must achieve [1].

The academic reinforcement consists of a feedback that looks for planning activities of academic improvement. These activities include 1. Reinforcement classes led by the same teacher or another teacher who teaches the same subject. 2. Individual tutorials with the same or another teacher who teaches the same subject. 3. Individual tutoring with an educational or expert psychologist based on the educational needs of the students. 4. Schedule of studies that students must complete at home with the help of their family [2].

In this context, although this reinforcement may be based on the perception of what the student requires when it is within the quality standards; usually, in educational institutions, periodic evaluations are taken to the students of Basic General Education, denominated block evaluation (up to 6 blocks per a five-month period) and according to the results, if the student does not report a satisfactory qualification, this one is forced to attend the so-called recovery classes.

In this study, it is assumed that students start the school year in similar conditions, therefore, they raise questions, such as: Why do groups of students, a priori, achieve different results? Why a group unlike another percentage, receives more kinds of recovery? How to measure the learning factor? How to establish a measure of need for recovery in one or another group? And finally, How to measure the probability of receiving classes of recovery before an assessment, knowing the academic background of previous assessments?

In this context, it is applied the model of learning theory, which is presented by J.G. Kalbfleisch in his book Probability and Statistical Inference second edition of 1985. To establish, in a case study, the learning factors and the need for recovery in two groups that are assumed to be exposed to similar learning conditions.
2 METHODOLOGY

2.1 Sample
The sample is made up of 30 eighth-year students of Basic General Education, 12 of one group and 18 of another one, of the High School Marqués de Selva Alegre, located in the province of Pichincha, Canton Rumiñahui, school period 2015-2016. Data were collected for a period of 14 consecutive weeks.

2.2 Instruments
The main instrument of analysis is the learning model presented by Kalbfleish, from which the parameters should be estimated according to the data obtained [3].

2.2.1 The verisimilitude function
The verisimilitude function can be understood in a proportional way to the probability of the desired event. Thus, the maximum verisimilitude method seeks to estimate the parameters that maximize the probability of having achieved the sample obtained. For the particular case of the proposed model, we have the following results.

Guessing that you have a model given by the Beta distribution:

\[
f_x(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}
\]

For, \(0 \leq x \leq 1, \alpha > 0, \beta > 0\)

The verisimilitude function:

\[
L(\alpha, \beta; x) = \prod_{i=1}^{n} f(x_i; \alpha, \beta)
\]

The log-verisimilitude ratio

\[
l(\alpha, \beta; x) = \sum_{i=1}^{n} \log f(x_i; \alpha, \beta)
\]

The parameters \(\alpha\) and \(\beta\) without estimates when maximizing the verisimilitude or log-verisimilitude function

Complementarily, the relative verisimilitude function is defined as:

\[
R(\alpha, \beta) = \frac{L(\alpha, \beta)}{L(\hat{\alpha}, \hat{\beta})}
\]

Where \(\hat{\alpha}, \hat{\beta}\) are the maximum verisimilitude estimators of \(\alpha, \beta\).

In this way, the verisimilitude contour of 14.7% (equivalent to 95% confidence) is defined as:

\[
C(\alpha, \beta) = \{(\alpha, \beta): R(\alpha, \beta) \geq 0.147\}
\]

2.2.2 Learning Model
Kalbfleisch, in his book: Probability and Statistical Inference: Volume 2, poses as an example of an application of the methods of estimation of parameters by maximum verisimilitude, an original example of the theory of learning, published by R.R. Bush and F. Mosteller. By analogy with the reported experiment, the sequence of tests (assessments) that an individual (student) reports in each evaluation is presented. So be it:
Where, in general, \( i = \{1, 2, \ldots, n\} \) and \( j = \{1, 2, \ldots, k\} \)

The probability of receiving recovery class prior to assessment \( j \) should be considered to be less than the probability in test \( j - 1 \).

In this way according to Kalbfleisch it is stated:

- \( P_j = \text{Probability that a student will receive recovery class before test } j, \text{ given his / her history up to test } j - 1. \) \text{ With } \( j \in \{0, 1, 2, \ldots, k\} \)
- \( X_j = \text{Number of times the student demonstrates learning prior to test } j, \text{ i.e.:} \)

\[
X_j = \sum_{i=0}^{j-1} Y_i
\]

- The number of times an individual has received recovery classes would be \( j - X_j \).
- It is assumed that \( Y_0 = 0 \), that is, at the beginning of a five-month period, all individuals require classes. Thus, \( P_0 = 1 \).
- For \( j > 0 \), \( P_j = A^{X_j} B^{j - X_j} \), with \( 0 \leq A \leq 1, 0 \leq B \leq 1, y \) where:
  - "A" is defined as the learning parameter and
  - "B" the need for recovery class

The probability of receiving recovery classes is reduced by the "A" factor if there was an apprenticeship in the \( j - 1 \) test, or by the "B" factor if there was a need for a recovery class until the \( j - 1 \) test. So:

- If \( A \) is small, then the effect of learning greatly reduces the possibility of future recovery class.
- If \( A = 1 \), Nothing has been learned.
- If \( A < B \), the individual learns more and needs less of a recovery class.

By the total probability theorem, it can be calculated, for an individual \( i \), the joint probability, as follows:

\[
f(Y_i, Y_{i1}, \ldots, Y_{ik}) = f(Y_i) f(Y_i | Y_0) f(Y_{i1} | Y_0, Y_i) \ldots
\]

With \( f(Y_0) = 1 \) and \( f(Y_i | Y_0, Y_{i1}, \ldots, Y_{i-1}) = P_i^{Y_i - Y_j}(1 - P_j)^{Y_j} \)

- The verisimilitude function is established as follows:

\[
\mathcal{L}(A, B | Y_0, Y_{i1}, \ldots, Y_{ik}) = \prod_{i=1}^{k} P_i^{1-Y_i}(1 - P_j)^{Y_j}
\]

- And the log-verisimilitude function:

\[
\ell(A, B) = \sum_{j=1}^{k} (1 - Y_j) \log P_j + Y_j \log(1 - P_j)
\]

- For \( n \) students is gotten the log-verisimilitude function:

\[
\ell(A, B) = \sum_{i=1}^{n} \sum_{j=1}^{k} (1 - Y_{ij}) \log P_{ij} + Y_{ij} \log(1 - P_{ij})
\]

With \( P_{ij} = A^{X_{ij}} B^{j - X_{ij}} \)
For the previously described, it is looked for to estimate the parameters A and B, along with their verisimilitude contours of 14.7%, in order to evaluate significant differences between groups of students and establish the reinforcement needs.

2.3 Procedure

The evaluations of 30 students (12 of one group and 18 of another one) and the times they stayed for recovery classes (0) and the times they did not stay to recovery class (1), during the 14 consecutive weeks of the first five-month. The structure of the database can be seen in Table 1.

Table 1. Scheme of the collected data table

<table>
<thead>
<tr>
<th>Grupo</th>
<th>Individual</th>
<th>E₀</th>
<th>E₁</th>
<th>E₂</th>
<th>...</th>
<th>Eₖ₋₁</th>
<th>Eₖ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>n₁</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>n₂</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Initially, it is estimated the proportion of success, the times they did not stay for recovery class, of each one of the students, it is assumed and then proved, that the proportions of success are adjusted to a Beta distribution, therefore, arises the estimation by maximum verisimilitude of the parameters of such distribution. The analysis of the parameters and their verisimilitude contours at 14.7%, equivalent to 95% confidence (Table 2), allow showing the existence of a significant difference in the proportion of compliance of each group.

Table 2. Compliance ratio scheme by individual

<table>
<thead>
<tr>
<th>Groups</th>
<th>Individual</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.55556</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.61111</td>
</tr>
<tr>
<td>...</td>
<td>n₁</td>
<td>0.77778</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.77778</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.72222</td>
</tr>
<tr>
<td>...</td>
<td>n₂</td>
<td>0.83333</td>
</tr>
</tbody>
</table>

In a second stage, the proposed model is evaluated with the original database. Given the algorithmic complexity of the model, programs are generated in free software R (current Version: 3.3.3), to estimate the parameters of each group in the two proposed stages. It is estimated a verisimilitude range or range of 14.7% verisimilitude, equivalent to 95% confidence, for the parameters. Depending on the overlap of the intervals or contours, it is decided on the difference or not of the parameters.

3 RESULTS

The proportion of students who indicate learning is shown in Fig. 1, that is, they do not stay in a recovery class, it is higher as time passes, however, G2 (Group 2) always shows a bigger proportion of students not taking recovery classes.
Initially, the proportion of success is analyzed, that is the proportion of times that each student does not take recovery class, the table 3, shows the basic statistics of the success rate, times when students do not stay for recovery. It can be observed that the G1 (Group 1) has lower indicators than those of the G2 group, so for example, the average proportion differs by slightly more than 0.15 points, the maximum proportion differs by slightly more than 0.06 points and so in the other indicators; It would seem then that the G2 group achieves better learning outcomes.

<table>
<thead>
<tr>
<th>Group</th>
<th>Minimums</th>
<th>Q1</th>
<th>Q2</th>
<th>Medium</th>
<th>Q3</th>
<th>Max</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>0.4667</td>
<td>0.5167</td>
<td>0.6000</td>
<td>0.5889</td>
<td>0.6667</td>
<td>0.8000</td>
<td>1.399928</td>
<td>9.738017</td>
</tr>
<tr>
<td>G2</td>
<td>0.6000</td>
<td>0.7333</td>
<td>0.7333</td>
<td>0.7444</td>
<td>0.8000</td>
<td>0.8667</td>
<td>3.426652</td>
<td>1.176432</td>
</tr>
</tbody>
</table>

The Kolmogorov-Smirnov test \( p > 0.05 \) and the Q-Q diagram showed that the assumption of adjustment to the Beta distribution with the parameters estimated by maximum verisimilitude (Table 3), can be assumed to be valid. Thus, the estimated beta density graphs and the 14.7% verisimilitude contours are shown in Figure 2 and 3.
These results show that the proportion of success is different in the analyzed groups. When constructing the ROC (Receiver Operating Characteristic) curve, it is clearly observed that in the G1 group the success rate is lower than in the G2 group.

![Fig. 4 Curve ROC by groups](image)

On the other hand, the application of the learning model proposed by J.G. Kalbfleisch presents the following estimators of parameters A and B, which are detailed in Table 4.

<table>
<thead>
<tr>
<th>Groups</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower limit</td>
<td>Upper limit</td>
</tr>
<tr>
<td>G1</td>
<td>0.584482</td>
<td>0.6526755</td>
</tr>
<tr>
<td>G2</td>
<td>0.4610368</td>
<td>0.535656</td>
</tr>
</tbody>
</table>

Whose contours of 14.7% of verisimilitude are shown in Fig. 5.

![Fig. 5 Verisimilitude contours for parameters A and B](image)

From the analysis it results that in both groups there are effects of the learning processes: A < B and A < 1, in group G2 there would be a better effect of the process because the value A is smaller: The smaller A, it implies greater learning effect.
4 CONCLUSIONS

From the analysis of the results, it is concluded that there is a significant difference between the A parameters of the two groups. This difference confirms that in G2 there are better learning outcomes, in addition, it can be stated, how the different learning methodologies applied in the classroom can be evaluated not only in terms of the number of students who are left or not to recovery classes (Fig.); but it is also possible to evaluate the effectiveness of the methodology used when analyzing the value of parameter A, which, as it is smaller, indicates better learning outcomes.

Statistical illiteracy may lead evaluation processes to contain only basic descriptive statistics, which, while showing some difference, do not validate it with technical argumentation. Even the uses of traditional methodologies of statistical inference are potentially misused at the time of such evaluations.

Therefore, it is advisable to implement alternative techniques, although it is the responsibility of those who carry out the evaluation processes, it is no less true that these should come from academia; and it is there that those who are immersed not only in the processes of teaching-learning but also in the management of data analysis tools, can contribute that it takes a lot of interdisciplinary work which accumulates the specificities of each profession, this work, no doubt must be done by those who manage the theoretical or applied research.

REFERENCES

