Nova Scientia
Revista de Investigación de la Universidad De La Salle Bajío

Mathematical modelling of student's cumulative learning
Modelo matemático del aprendizaje acumulativo de un estudiante

Alma Rocío Sagaceta Mejía [1] ✉ - Julián Alberto Fresán Figueroa [2] - Ehyter Matías Martín González [3]

[1] Universidad Iberoamericana, Ciudad de México. Departamento de Física y Matemáticas
[2] Universidad Autónoma Metropolitana, Unidad Cuajimalpa. Departamento de Matemáticas Aplicadas y Sistemas
[3] Universidad de Guanajuato. Departamento de Matemáticas
✉Correspondence author: alma.sagaceta@ibero.mx

Abstract

In this paper we propose a model to study the learning process of one student during a course. We formulate a stochastic model based on the quality of the teacher's class and the affinity of the student to understand the sessions, under the assumption that previous sessions have some influence in the understanding of the next sessions. The afore mentioned assumption implies that the process is not a Markov process. We derive some recursive expressions for the distribution of the number of sessions that the student comprehends. Furthermore, we study the convergence of this distribution and illustrate its speed of convergence through some numerical examples. Finally, we apply these results to propose a methodology to estimate the quality of this kind of courses.

Keywords: stochastic model; learning; non-markovian process; quality of a course; mathematical models; learning processes; courses; educational quality; sessions; distribution; convergence; education; formation

Resumen

En este artículo se propone un modelo para estudiar el proceso de aprendizaje de un alumno durante un curso. Se estableció un modelo estocástico basado en la calidad de la clase del profesor y la afinidad del alumno para comprenderla, bajo el supuesto de que las clases que éste tuvo previamente influyen en su comprensión de las posteriores. La suposición antes mencionada implica que el proceso no es un proceso de Markov. Se obtuvieron algunas expresiones recursivas para la distribución del número de clases que el alumno comprende. Además, se estudió la convergencia de la distribución obtenida y se ilustró la velocidad de convergencia a través de algunos ejemplos numéricos. Finalmente, se aplicaron estos resultados para dar una estimación de la calidad de este tipo de cursos.

Palabras clave: modelo estocástico; aprendizaje; proceso no-markoviano; calidad de un curso; modelos matemáticos; procesos de aprendizaje; cursos; calidad educativa; sesiones; distribución; convergencia; educación; formación

1. Introduction

Simple methods from physics and mathematics have been recently adopted to model, as mathematical metaphors, a wide range of social phenomena and social systems [3, 9, 13, 15, 16, 20]. In particular, stochastic models have been widely used to study learning, particularly Markov models which are a fundamental technique in the mathematical psychology. Markov models are discussed in virtually all the mathematical psychology textbooks (viz., [2]; [5]; [10]; [12]; [17]; [19]; [21]). However, to the best of our knowledge, this kind of models have not been comprehensively used to study the way students learn in a certain class. How student´s behavior affects how well they learn is an old question and has been widely studied in several contexts since people started to concern about teaching and pedagogy [6, 11, 14, 18].

In the present work we propose a model of how a student learns through a course. We assume that the course is presented in a number of sessions such that session j helps the student's understanding of session $j + 1$. This assumption implies that the given model does not rely on a Markov process (see e.g. [8]), because session j depends strongly on all the $j - 1$ previous sessions.

In this paper, the learning processes is modelled with a mathematical perspective, and it is organized as follows: A general description of the model is given in Method section. The mathematical manipulation of the model and some results are presented in Main results section. Afterwards, in Estimation of the quality of the course q section

we apply this model to estimate the quality of a course and present some numerical examples. Finally, conclusions and future work are presented in the last section.

2. Methods, techniques, and instruments

Our base for the suggested model is the modelling of learning presented in [4], which is based on previous works by William K. [7], who proved the accuracy of Markov chains to gain important information on the process of learning. The assumption made by William K. Estes is that, given that the learning process at time n resulted in a successfully understood session, this will influence the result of the $n + 1$-session but will be irrelevant for sessions at times further than $n + 1$. This assumption has proven to be fine for some particular cases (see for instance [[1]]). However, if we consider a course with topics in mathematics, due to the natural structure of this course, the successful or unsuccessful learning of one session has influence not only in the next one, but it might also influence all further sessions. One easy example would be failing to learn sums of fractions, then learning how to solve systems of linear equations with integer coefficients and then moving to linear equations with fractional coefficients. In this context, student might have learned the way to solve systems of linear equations with integer coefficients, but they might fail at learning the same procedure when the coefficients are fractional.

There are some works where independence is not assumed, for instance [4] where they use Markov chains to model the learning process. In this case, the transition probabilities for the associated Markov chain are not constant across the time, although the assumption of a future result being influenced only by the result at present time prevails.

The following data, in figure 1, corresponds to a class taught by one of the authors of the present work. In this class, twenty sessions were evaluated through homework's graded from 0 to 10. The plots correspond to 9 students who took this class, and we note that the values of each grade are not uniformly distributed, which would be the case if such values were independent.
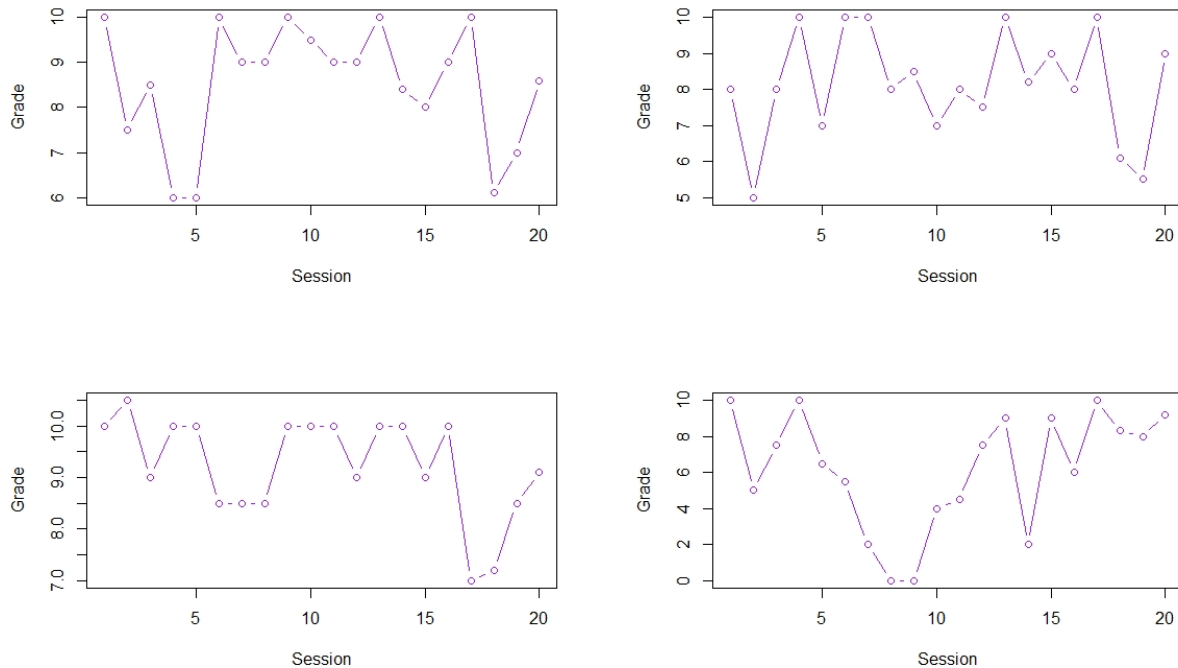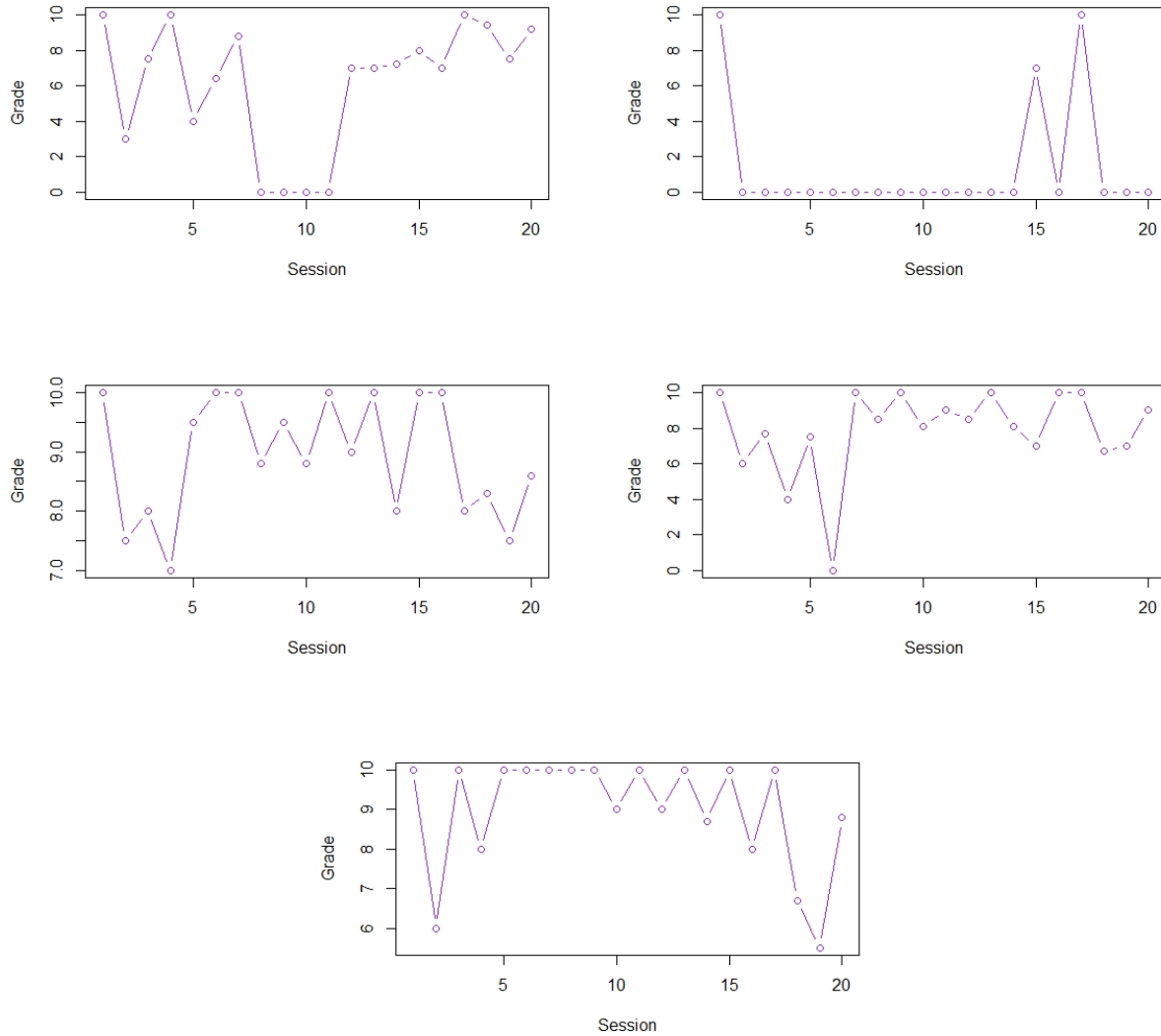
Figure 1. *Data obtained from homework of 9 students in 20 sessions, graded from 0 to 10.*
Figura 1. *Datos obtenidos de las tareas de 9 estudiantes en 20 sesiones, evaluadas de 0 a 10.*

The data above show that assuming a binomial distribution for the number of topics understood by a student is not always realistic. However, it might be the case that dependence grows very small with the understanding of each session, a situation which might reflect that the student eventually had to learn all the material from previous sessions, which they did not understand at first. This could be the situation presented, for instance, in the plots in positions (2,1) and (2,2). In this case a binomial model in the long-time scenario might be accurate.

With all the previous arguments in mind, we consider the case when the learning process has dependence not only on the present session, but also on every previous session. This implies that the probability of successfully learning the content of each session is varying over the time, according to some parameter which measures the influence that previously learned topics from previous sessions have in the present result. This also models the fact that students are accumulating knowledge across time, which helps them to understand in an easier way depending on whether or not they understood all or just some of the previous sessions.

We also assume that, however, after a sufficiently large number of successes or failures, the parameter which measures the influence of past topics will decrease and hence we prove that this results in an asymptotically binomial model for which quantities of interests can be approximated in a simple way.

The initial assumption removes the restriction imposed by the Markov property and provides a different approach in the modelling of learning processes. Our model considers the following assumptions:

1) The course consists in a finite number n of sessions, such that session j helps the student´s understanding of session $j + 1$.

2) The student's learning process is independent of the other students.

3) The lecturer teaches each session according to a *quality parameter q* which represents the quality of the sessions. This parameter remains constant along the course.

4) The student either learns or not the content of each session according to a certain probability distribution F.

5) Each time the student learns or not, the argument of F gets modified by an addend ε, which is assumed to be positive and fixed during the entire course and reflects how the comprehension of the content of a session influences the next session. Hence, we call ε the *dependence parameter*.

6) We wish to avoid the situation when the student understands the last sessions of the course with probability 1, due to the cumulative effect of ε after some point, therefore we assume that ε is $o(f(n))$, for a properly chosen function f depending on the total number of sessions n.

Assumption 2 implies that each student is isolated. The influence of the interaction between the students may be considered for future work under additional assumptions on the dependence parameter ε. Nevertheless, this assumption applies, for instance, to situations like homeschooling, tutorials, postgraduate studies, online classes, and introverted students.

In all this work, the parameter q is used a $1 - q$. The reason for this is merely technical and it means that, in the case of a distribution supported in $(0,1)$, $1 - q$ intuitively measures how poor the quality of sessions was, i.e., the poorer the quality the more difficult it will be for the student to understand the session.
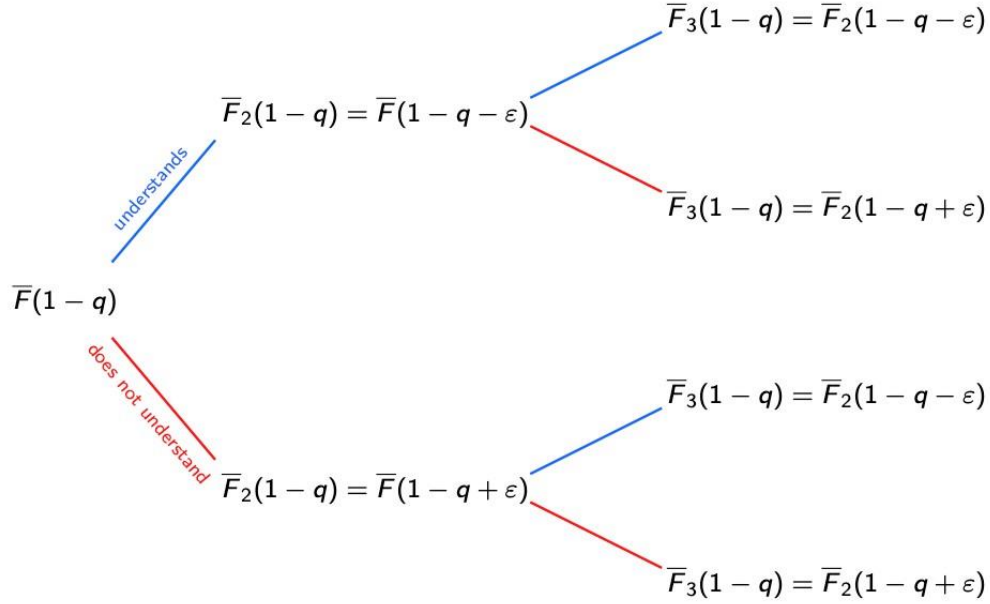
As we will see in one of our simulations, if the distribution F is supported in the whole real line, the value of q might be negative meaning another scale to denote low quality sessions.

The event in which the student understands the first session has a probability given by $\overline{F}(1 - q)$, where $\overline{F} := 1 - F$.

From the second session until the end of the course, if the student *understood* session j, they understand the next one with probability $\overline{F}_{j+1}(1 - q) := \overline{F}_j(1 - q - \varepsilon)$. Here F_1 refers to the initial distribution F and each F_j for $j \geq 2$ is constructed conditioned on the result of all the previous $j - 1$ sessions.

Similarly, if the student did not understand session j, they understand the next one with probability $F_{j+1}(1 - q) := F_j(1 - q + \varepsilon)$. These changes can be visualized in figure 2.

Thus, the dependence parameter, defined in Assumption 5, helps us to keep track on how the results of session j influences session $j + 1$. Whenever a student understands the session j, it is a little more likely that they understand the next one. An analogous situation occurs if the student does not understand session j. In the following sections we manipulate this model to obtain some results related to the distribution of the number of sessions that the student understands along the course.

$$\overline{F}_3(1-q) = \overline{F}_2(1-q-\varepsilon)$$

$$\overline{F}_2(1-q) = \overline{F}(1-q-\varepsilon)$$

understands

$$\overline{F}_3(1-q) = \overline{F}_2(1-q+\varepsilon)$$

$$\overline{F}(1-q)$$

does not understand

$$\overline{F}_3(1-q) = \overline{F}_2(1-q-\varepsilon)$$

$$\overline{F}_2(1-q) = \overline{F}(1-q+\varepsilon)$$

$$\overline{F}_3(1-q) = \overline{F}_2(1-q+\varepsilon)$$

Figure 2. *Outline of the model for $n = 3$.*
Figura 2. *Diagrama del modelo para $n = 3$.*

3. Results and discussion

We let $Y(j)$ be Bernoulli random variables defined as follows: For session 1, $Y(1)$ is Bernoulli with parameter $\overline{F}(1-q)$ for the given initial continuous distribution F. For session 2, by the Law of Total Probability we have:

$$\mathbb{P}[Y(2) = 1] = \mathbb{P}[Y(2) = 1|Y(1) = 1]\overline{F}(1-q) + \mathbb{P}[Y(2) = 1|Y(1) = 0]F(1-q). \qquad (1)$$

Given that the student understood session 1, the probability that they understand session 2 is given by:

$$\mathbb{P}[Y(2) = 1|Y(1) = 1] = \overline{F}(1-q-\varepsilon).$$

Similarly, if the student did not understand session 1, the probability that they understand session 2 is:

$$\mathbb{P}[Y(2) = 1|Y(1) = 0] = \overline{F}(1-q+\varepsilon).$$

It follows that (1) is equivalent to:

$$\mathbb{P}[Y(2) = 1] = \overline{F}(1-q-\varepsilon)\overline{F}(1-q) + \overline{F}(1-q+\varepsilon)F(1-q).$$

For the general setting, we denote by F_j the probability distribution such that:

$$\overline{F}_j(x) = \overline{F}_{j-1}(x-\varepsilon)p_{j-1} + \overline{F}_{j-1}(x+\varepsilon)\big(1-p_j-1\big),$$

Provided that $\overline{F}_{j-1}(x-\varepsilon)$ and $\overline{F}_{j-1}(x+\varepsilon)$ do not equal zero or one. In the equation above, $p_k := \mathbb{P}[Y(k) = 1]$ for $k > 1$ with $p_1 = \overline{F}(1-q)$ and $F_1 := F$. Using this notation, we obtain:

$$p_j = \overline{F}_{j-1}(1 - q - \varepsilon)p_{j-1} + \overline{F}_{j-1}(1 - q + \varepsilon)(1 - p_{j-1}). \qquad (2)$$

Throughout the rest of this paper, we assume that $1 - q \pm n\varepsilon$ is such that $0 < F(1 - q \pm n\varepsilon) < 1$, for each n. In the following result we derive a recursive expression for the probabilities $\{p_m, 1 < m \leq n\}$.

Theorem 1. The general expression for $\{p_m, 1 < m \leq n\}$ reads:

$$p_m = \overline{F}_1(1 - q) \prod_{j=1}^{m-1}[\overline{F}_j(1 - q - \varepsilon) - \overline{F}_j(1 - q + \varepsilon)] + \sum_{j=i+1}^{m-1}[\overline{F}_j(1 - q - \varepsilon) - \overline{F}_j(1 - q + \varepsilon)]. \qquad (3)$$

Proof. See Appendix ∎

From this point on, we drop the notation F_1 and write F instead.

We let B_n denote the number of sessions, not necessarily consecutive, from a total of n that the student understood. We are interested in the probability function of B_n, $\{\mathbb{P}[B_n = k], 0 \leq k \leq n\}$, for which we consider the following particular cases.

Case 1: $n = 3$ and $k = 0$.
This is the case when the student understands 0 sessions out of 3. The probability reads:

$$\mathbb{P}[B_3 = 0] = F(1 - q)F(1 - q + \varepsilon)F(1 - q + 2\varepsilon).$$

Case 2: $n = 3$ and $k = 1$.
If student understand only one session (i.e., $k = 1$), the probability is the sum of the following cases.

2.1. The student understands only the first session:

$$\overline{F}(1 - q)F(1 - q - \varepsilon)F(1 - q).$$

2.2. Student i understands only the second session:

$$F(1 - q)\overline{F}(1 - q + \varepsilon)F(1 - q).$$

2.3. And student i understands only the third session:

$$F(1 - q)F(1 - q + \varepsilon)\overline{F}(1 - q + 2\varepsilon).$$

Note that the probabilities in cases 2.1 and 2.2 equal $\mathbb{P}[B_2 = 0]F(1 - q)$ and the probability of case 2.3 corresponds to $\mathbb{P}[B_2 = 1]F(1 - q + 2\varepsilon)$. Hence:

$$\mathbb{P}[B_3 = 1] = \mathbb{P}[B_2 = 0]F(1 - q) + \mathbb{P}[B_2 = 0]F(1 - q).$$

Case 3: $n = 3$ and $k = 2$. If student understand two sessions (i.e., $k = 2$), the probability is the sum of the following cases.

3.1. Student i understands the first and second sessions but not the third:

$$\overline{F}(1 - q)\overline{F}(1 - q - \varepsilon)F(1 - q - 2\varepsilon).$$

3.2. Student i understands the second and third sessions but not the first:

$$F(1-q)\overline{F}(1-q+\varepsilon)\overline{F}(1-q).$$

3.3. Student i understands the first and third sessions but not the second:

$$\overline{F}(1-q)F(1-q-\varepsilon)\overline{F}(1-q).$$

Note that cases 2 and 3 correspond to $\mathbb{P}[B_2=1]\overline{F}(1-q)$ and the first case equals $\mathbb{P}[B_2=2]F(1-q-2\varepsilon)$. Hence:

$$\mathbb{P}[B_3=2]=\mathbb{P}[B_2=1]\overline{F}(1-q)+P[B_2=2]\overline{F}(1-q-2\varepsilon).$$

Case 4: $n=3$ and $k=3$.
This is the case when student i understands all sessions. This probability is given by:

$$\mathbb{P}[B_3=3]=\overline{F}(1-q)\overline{F}(1-q-\varepsilon)\overline{F}(1-q-2\varepsilon).$$

The recursive behavior observed in the probability function of B_3 is generalized in the following theorem.
Theorem 2: *Let $n\geq 3$ be an integer. The probability function of the random variable B_3 satisfies the following relations.*

1. $\mathbb{P}[B_n=0]=\mathbb{P}[B_{n-1}=0]F(1-q+(n-1)\varepsilon),$

2. $\mathbb{P}[B_n=n]=\mathbb{P}[B_{n-1}=n-1]F(1-q-(n-1)\varepsilon),$

3. $\mathbb{P}[B_n=k]=\mathbb{P}[B_{n-1}=k]F(1-q+(n-1-2k)\varepsilon)+\mathbb{P}[B_{n-1}=k-1]\overline{F}(1-q+(n-1-2(k-1))\varepsilon).$

Proof. See Appendix ∎

Equations in Theorem 2 can be written as a single matrix equation. Let $\overrightarrow{B_n}\in\mathbb{M}_{1,n+1}$ be given by:

$$\overrightarrow{B_n}=[\mathbb{P}[B_n=0],\mathbb{P}[B_n=1],\mathbb{P}[B_n=2],\dots,\mathbb{P}[B_n=n]],$$

And denote by $\mathcal{M}_n\in\mathbb{M}_{1,n+1}$ the matrix such that:

$$(\mathcal{M}_n)_{a,b}=\begin{cases}F(1-q-(n-1-2a)\varepsilon),&for\ a=b\\ \overline{F}(1-q-(n-1-2a)\varepsilon),&for\ a=b-1\\ 0,&otherwise.\end{cases}\tag{4}$$

Using the notation above we note that $\overrightarrow{B_n}=\mathcal{M}_n\cdot\overrightarrow{B}_{n-1}$, therefore:

$$\overrightarrow{B_n}=\prod_{k=1}^{n}\mathcal{M}_k\cdot\overrightarrow{B}_{0,i},$$

Where: $\overrightarrow{B_0}=[F(1-q),\overline{F}(1-q)]$. This representation is used for some numerical examples in the figure section.

The explicit distribution of B_n is not easy to obtain even in simple cases (such as the case when F is a uniform distribution). Nevertheless, in the following result we provide a simple asymptotic expression for this distribution.

Theorem 3: *Let $\{p_n(k), k = 0,1, \dots, n\}$ denote the probability function of a Binomial (n, p) distribution, with $p :=$ $\overline{F}(1 - q)$. Suppose n, ε are such that $n^2 \varepsilon \to 0$ as $n \to \infty$ and F is absolutely continuous with density f such that f is continuous at $1 - q$, then*:

$$\lim_{n \to \infty} \frac{\mathbb{P}[B_n = k]}{p_n(k)} = 1, \qquad \forall k \in \{0,1, \dots, n\}.$$

Proof. See Appendix ∎

Theorem 3 says that for a sufficiently large number of sessions the dependence becomes less relevant, hence the number of sessions that each student understands behaves like a binomial distribution in which the occurrence of successes is independent. The convergence referenced in Theorem 3 does not imply, in general, the convergence between the means of B_n and the limit binomial random variable. However, in this case such convergence holds.

Theorem 4: $\dfrac{E[B_n]}{n\overline{F}(1-q)} \to 1$ as $n \to \infty$.

Proof. See Appendix ∎

Three numerical examples of the result above are given in figure III (Appendix), where we note that both means, the simulated and the approximated one, behave similarly. On the other hand, the values of the quotient $\dfrac{\mathcal{E}[B_n]}{n\overline{F}(1-q)}$ are mostly in the interval $(0.98, 1.02)$.

Estimation of the quality of the course q: It is possible to use the approximation in Theorem 4 to estimate the quality parameter q. This is stated and proved in the following result.

Theorem 5: Assume that $\overline{F}$ has a continuous inverse function $\overline{F}^{-1}$ and set $q_n := 1 - \overline{F}^{-1}\left(\dfrac{E[B_n]}{n}\right)$. Then $q_n \to q$ as $n \to \infty$.

Proof. See Appendix ∎

Example 1. For the data considered at the beginning of Section 2, we have 22 students taking the same course and we use their grades to determine whether they understood the lessons or they did not (we assume that grades equal or greater than 6 indicate the student understood the lesson).

The whole exploratory data indicated that the grades of students are not identically distributed and thus, we approximated the value of q for each one of these students using the result of Theorem 4 and the empirical estimator of $E[B_n]$. The development of a statistical tool to test the goodness of fit from data may be an interdisciplinary option for future work. These results are presented in the table below.

Student	1	2	3	4	5	6	7	8	9	10	11
$1 - q$	9.50	9.00	10.00	9.00	9.20	0.00	10.00	9.00	10.00	10.00	9.50

Student	12	13	14	15	16	17	18	19	20	21	22
$1 - q$	9.00	10.00	9.70	8.50	10.00	9.50	10.00	9.50	10.00	9.50	9.20

The average of this estimations is 9.1. *In the official instrument, made by the university, where the students grade the quality of the course, the same course was graded* 9.6. ∎

By Theorem 5 we have:

$$q_n \simeq 1 - \overline{F}^{-1}\left(\frac{E[B_m]}{n}\right) \Rightarrow q_n \simeq log\left(\frac{E[B_n]}{n-E[B_n]}\right) + 1. \qquad (5)$$

This estimation of q under this model is illustrated in figure IV (Appendix) in plots (a) and (b), where we also consider a standard normal distribution (plots (c) and (d)). We note that the values of q in these examples provide a criteria to determine if the quality of the course' sessions has been poor or not. For instance, as we see from the perpendicular dashed lines, $q = 1$ implies a course in which the student understood just half the sessions, which might be considered as low quality course.

In figure I (Appendix) q and simulate the exact distribution of B_n. Then we compare it to the approximating binomial distribution given in Theorem 3 for number of sessions given by $n = 10,30,60,100$ (plots in blue, red, green, and purple respectively). The light line corresponds to the approximating binomial distribution while the dark line represents the simulated exact distribution. We see in these plots that the convergence to the binomial distribution is quite fast, and it grows faster when $q = 0.5$.

Our numerical examples show that the speed of convergence depends on the value of q, as it can be seen in figure II (Appendix) q with fixed n and compare the exact distribution of B_n to the approximating binomial distribution.

Another point worth mentioning is that the distributions behave symmetrically with respect to $q = 0.5$. In this case, the convergence to the binomial distribution is faster than in the other cases. In fact, the cases when q is nearly 0 or 1, present a slower convergence to the binomial distribution. This may imply that the dependence is stronger when the quality of the class is low or high.

4. Conclusions

In this work we present a model to study the behavior of the understanding of a student along several sessions of a course without interaction with other students. In particular, we study the case where the dependence between sessions is relatively small compared to the total number of sessions, as in seminars or panoramic courses. We obtained a recursive expression for the distribution of the number of sessions that the student understands along the course and showed that when the dependence parameter is small, this distribution has a binomial approximation. The speed of convergence of the approximation depends on the number of sessions and the quality of them. Even though this is a simple model, it can be fruitfully extended to consider many more situations, some of which we list below:

1) The environment we considered assumes the student has no interaction with their classmates. However, several studies have shown that collaborative learning provides better results for students. It would be very interesting to modify the model to consider this situation.
2) We studied the case when ε is constant and relatively small compared to n, but this may not always be the case, as in some science classes. Therefore, it would be useful to consider the cases when ε changes according to the sessions themselves or according to the number of sessions previously understood.
3) All the results obtained in this work were made for q constant, but the value of q can vary along the course due to exhaustion and motivation of the student and the teacher.
4) It would be interesting to test the model with more real world data and develop some statistical procedures for the model to be fitted and validated.

5. Supplementary Information
Appendix.

6. Acknowledgements
No.

Authors Information

Alma Rocío Sagaceta Mejía [1] [ID] 0000-0001-8265-3379
Julián Alberto Fresán Figueroa [2] [ID] 0000-0002-9834-8157
Ehyter Matías Martín González [3] [ID] 0000-0003-3824-699X

Contribution of the authors in the development of the work

The authors declare that they contributed equally.

Interest conflict

The authors declare that there is no conflict of interest.

References

[1] Bower GH. Application of a model to paired-associate learning. Psychometrika. 1961; 26(3): 255-280. Disponible en: https://link.springer.com/article/10.1007/BF02289796

[2] Busemeyer JR, Diederich A. Cognitive modeling. Sage; 2010.

[3] Cajueiro DO. Enforcing social behavior in an Ising model with complex neighborhoods. Physica A: Statistical Mechanics and its Applications, 2011; 390(9):1695 – 1703. Disponible en: https://www.sciencedirect.com/science/article/pii/S0378437111000501

[4] Chechile RA, Sloboda LN. Reformulating Markovian processes for learning and memory from a hazard function framework. Journal of Mathematical Psychology, 2014; 59: 65–81. Special Issue in Honor of William K. Estes. Disponible en: https://www.sciencedirect.com/science/article/abs/pii/S0022249613000862

[5] Coombs CH, Dawes RM, Tversky A. Mathematical psychology: An elementary introduction. 1970.

[6] Dufresne RJ, Gerace WJ, Leonard WJ, Mestre JP, Wenk L. Classtalk: A classroom communication system for active learning. Journal of Computing in Higher Education. 1996 Mar; 7(2):3–47. Disponible en https://link.springer.com/article/10.1007/BF02948592.

[7] Estes WK, Suppes P. Foundations of stimulus sampling theory. 1974.

[8] Huang YM, Huang TC, Wang KT, Hwang WY. A Markov-based Recommendation Model for Exploring the Transfer of Learning on the Web. Journal of Educational Technology and Society. 2009; 12(2):144–162. Disponible en https://www.jstor.org/stable/jeductechsoci.12.2.144.

[9] Klemm K. Nonequilibrium transitions in complex networks: A model of social interaction. Physical Review E, 2003; 67(2). Disponible en: https://journals.aps.org/pre/abstract/10.1103/PhysRevE.67.026120

[10] Laming F, Straile R. Photoresist cleaning. Journal of The Electrochemical Society, 1973; 120(2):292.

[11] Lee YA. Third turn position in teacher talk: Contingency and the work of teaching. Journal of Pragmatics. 2007; 39(6):1204 – 1230. Disponible en: http://www.sciencedirect.com/science/article/pii/S037821660600227X.

[12] Levine G, Burke CJ. Mathematical model techniques for learning theories. 1972.

[13] Li C, Liu F, Li P. Ising model of user behavior decision in network rumor propagation. Discrete Dynamics in Nature and Society. 2018 Aug; 2018:1–10. Disponible en: https://www.hindawi.com/journals/ddns/2018/5207475/.

[14] McCroskey JC. Classroom consequences of communication apprehension. Communication Education. 1977 Jan; 26(1):27–33. Disponible en: http://dx.doi.org/10.1080/03634527709378196.

[15] Nakayama S and Nakamura Y. A fashion model with social interaction. Physica A: Statistical Mechanics and its Applications. 2004; 337(3):625 – 634. Disponible en: http://www.sciencedirect.com/science/article/pii/S037843710400113X.

[16] Ostilli M, Yoneki E, Leung IXY, Mendes JFF, Lió P, Crowcroft J. Ising model of rumour spreading in interacting communities. University of Cambridge, Computer Laboratory; 2010. UCAM-CL-TR-767. Disponible en: https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-767.pdf.

[17] Restle F, Greeno JG. Introduction to mathematical psychology. 1970.

[18]   Richmond VP. Communication in the classroom: Power and motivation. Communication Education. 1990 Jul; 39(3):181–195. Disponible en: http://dx.doi.org/10.1080/03634529009378801.

[19]   Roberts FS. Discrete mathematical models, with applications to social, biological, and environmental problems. 1976.

[20]   Stauffer D. Social applications of two-dimensional ising models. American Journal of Physics. 2008 Apr; 76(4):470–473. Disponible en: https://doi.org/10.1119/1.2779882

[21]   Wickens TD. Models for behavior: Stochastic processes in psychology. WH Freeman & Co Ltd; 1982.
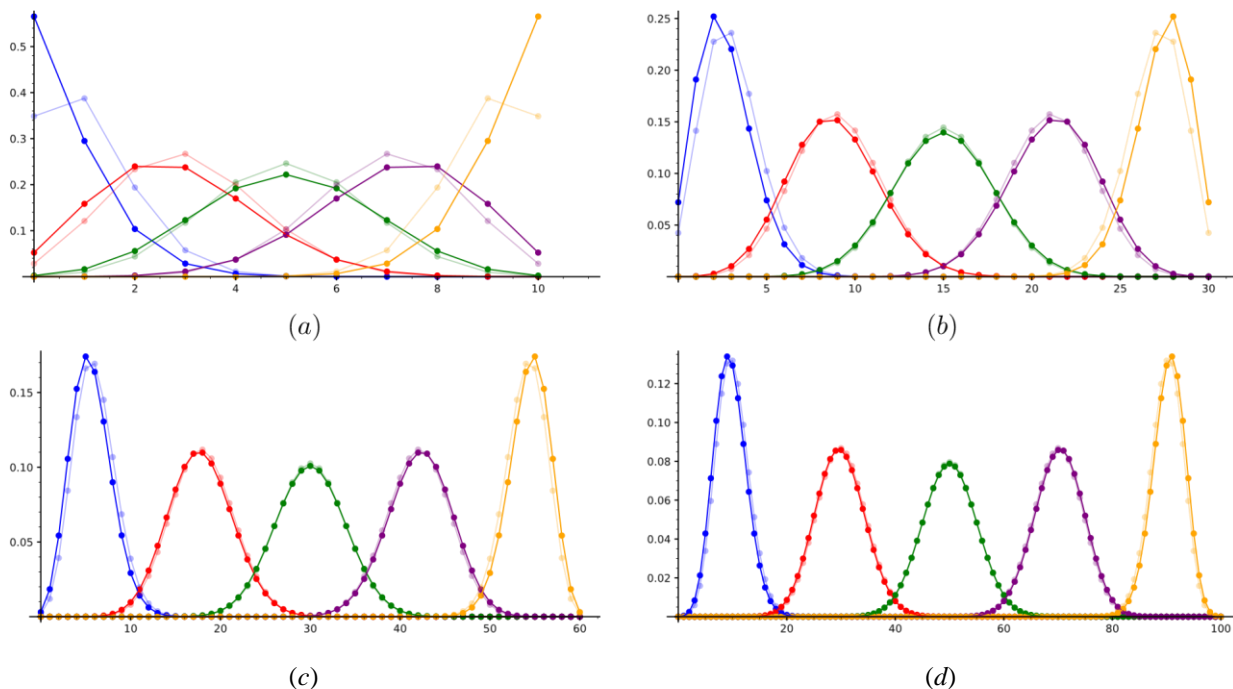
Appendix
Appendix A. Figures.
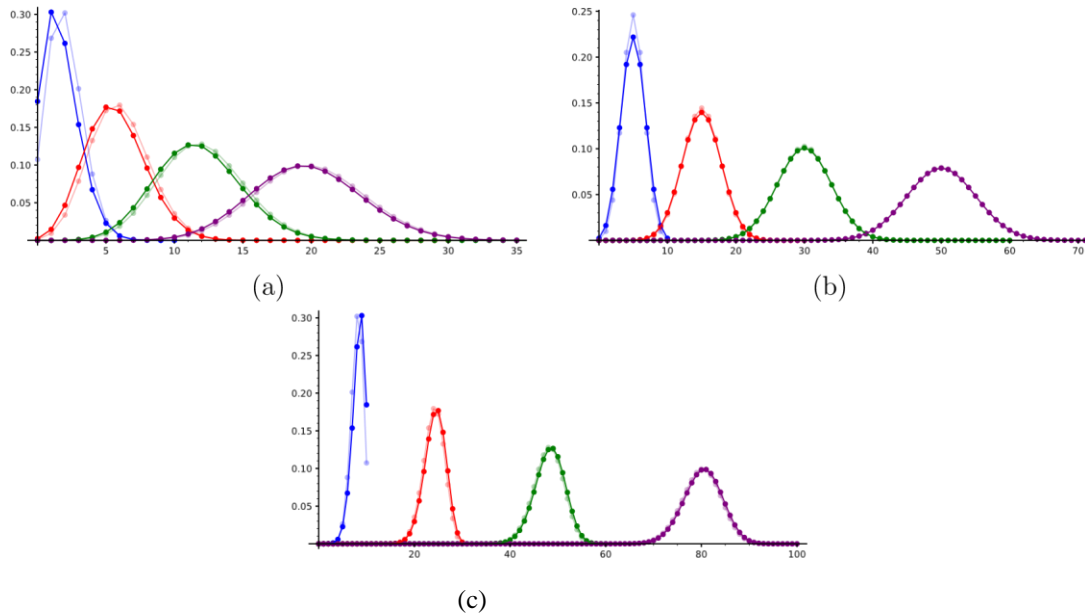
Figure I. *In this figure are plotted some simulations of B_n (dark line) and its approximating binomial distribution (light line). In all the plots, it was considered: $\varepsilon = 1/n^2$ and $n = 10$ in blue; $n = 30$ in red; $n = 60$ in green and $n = 100$ in purple. Plot* (a) *is made for $q = 0.2$; plot* (b) *is made for $q = 0.5$, and plot* (c) *for $q = 0.8$.*
Figura I. *En esta figura se grafican algunas simulaciones de B_n (línea oscura) y su aproximación a una distribución binomial (línea clara). En todas las gráficas se consideró: $\varepsilon = 1/n^2$ y $n = 10$ en azul; $n = 30$ en rojo; $n = 60$ en verde y $n = 100$ in morado. La gráfica* (a) *está hecha con $q = 0.2$; la* (b) *con $q = 0.5$, y la* (c) *con $q = 0.8$.*

Figure II. *In this figure are plotted some simulations of B_n (dark line) and its approximating binomial distribution (light line). In all the plots, it was considered: $\varepsilon = 1/n^2$ and $q = 0.1$ in blue; $q = 0.3$ in red; $q = 0.5$ in green; $q = 0.7$ in purple and $q = 0.9$ in yellow. Plot* (a) *is made for $n = 10$; plot* (b) *is made for $n = 30$; plot* (c) *is made for $n = 60$, and plot* (d) *for $n = 100$*

Figura II. *En esta figura se grafican algunas simulaciones de B_n (línea oscura) y su aproximación a una distribución binomial (línea clara). En todas las gráficas se consideró: $\varepsilon = 1/n^2$ y $q = 0.1$ en azul; $q = 0.3$ en rojo; $q = 0.5$ en verde; $q = 0.7$ in morado y $q = 0.9$ en amarillo. La gráfica* (a) *está hecha con $n = 10$; la* (b) *con $n = 30$; la* (c) *con $n = 60$, y la* (d) *con $n = 100$.*
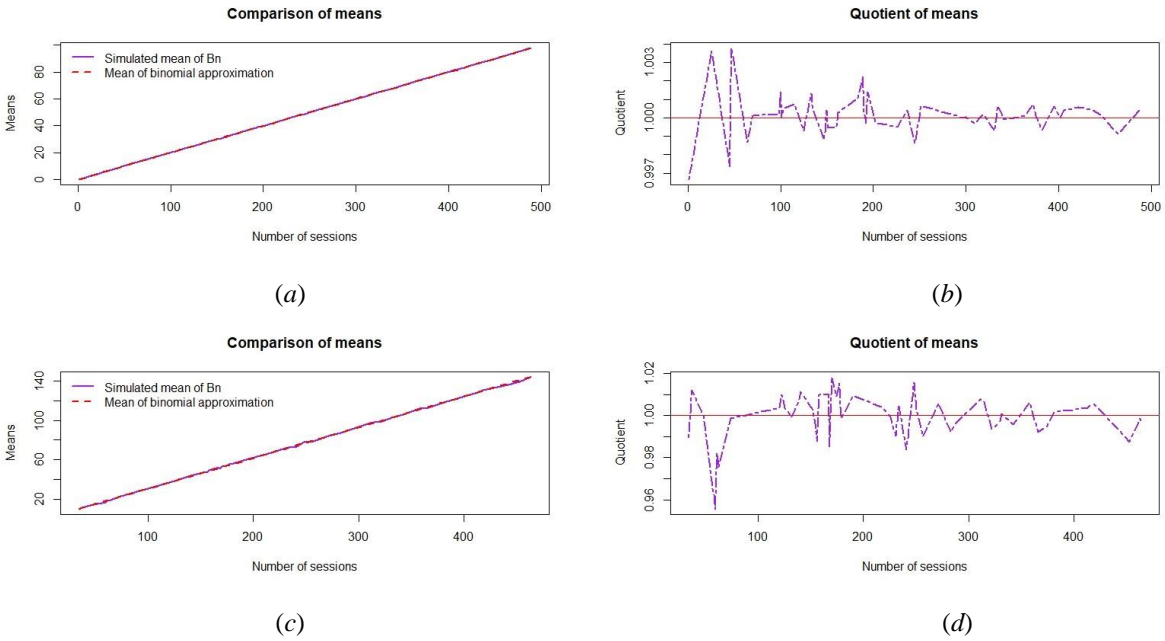
Figure III. *It was simulated the mean of B_n for n up to 500 and it was compared with the mean of the binomial approximation for $q = 0.2$. Similarly, the quotients were plotted $\frac{E[B_n]}{n\overline{F}(1-q)}$ for the cases when F is the uniform distribution* (a)-(b) *and the logistic distribution* (c)-(d).

Figura III. *Se simuló la media de B_n para n hasta 600 y se comparó con la media de la aproximación binomial para $q = 0.2$. De igual manera se graficaron los cocientes $\frac{E[B_n]}{n\overline{F}(1-q)}$ para los casos en que F es la distribución uniforme* (a)-(b) *y la distribución logística* (c)-(d).

(a)

(b)

(c)

(d)
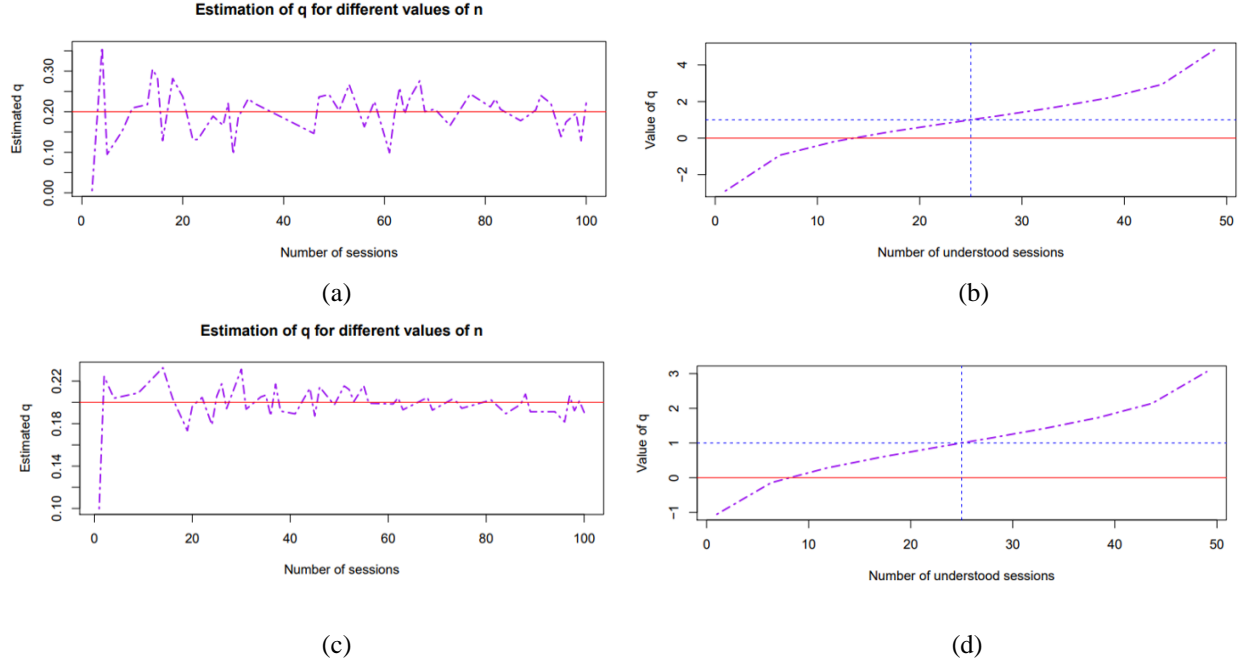
Figure IV. *In this plot it is illustrated the estimation of q using Theorem 5 for* (a)-(b) *logistic model and* (c)-(d) *standard normal model. Plots* (b) *and* (d) *show the behavior of q varying the values* $E\left[B_n\right]$ *when* $n = 50$.
Figura IV. *En esta gráfica se ilustra la estimación de q usando el Teorema 5 para* (a)-(b) *una distribución logística y* (c)-(d) *una distribución normal estándar. Las gráficas* (b) *y* (d) *muestran el comportamiento de q mientras varían los valores de* $E\left[B_n\right]$ *cuando* $n = 50$.

Appendix B. Proof of the main theorems
Proof of Theorem 1. The result holds for $m = 1$ by the definition of p_1. We proceed by induction, assuming that for an integer $k \geq 1$:

$$p_k = \overline{F}_1(1-q)\prod_{j=1}^{k-1}\left[\overline{F}_j(1-q-\varepsilon)-\overline{F}_j(1-q+\varepsilon)\right] + \sum_{i=1}^{k-1}\overline{F}_i(1-q+\varepsilon)\prod_{j=i+1}^{k-1}\left[\overline{F}_j(1-q-\varepsilon)-\overline{F}_j(1-q+\varepsilon)\right].$$

By equation (2):

$$p_{k+1} = \overline{F}_k(1-q-\varepsilon)p_k + \overline{F}_k(1-q+\varepsilon)(1-p_k) = p_k\left[\overline{F}_k(1-q-\varepsilon)-\overline{F}_k(1-q+\varepsilon)\right] + \overline{F}_k(1-q+\varepsilon).$$

Hence, we obtain from the induction hypothesis:

$$p_{k+1} = \left[\,\overline{F}_k(1-q-\varepsilon) - \overline{F}_k(1-q+\varepsilon)\right]\overline{F}_1(1-q)\prod_{j=1}^{k-1}\left[\,\overline{F}_j(1-q-\varepsilon) - \overline{F}_j(1-q+\varepsilon)\right]$$

$$+ \left[\,\overline{F}_k(1-q-\varepsilon) - \overline{F}_k(1-q+\varepsilon)\right]\sum_{i=1}^{k-1}\overline{F}_i(1-q+\varepsilon)\prod_{j=i+1}^{k-1}\left[\,\overline{F}_j(1-q-\varepsilon) - \overline{F}_j(1-q+\varepsilon)\right]$$

$$+ \overline{F}_k(1-q+\varepsilon)$$

$$= \overline{F}_1(1-q)\prod_{j=1}^{k}\left[\overline{F}_j(1-q-\varepsilon) - \overline{F}_j(1-q+\varepsilon)\right]$$

$$+ \sum_{i=1}^{k}\overline{F}_i(1-q+\varepsilon)\prod_{j=i+1}^{k}\left[\,\overline{F}_j(1-q-\varepsilon) - \overline{F}_j(1-q+\varepsilon)\right].$$

The result now follows ∎

Proof of Theorem 2.

1. If the student has not understood the first $n-1$ sessions from a total of n, it follows from the construction of the model that the student's parameter for understanding the nth-session becomes $1-q+(n-2)\varepsilon$. Hence:

$$\mathbb{P}[B_n = 0] = \mathbb{P}[B_{n-1} = 0, Y(n) = 0] = \mathbb{P}[Y(n) = 0|B_{n-1} = 0]\mathbb{P}[B_{n-1} = 0] = F(1-q+(n-1)\varepsilon)\mathbb{P}[B_{n-1} = 0].$$

2. Let A_n denote the number of sessions that the student has not understood from a total of n. Then the event $B_n = n$ is the same as $\{A_n = 0\}$ and hence $\mathbb{P}[B_n = n] = \mathbb{P}[A_n = 0]$ Now the result in 1 yields:

$$\mathbb{P}[B_n = n] = \mathbb{P}[A_{n-1} = 0]\big(1 - F(1-q+(n-1)\varepsilon)\big) = \mathbb{P}[B_{n-1} = n-1]\,\overline{F}\,(1-q+(n-1)\varepsilon).$$

3. Let $U_k(n):= \{(x_1,...,x_n) \in \{0,1\}^n : x_1 + \cdots + x_n = k\}$, then:

$$\mathbb{P}[B_n = k] \tag{6}$$

$$= \mathbb{P}[B_{n-1} = k, Y(n) = 0] + \mathbb{P}[B_{n-1} = k-1, Y(n) = 1]$$

$$= \sum_{(x_1,...,x_{n-1})\in U_k(n-1)} \mathbb{P}[Y(1) = x_1, ..., Y(n-1) = x_{n-1}, Y(n) = 0]$$

$$+ \sum_{(x_1,...,x_{n-1})\in U_k(n-1)} \mathbb{P}[Y(1) = x_1, ..., Y(n-1) = x_{n-1}, Y(n) = 1]$$

$$= \sum_{(x_1,...,x_{n-1})\in U_k(n-1)} \mathbb{P}[Y(n) = 0 \mid Y(1) = x_1, ..., Y(n-1) = x_{n-1}]\mathbb{P}[Y(1) = x_1, ..., Y(n-1) = x_{n-1}]$$

$$+ \sum_{(x_1,...,x_{n-1})\in U_k(n-1)} \mathbb{P}[Y(n) = 1 \mid Y(1) = x_1, ..., Y(n-1) = x_{n-1}]\mathbb{P}[Y(1) = x_1, ..., Y(n-1)$$
$$= x_{n-1}]. \tag{7}$$

Note that, given the configuration $Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}$ in which the student understood exactly k sessions, we have added k times ε to the quality parameter q. Moreover, the maximum number of times we may add or subtract ε in a total of n sessions equals $n-1$, since we do not add or subtract anything in session one. Hence if the student understood k of n sessions, they did not understand $n-k$ and we have subtracted $n-1-k$ times ε. This means that understanding session n depends on the parameter:

$$q + k\varepsilon - (n-1-k)\varepsilon = q - (n-1-2k)\varepsilon.$$

It follows that the student does not understand session n with probability $F(1 - q - (n-1-2k)\varepsilon)$. Since this holds for any given configuration $Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}$, in which the student has understood exactly k sessions, we obtain:

$$\sum_{(x_1,\ldots,x_{n-1}) \in U_k(n-1)} \mathbb{P}[\, Y(n) = 0 \mid Y(1) = x_1, \ldots, Y(n-1) = x_{n-1} \,] \mathbb{P}[Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}]$$

$$= F(1 - q - (n-1-2k)\varepsilon) \sum_{(x_1,\ldots,x_{n-1}) \in U_k(n-1)} \mathbb{P}[Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}]$$

$$= F(1 - q - (n-1-2k)\varepsilon)\mathbb{P}[B_{n-1} = k]. \tag{8}$$

Analogously:

$$\sum_{(x_1,\ldots,x_{n-1}) \in U_k(n-1)} \mathbb{P}[\, Y(n) = 1 \mid Y(1) = x_1, \ldots, Y(n-1) = x_{n-1} \,] \mathbb{P}[Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}]$$

$$= \overline{F}\big(1 - q - \big(n-1-2(k-1)\big)\varepsilon\big)\mathbb{P}[\, B_{n-1} = k-1]. \tag{9}$$

The result follows by substituting equations (8) and (9) in equation (7).

Proof of Theorem 3. First, we prove the case when $k \notin \{0, n\}$. Following the arguments in the proof of Theorem 2, we have:

$$\mathbb{P}[B_n = k] = F(1 - q - (n-1-2k)\varepsilon) \sum_{(x_1,\ldots,x_{n-1}) \in U_k(n-1)} \mathbb{P}[Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}]$$

$$+ \overline{F}\big(1 - q - \big(n-1-2(k-1)\big)\varepsilon\big) \sum_{(x_1,\ldots,x_{n-1}) \in U_{k-1}(n-1)} \mathbb{P}[Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}], \tag{10}$$

Where:

$$U_k(n) = \{(x_1, \ldots, x_n) \in \{0,1\}^n : x_1 + \cdots + x_n = k\}.$$

From the hypothesis $n^2\varepsilon \to 0$, it follows that $\varepsilon \to 0$. Hence, by the assumption of continuity of F, the following convergences as $n \to \infty$ hold:

$$F(1 - q - (n-1-2k)\varepsilon) \to F(1-q) \ and \ \overline{F}\big(1 - q - \big(n-1-2(k-1)\big)\varepsilon\big) \to \overline{F}(1-q).$$

Note from the matrix representation given in (4) that $\mathbb{P}[Y(1) = x_1, \ldots, Y(n-1) = x_{n-1}]$ can be expressed as:

$$p_{n-1,k} := \prod_{j=1}^{n-1} \overline{F}^{v_1(j)}(1 - q - u_1(j)\varepsilon + u_0(j)\varepsilon)F^{1-v_1(j)}[1 - q - u_1(j)\varepsilon + u_0(j)\varepsilon],$$

Where:

$$u_1(j) = \sum_{a=1, x_a=1}^{j}(-1)^{x_a}, \quad u_0(j) = \sum_{a=1, x_a=0}^{j}(-1)^{x_a},$$

And $v_1(j) = 1$ if student understood session j. Now we have the following bounds for $p_{n-1,k}$:

$$\overline{F}^k(1 - q + u_0(j)\varepsilon)F^{n-1-k}(1 - q - u_1(j)\varepsilon) \le p_{n-1,k} \le \overline{F}^k(1 - q - u_1(j)\varepsilon)F^{n-1-k}(1 - q + u_0(j)\varepsilon). \quad (11)$$

Since the terms with the tail $\overline{F}$ converge to $\overline{F}(1 - q)$ and their exponents do not depend on n, we only need to prove that:

$$\frac{F^{n-1-k}(1 - q - u_1(j)\varepsilon)}{F^{n-1-k}(1 - q)} \to 1, \quad n \to \infty,$$

Or equivalently:

$$(n - 1 - k)\log\left(\frac{F(1 - q - u_1(j)\varepsilon)}{F(1 - q)}\right) \to 0, \quad n \to \infty.$$

Using the hypothesis $n^2\varepsilon \to 0$ we may write $\varepsilon = cn^{-2-\eta}$ for some $c, \eta > 0$. Applying L'Hôpital's rule we obtain:

$$\lim_{n\to\infty} n\log\left(\frac{F(1 - q - u_1(j)\varepsilon)}{F(1 - q)}\right) = cu(j)(2 + \eta)\lim_{n\to\infty}\frac{n^{-3-\eta}f(1 - q - u_1(j)cn^{-2-\eta})}{(-n^{-2})F(1 - q - u_1(j)cn^{-2-\eta})} = 0.$$

From the limit above and (11) it follows that:

$$\frac{p_{n-1,k}}{\overline{F}^k(1 - q)F^{n-1-k}(1 - q)} \to 1, n \to \infty. \quad (12)$$

Now let us consider the term:

$$\sum_{(x_1,\dots,x_{n-1})\in U_k(n-1)} \mathbb{P}[Y(1) = x_1, \dots, Y(n - 1) = x_{n-1}].$$

Since $|U_k(n - 1)| = \binom{n-1}{k}$, using the result in equation (12), for an arbitrary $\beta > 0$ and sufficiently large n we have:

$$1 - \beta \le \sum_{(x_1,\dots,x_{n-1})\in U_k(n-1)} \frac{\mathbb{P}[Y(1) = x_1, \dots, Y(n - 1) = x_{n-1}]}{\binom{n-1}{k}\overline{F}^k(1 - q)F^{n-1-k}(1 - q)} \le 1 + \beta.$$

The result follows by letting $n \to \infty$ and $\beta \to 0$. The result for the second term in equation (10) is obtained analogously. Now we proceed in a similar way to prove that $\mathbb{P}[B_n = 0]$ is asymptotically equivalent to $p_n(0)$. It might be easily checked that $\mathbb{P}[B_n = 0] = F(1-q) \prod_{j=1}^{n-1} F(1-q+j\varepsilon)$, hence:

$$1 = \left(\frac{F(1-q)}{F(1-q)}\right)^n \leq \frac{F(1-q) \prod_{j=1}^{n-1} F(1-q+j\varepsilon)}{F^n(1-q)} \leq \left(\frac{F(1-q+n\varepsilon)}{F(1-q)}\right)^{n-1}.$$

Using the representation $\varepsilon = cn^{-\eta-2}$ and L'Hôpital's rule again, we obtain:

$$\lim_{n\to\infty} (n-1) \log\left(\frac{F(1-q+n\varepsilon)}{F(1-q)}\right) = -c(\eta+2) \lim_{n\to\infty} (n-1)^2 \, n^{-\eta-2} F(1-q+cn^{-\eta-1}) f(1-q+cn^{-\eta-1}) = 0.$$

The proof for $\mathbb{P}[B_n = n]$ is analogous. ∎

Proof of Theorem 4. We let $a_n = \sum_{j=1}^n j \mathbb{P}[B_n = j]$ and $b_n = n\overline{F}(1-q)$ and calculate the limit:

$$\lim_{n\to\infty} \frac{a_{n+1} - a_n}{b_{n+1} - b_n}. \qquad (13)$$

According to Stolz-Césaro Theorem, if the limit above exists, then $\frac{a_n}{b_n}$ converges and its limit coincides with that in (13).

First, we calculate the difference $a_{n+1} - a_n$. It follows from the relations in Theorem 2 that:

$$a_{n+1} - a_n = \sum_{j=1}^n j \mathbb{P}[B_n = j] F(1-q+(n-1-2j)\varepsilon) + \sum_{j=1}^n j \mathbb{P}[B_n = j-1] \overline{F}\big(1-q+\big(n-1-2(j-1)\big)\varepsilon\big)$$
$$+ (n+1)\mathbb{P}[B_{n+1} = n+1] - \sum_{j=1}^n j \mathbb{P}[B_n = j].$$

By adding and subtracting 1 in the second term above and changing the index $j-1$ to j, we obtain:

$$a_{n+1} - a_n = \sum_{j=1}^{n} j\mathbb{P}[B_n = j]F(1 - q + (n - 1 - 2j)\varepsilon)$$

$$+ \sum_{j=1}^{n-1} j\mathbb{P}[B_n = j]\overline{F}(1 - q + (n - 1 - 2j)\varepsilon)$$

$$+ \sum_{j=1}^{n} P[B_n = j - 1]\overline{F}\big(1 - q + (n - 1 - 2(j - 1))\varepsilon\big)$$

$$+ (n + 1)P[B_{n+1} = n + 1] - \sum_{j=1}^{n} jP[B_n = j]$$

$$= -nP[B_n = n]\overline{F}(1 - q + (n - 1 - 2n)\varepsilon)$$

$$+ \sum_{j=0}^{n-1} P[B_n = j]\overline{F}(1 - q + (n - 1 - 2j)\varepsilon)$$

$$+ (n + 1)P[B_n = n]\overline{F}(1 - q - n\varepsilon)$$

$$= -(n + 1)P[B_n = n]\overline{F}(1 - q - (n + 1)\varepsilon)$$

$$+ \sum_{j=0}^{n} P[B_n = j]\overline{F}(1 - q + (n - 1 - 2j)\varepsilon)$$

$$+ (n + 1)P[B_n = n]\overline{F}(1 - q - n\varepsilon)$$

$$= (n + 1)P[B_n = n]\left(\overline{F}(1 - q - n\varepsilon) - \overline{F}(1 - q - (n + 1)\varepsilon)\right)$$

$$+ \sum_{j=0}^{n} P[B_n = j]\overline{F}(1 - q + (n - 1 - 2j)\varepsilon). \quad (14)$$

By relation 2 in Theorem 2, $\mathbb{P}[B_n = n]$ tends to zero. Furthermore, using the hypothesis on ε and the density of F we obtain from L'Hôpital's rule that:

$$\lim_{n \to \infty} \frac{\overline{F}(1 - q - n\varepsilon) - \overline{F}(1 - q - (n + 1)\varepsilon)}{(n + 1)^{-1}}$$

$$= -\lim 1_{n \to \infty} (n + 1)^2 \varepsilon[f(1 - q - n\varepsilon) - f(1 - q - (n + 1)\varepsilon)]$$

$$= 0.$$

It follows that the first term in the right-hand side of (14) tends to zero. On the other hand, for $0 \le j \le n$ we have:

$$\overline{F}(1 - q + (n - 1)\varepsilon) \le \overline{F}(1 - q + (n - 1 - 2j)\varepsilon) \le \overline{F}(1 - q - (n + 1)\varepsilon).$$

Hence, using that $\sum_{j=0}^{n} \mathbb{P}[B_n = j] = 1$ we obtain:

$$\overline{F}(1 - q + (n - 1)\varepsilon) \le \sum_{j=0}^{n} P[B_n = j]\overline{F}(1 - q + (n - 1 - 2j)\varepsilon) \le \overline{F}(1 - q - (n + 1)\varepsilon),$$

Therefore:

$$\lim_{n\to\infty}(a_{n+1}-a_n)=\lim_{n\to\infty}\sum_{j=0}^{n}P[B_n=j]\overline{F}(1-q+(n-1-2j)\varepsilon)=\overline{F}(1-q).$$

The limit above and the equality $b_{n+1}-b_n=\overline{F}(1-q)$ yield:

$$\lim_{n\to\infty}\frac{a_{n+1}-a_n}{b_{n+1}-b_n}=1.$$

The result now follows by Stolz-Césaro's Theorem. ■

Proof of Theorem 5. First, we show that $\frac{E[B_n]}{n}\to\overline{F}(1-q)$. By Theorem 4 above, for an arbitrary $\varepsilon>0$ and sufficiently large n we have:

$$\left|\frac{E[B_n]}{n\overline{F}(1-q)}-1\right|<\varepsilon\Rightarrow\overline{F}(1-q)\left|\frac{E[B_n]}{n\overline{F}(1-q)}-1\right|<\overline{F}(1-q)\Rightarrow\left|\frac{E[B_n]}{n\overline{F}(1-q)}-\overline{F}(1-q)\right|<\varepsilon,$$

Where in the last implication we have used that $\overline{F}(1-q)<1$. Now under the assumption that $\overline{F}^{-1}$ exists and it is continuous we obtain:

$$1-\lim_{n\to\infty}\overline{F}^{-1}\left(\frac{E[B_n]}{n}\right)=1-\overline{F}^{-1}\left(\overline{F}(1-q)\right)=q.$$

■

Appendix C. The case of the uniform distribution

In this section we present some important quantities in the particular case when the students initial distribution for understanding is uniform in [0,1]. Throughout this section we assume that ε is such that $[1-q-(n-1)\varepsilon,1-q+(n-1)\varepsilon]\subseteq[0,1]$, which we name as Hypothesis 1.

Proposition C.1. Let F be the uniform distribution over $(0,1)$ and assume Hypothesis 1 holds. Then for $i\leq n$ we have:

$$\overline{F}_i(1-q-m\varepsilon)=(1+2\varepsilon)^{i-1}\left\{q-\tfrac{1}{2}\right\}+\tfrac{1}{2}+m\varepsilon.$$

Proof. Recall that $F_1=F$. The result holds for $i=1$, since:

$$\overline{F}_1(1-q-m\varepsilon)=q+m\varepsilon.$$

We proceed by induction, assuming that for an integer $j\geq1$ we have:

$$\overline{F}_j(1-q-m\varepsilon)=(1+2\varepsilon)^{j-1}\left\{q-\frac{1}{2}\right\}+\frac{1}{2}+m\varepsilon.$$

By the Law of Total Probability and induction hypothesis:

$$\overline{F}_{j+1}(1 - q - m\varepsilon) = \overline{F}_j(1 - q - (m + 1)\varepsilon)\overline{F}_j(1 - q) + \overline{F}_j(1 - q - (m - 1)\varepsilon)\left(1 - \overline{F}_j(1 - q)\right)$$

$$= \left[(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\} + \frac{1}{2} + (m + 1)\varepsilon\right]\left[(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\} + \frac{1}{2}\right]$$

$$+ \left[(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\} + \frac{1}{2} + (m - 1)\varepsilon\right]\left[\frac{1}{2} - (1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\}\right]$$

$$= \frac{1}{2}(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\} + \frac{1}{2}\left[\frac{1}{2} + (m + 1)\varepsilon\right] + \left[(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\}\right]\left[\frac{1}{2} + (m + 1)\varepsilon\right]$$

$$+ \frac{1}{2}(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\} + \frac{1}{2}\left[\frac{1}{2} + (m - 1)\varepsilon\right] - \left[(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\}\right]\left[\frac{1}{2} + (m - 1)\varepsilon\right]$$

$$= (1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\} + \frac{1}{2} + m\varepsilon + 2\varepsilon\left[(1 + 2\varepsilon)^{j-1}\left\{q - \frac{1}{2}\right\}\right] = (1 + 2\varepsilon)^{j}\left\{q - \frac{1}{2}\right\} + \frac{1}{2} + m\varepsilon.$$

Hence the result follows. ∎

Theorem C.1. Let F be the uniform distribution over $(0,1)$ and assume Hypothesis 1 holds. Then for $m \leq n$ we have:

$$p_m = \frac{1}{2} + (1 + 2\varepsilon)^{m-1}\left(q - \frac{1}{2}\right).$$

Proof. For $1 \leq j \leq m - 1$, it follows from Proposition C.1 that:

$$\overline{F}_j(1 - q - \varepsilon) - \overline{F}_j(1 - q + \varepsilon) = 2\varepsilon.$$

Since $\overline{F}(1 - q) = q$, by Theorem 1 we have:

$$p_m = \overline{F}_1(1 - q)\prod_{j=1}^{m-1}\left[\overline{F}_j(1 - q - \varepsilon) - \overline{F}_j(1 - q + \varepsilon)\right] + \sum_{i=1}^{m-1}\overline{F}_i(1 - q + \varepsilon)\prod_{j=i+1}^{m-1}\left[\overline{F}_j(1 - q - \varepsilon) - \overline{F}_j(1 - q + \varepsilon)\right]$$

$$= q[2\varepsilon]^{m-1} + \left\{\frac{1}{2} - \varepsilon\right\}\sum_{i=1}^{m-1}(2\varepsilon)^{m-1}(2\varepsilon)^{-i} + \left\{q - \frac{1}{2}\right\}\sum_{i=1}^{m-1}(1 + 2\varepsilon)^{i-1}[2\varepsilon]^{m-1-i}$$

$$= [2\varepsilon]^{m-1}\left[q + \left\{\frac{1}{2} - \varepsilon\right\}\sum_{i=1}^{m-1}(2\varepsilon)^{-i} + \frac{q - \frac{1}{2}}{1 + 2\varepsilon}\sum_{i=1}^{m-1}\left(\frac{1 + 2\varepsilon}{2\varepsilon}\right)^{i}\right]. \tag{15}$$

Using the identities:

$$\sum_{i=1}^{n-1}\left(\frac{1 + 2\varepsilon}{2\varepsilon}\right)^{i} = [1 + 2\varepsilon]\left[\left(\frac{1 + 2\varepsilon}{2\varepsilon}\right)^{n-1} - 1\right] \text{ and } \sum_{i=1}^{n-1}(2\varepsilon)^{-i} = \left[\frac{1}{1 - 2\varepsilon}\right]\left[\left(\frac{1}{2\varepsilon}\right)^{n-1} - 1\right],$$

The right hand of equation (15) becomes:

$$[2\varepsilon]^{m-1}\left[q + \left(\frac{1}{2} - \varepsilon\right)\left[\frac{1}{1 - 2\varepsilon}\right]\left[\left(\frac{1}{2\varepsilon}\right)^{m-1} - 1\right] + \frac{q - \frac{1}{2}}{1 + 2\varepsilon}[1 + 2\varepsilon]\left[\left(\frac{1 + 2\varepsilon}{2\varepsilon}\right)^{m-1} - 1\right]\right]$$

$$= [2\varepsilon]^{m-1}\left[\frac{1}{2}\left(\frac{1}{2\varepsilon}\right)^{m-1} + q\left(\frac{1 + 2\varepsilon}{2\varepsilon}\right)^{m-1} - \frac{1}{2}\left(\frac{1 + 2\varepsilon}{2\varepsilon}\right)^{m-1}\right]$$

$$= \left[\frac{1}{2} + q(1 + 2\varepsilon)^{m-1} - \frac{1}{2}(1 + 2\varepsilon)^{m-1}\right].$$

Hence, we obtain:

$$p_m = \frac{1}{2} + (1 + 2\varepsilon)^{m-1}\left(q - \frac{1}{2}\right),$$

And the result follows. ∎