HIGH-SCHOOL STUDENTS' PROBABILISTIC REASONING WHEN WORKING WITH RANDOM INTERVALS

RAZONAMIENTO PROBABILÍSTICO DE ESTUDIANTES DE BACHILLERATO CUANDO TRABAJAN CON INTERVALOS ALEATORIOS

M en C Sandra Areli Martínez Pérez
Centro de Investigación y Estudios Avanzados del IPN
sandra.martinezperez@cinvestav.mx

Dr. Ernesto A Sánchez Sánchez
Centro de Investigación y Estudios Avanzados del IPN
esanchez@cinvestav.mx

This work reports the results of a research aimed to know the probabilistic reasoning of high-school students when they deal with the notion of random intervals. An activity was carried out involving students between ages 16 and 17 who built random intervals through physical and computational simulations. The research question guiding this work was: Which reasoning do students exhibit when they estimate the probabilities of events related to the experience of creating random intervals from a frequentist approach? From the data analysis, partly based on the Grounded Theory, four categories were established. They suggest that the patterns observed in this work are likely present in situations demanding the frequentist approach to probability.

Keywords: Reasoning, Frequentist approach, probability, random intervals

Problem statement

The educational research on the probabilistic reasoning of students is a complex field since it involves an abundance of concepts, innovative instructional proposals, conceptions, misconceptions, and difficulties as well as a number of methodological approaches and conceptual frameworks (Jones et al., 2007; Chernoff & Sriraman, 2014). It has been recently stressed that educators and teachers must research and document the implementation of innovative approaches and materials in class to allow for a close integration of probability and statistics (Chernoff, Paparistodemou, Bakogianni, Petocs, 2016; Langrall, Makar, Nilson, Shaughnessy, 2017). It has been suggested to give probability teaching a modeling approach (Pfannkuch et al., 2016) starting from extra-mathematical situations or contexts of the natural or social reality. That is, modeling will allow students to acquire or create probabilistic concepts when solving problems emerged from real situations. For this supposition to be feasible, it is necessary to start with simple random situations that can be repeated under a set of well-defined conditions. Then, they will allow for the observation of patterns of outcomes. In our opinion, situations with coins, dice, urns, and roulettes can play a mediating role in the acquisition of probabilistic concepts and the fundamentals of modeling (Sharma, 2016). In this work, we assume the hypothesis that the situations based on random devices, with the aid of digital gadgets, can be mediators between abstract concepts and real situations. They can also be the support of fertile situations to contribute to the integrated learning of probability and statistics.

Traditionally, introductory high-school courses deal with probability and statistics separately. However, statistical analyses must include probabilistic reasoning since it allows to handle situations of uncertainty and variability intrinsic to the phenomena studied by statistics. Still, some approaches of probability teaching avoid developing reasoning on uncertainty and variability handling instead of promoting it and focus on more formal aspects (set theory), calculus (classic approach of probability), and technical elements (combinatorics). One way of including uncertainty and variability in probability classes is to organize situations and present problems that produce data following an unknown distribution. So, students have to analyze them to obtain conclusions, as in the estimation of the probability of events.

In: Sacristán, A.I., Cortés-Zavala, J.C. & Ruiz-Arias, P.M. (Eds.). (2020). *Mathematics Education Across Cultures: Proceedings of the 42nd Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Mexico. Cinvestav / AMIUTEM / PME-NA. https:/doi.org/10.51272/pmena.42.2020

Saldanha (2016) and Prodromou (2016) argue that the study of samples of a distribution obtained through a digital simulation provides an adequate resource to link probability to statistics, especially from a frequentist approach. For that reason, we consider there should be research on how students reason when facing problems that link probability to statistics and that the frequentist approach of probability must be further studied.

Research question

Given that probability is studied in preuniversitary levels, it is important to explore the possibilities that technology offers to study its link to inference in high school. Particularly, this work explores the idea of introducing the notion of random intervals (RI). The aim is to set a background that helps reasoning on confidence intervals (CI), and so a research question was formulated:

Which reasonings do students show when they estimate the probability of events related to the experience of creating random intervals through a frequentist approach?

Background

The work presented here calculates probabilities from the frequentist approach of probability. So, we review some research that includes the frequentist approach of probability with technology and particularly focus on the work by Ireland and Watson (2009). The work explores the understanding of elementary-school students (ages 10−12) regarding the connection between theoretical probability and experimental probability (frequentist approach of probability) after students work with manipulatives (coins, dice) and the software ThinkerPlots. Ireland and Watson propose a framework to interpret the students' understanding as a continuum from concrete (experimental) to abstract (theoretical) in which manipulatives, the simulator, and the Law of Large Numbers are especially important. The findings lead them to conclude that it is necessary to explicitly teach the connection between theoretical and experimental probability; it is not enough for students to observe the behavior of the outcomes from simulations to achieve such connection.

Another study on the frequentist approach was that by Stohl and Tarr (2002). The authors report an instructional sequence with the aim of assessing how technological tools allow for and limit the development of the notion of inference from probabilistic situations. The participants were 23 students in sixth grade (ages 11−12) who worked in pairs and used the computational tool Probability Explorer to formulate and evaluate inferences during a 12-day teaching period. Among the conclusions, the authors state that the tasks designed in the context of games of chance and urns (one in the context of fishing in a lake) allowed the students to perceive relationships between empirical and theoretical probability as well as the role of the sample size in such relationships.

Conceptual framework

For the aims of this work, we have chosen concepts referring to two dimensions: content and cognition. From the first, content related to random intervals is presented and the same is done with probabilistic reasoning from the second.

Mathematical content relevant to the study

The content of this study refers to the estimation of probabilities through the frequentist approach of the probability of events related to the experience of generating random intervals. Then, some concepts on the frequentist approach of probability should be reviewed.

A repetitive phenomenon is that which can be repeated under a set of given conditions such that every repetition of the phenomenon is considered equivalent to its predecessors. Particularly, a random experience E is a repetitive phenomenon in which a characteristic is observed to change from one repetition to another and cannot be predicted; still, the set of all potential outcomes can be determined (sample space). Consider a random experience E and its sample space S while an event A is a subset of S. If experience E is repeated N times and n is the number of times event A occurs, then the quotient $n/_N$ is the relative frequency of A. The relative frequency of an event depends on N such that it varies as the values of N change. The most important characteristic of the relative frequencies of an event is that they converge in a given number as the number of repetitions grows indefinitely.

A random interval is defined as an interval in which at least one of the end points is a random variable, and so it gives rise to a family of intervals. Since we cannot assume that the students master the topic of distributions, the task design was carried out based on the experiment of drawing a ball from an urn containing 10 numbered balls from 0 to 9. If the random variable is the number printed on the ball, its distribution is the discrete uniform distribution taking the values $x = 0,1,2,\ldots,9$ and a probability $p = \frac{1}{10}$. Students are asked to observe the events of the type "E_C = the interval I_x contains number C, where C is any number between 0 and 9." Table 1 shows the probabilities of events E_C. These events are not mutually exclusive; hence the sum of their probabilities is higher than 1.

Table 1. Probabilities that the event E_c occurs

Event	E_0	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9
Probability	$\frac{9}{45}$	$\frac{17}{45}$	$\frac{23}{45}$	$\frac{27}{45}$	$\frac{29}{45}$	$\frac{29}{45}$	$\frac{27}{45}$	$\frac{23}{45}$	$\frac{17}{45}$	$\frac{9}{45}$

Probabilistic reasoning

Batanero et al. (1996) define probabilistic reasoning as a mode of reasoning that refers to judgements and decision making under uncertainty; therefore, it is relevant to real life. This reasoning includes the ability to: identify random events in nature, technology, and society; analyze the conditions of such events and derive suppositions to make an adequate modeling; construct mathematical models and explore different scenarios and outcomes; and apply mathematical methods and procedures of probability and statistics.

Methodology

Participants

The participants of this research were 16 students between ages 16 and 17 from a public school in Mexico City who had never taken a formal course on probability. One of the researchers was the class teacher and conducted the activity.

Activity

The activity was created using the principles of design to support the students' statistical reasoning proposed by Cobb and McClain (2004). Those principles must consider five aspects: 1) central statistical ideas (in this work, we focus on population, sample, random intervals, relative frequency, probability, and law of large numbers), 2) the instructional activity (the questions were formulated to observe whether the students created a conception of the central ideas), 3) the classroom activity structure (it must start by pointing out relevant aspects, variables to consider and how they will be mediated, the topic to cover, the activity development, and the students' discussion on the data obtained), 4) computer tools used by the students (in this work, an applet was made using Fathom), and 5) the classroom discourse (it refers to the language used, which should cover the possible judgements that students make on the central ideas).

For the development of the activity, the following situation was presented to the students:

Two balls are drawn from an urn containing 10 balls numbered from 0 to 9. Consider the interval formed by the integer values found between the minimum and maximum of the values drawn (considering the end points). Which is the probability that the value contain number 8?

*Containing number 8 means that 8 is between the minimum and maximum values or it is one of them.

Technology

Biehler (2013) states that, if it is to be considered in statistics teaching, a digital tool should allow students to perform several actions as: quickly dragging and dropping variables in a graph to visualize distributions and relationships between variables; visualizing in real time how data and parameters change dynamically, affecting measures and related representations; and linking multiple data representations to informally observe statistical trends. Fathom allows for these actions and more.

The applet provided to the students (Figure 2) works as follows:

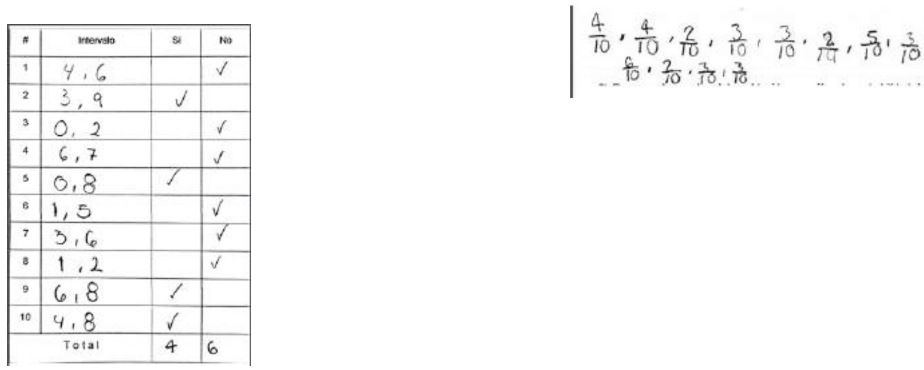
In an urn called *collection 1*, 10 balls numbered from 0 to 9 are placed. In *Sample of Collection* 1 is a sample of size 2, symbolizing the two values drawn from the urn with which the interval *[min, max]* is created. A measure, *belongs to*, is defined and consists in using the function *if* (*Si*) when number 8 is in the interval or *No* in the opposite case. A collection of measures in different sizes is taken and the tool *Summary* shows the number of *If* and *No* in a certain number of repetitions of the event. The plot shows the behavior of the relative frequencies.

Results and data analysis

For the analysis of the students' responses, we sought words or ideas that were common and placed them in codes. This response grouping process is proposed in the Grounded Theory by Birks and Mills (2015).

For instance, when they created intervals in the applet, students were asked what they observed in the plot. Some of the responses were "the dots are too scattered," "the dots generated by the data are too separated." They were placed in the code *scattering* since the students described how the dots in the plot were placed using the words "scattered" and "separated." Another type of response was "The plot is more constant when we placed more intervals." In these responses, the student observed that the dots in the plot (relative frequencies) were close to a constant value when many intervals were generated. These responses were placed in the code *Tendency to regularity*. Finally, in responses as "In the plot, the dots seem closer when we place more intervals and they are separated when there are fewer," we observed that students managed to see there was a difference when the number of intervals increased. These responses were placed in the code *Variability*. This analysis process was used to group the responses to all the questions in the activity.

The activity was divided in two parts; in the first one, the students carried out a physical simulation. To do so, they were given a bag with 10 balls numbered from 0 to 9. Each student obtained 10 intervals by drawing two balls without replacement. They wrote down the interval formed by the two values obtained $[x_{min}, x_{max}]$ on a table (Figure 1.a) and determined whether number 8 belonged to the interval.



#	Intervalo	Sí	No
1	4, 6		✓
2	3, 9	✓	
3	0, 2		✓
4	6, 7		✓
5	0, 8	✓	
6	1, 5		✓
7	3, 6		✓
8	1, 2		✓
9	6, 8	✓	
10	4, 8	✓	
Total		4	6

a b

Figure 1. a) Example of table filled in by students and b) group results.

Then, the students were asked to write down the relative frequencies of their classmates (Figure 1.b), and they were also asked about the probability that the interval contained the number 8. Some responded they had to obtain the mean or average (the best approximation), others expressed they had to pay attention to the one repeated the most (mode), and other students used the total relative frequency since they observed that 49 out of 160 intervals contained number 8. At the end of the physical simulation, the students were asked what would happen if they obtained 100 or 1 000 intervals to promote the use of the software. Three predictions were presented for this question: 1) proportionality (8/16)[1], which indicates that the number of favorable cases in 100 intervals must be proportional to 49/160; that is, approximately 30 in 100 intervals and 300 in 1 000; 2) approximation (3/16), where students propose a range of possible values around the proportional value of the frequency (values around 30 are suggested for 100 intervals while 300 are indicated for 1 000); and 3) the attention bias in favorable cases, in which only absolute frequencies, for instance, are said to "increase." The "approximation" responses are the most appropriate given that they take into account data on the relative frequency and predict a certain variability; attention bias is the response with the least quality.

In the second part of the activity, the students use an applet created in Fathom (Figure 2) by the authors and are explained how it works. Once they are familiarized with the software, the students are asked to create blocks of 5, 10, 20, … intervals, observe the corresponding plot, and write down their observations. Responses were classified in three codes, but it must be highlighted that the language used in their responses was mostly geometric and not probabilistic. They refer to the "dots" in the plot without providing any indication that they represent "relative frequencies:" 1) dispersion (5/16), they only say that the first dots in the plot are too separated; 2) a trend towards regularity (3/16), they only say that the "last" dots in the plot make a constant straight line; and 3) variability (4/16), they only deal with the separation between the dots at the beginning and their tendency to create a constant straight line at the end. The responses with the best quality are those classified in variability.

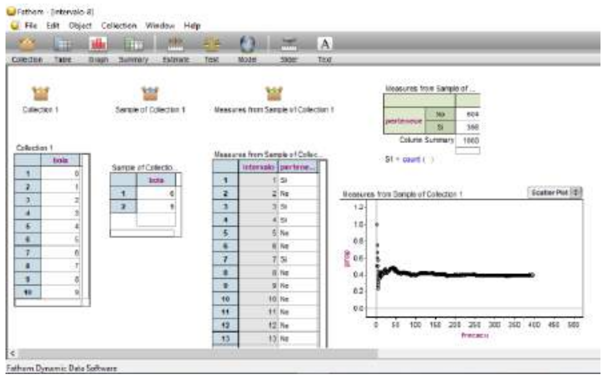


Figure 2. View of applet created by the students using Fathom.

Once the results of the applet were observed represented in the plot, the students were asked to tell the probability that an event contain the number 8. The responses to this question were classified in three codes: 1) relative frequency (8/16) when the students responded with the relative frequency that they obtained when they used the applet (considering 20, 100, and 1 000 tries); 2) approximation (6/16), when the students provided an interval or range within the relative frequency (as in "it is found around 0.306"). To know whether they were sensitive to the number of repetitions of the experiment (or number of intervals), the students were directly asked: What is the difference when there are a few and many intervals? The analysis of the responses led us to the codes above: 1) variability (7/16), 2) tendency to regularity (4/16), and 3) dispersion (1/16). Still, analyzing the responses according to the sense of their expressions, we proposed three additional codes: 1) geometric language (8/16) where they describe the behavior of the plot and not that of the relative frequencies, using terms as "dot," "straight line," or "constant" but not probabilistic terms; 2) variable probability (3/16), when they use the term probability in the same way as relative frequency, meaning that it changes as the number of experiments increases (for example "*The probability is more constant when the interval is greater*"); and 3) a priori probability (2/16), when, from their expression, we understand frequencies tend to a certain number ("*The more intervals there are, the better defined the constant is and also the probability we look for"* and "*The more dots there are, the closer we are to the probability*"). In these responses, students noticeably make a difference between relative frequencies and probability. Therefore, they are closer to the correct probability interpretation.

Finally, they are once again asked a prediction question: "If you had 1 000 intervals (without using the applet), how many of them would contain the number 8? Why?" The responses were classified in: 1) approximation (6/16), when students provide a range of 300 or say "*approximately around 300*;" 2) frequency (5/16), when they provide the frequencies obtained using the software; and 3) approximation to the value found using the software (4/16), when they say that they would obtain something similar to the result provided in the applet.

Discussion and conclusions

During the coding process, the synthesis of features present in several responses allow us to propose four categories that can provide a global notion of the progress and difficulties the students face in the process of conceptualizing not only the random intervals but also the probability in itself. They are:

Sensitivity to variability expressed in two ways. The first consists in accepting that, in prediction problems, it is impossible to accurately predict the number of favorable cases of an event in a series of repetitions of an experiment, but the relative frequency is known to be close to the probability. The second one consists in knowing that successive relative frequencies change greatly in a few repetitions of the experiment, while in the long run, they are stabilized around a constant. Our observations indicate that the activities using the software allow students to perceive variability. This is especially revealed when the simulation process is accompanied by the representation of the trajectory of the relative frequencies and also when, in prediction problems, several students do not provide exact values but intervals or ranges in which results can be found. This achievement is important since it is the basis to subsequently understand the law of large numbers.

No conservation of probability. It consists in believing that the probability of an event changes according to the implementations of the random experience. This phenomenon is similar to that of no conservation of the quantities described by Piaget (Gisnburg & Opper, 1988, p. 149). In the present case, the students do not use the term "relative frequency", preferring instead to use the term "probability." By doing so, they accept the notion that probability changes according to the number of experiments done. Although this is apparently a matter of terminology, it reflects the fact that the students are confused and do not conceive probability as a constant number related to an event.

Descriptive probability. It consists in believing that probability only offers information on the random experiences that have already occurred without reporting the future implementations of such experience. It is related to the previous category in that probability only describes a past state, and, when making new experiments, the state will change; therefore, probability will also change. Under this belief, there can be the misconception that probability does not allow for predictions. It is believed that the constant achieved when many experiments are repeated only occurs when a series of experiments are conducted. However, there is no assurance that the constant will be the same when carrying out other repetitions of the same random experience. This conception can be even more present in real random experiences different from games of chance; for example, social, medical, or weather problems.

Absence of a probabilistic language. It consists in describing a procedure or probabilistic result using non-probabilistic terms associated to a representation. In the students' responses to the question What is observed? Several students use geometric terms (dots, closeness/distance, constant, straight line) in the trajectories of relative frequencies without making any reference to relative frequencies and probability. The tendency to not use probabilistic terms to say what trajectories means raises the question of whether students understand the probabilistic meaning of trajectories of relative frequencies; that is, whether they interpret such representations as relative frequencies trending towards probability.

The four categories that emerged from the data of the exploration of the students' reasoning are more general to the particular situation of random intervals studied. They can also make sense in probability situations where there are problems to be solved through digital probabilistic simulations and a frequentist approach of probability. Indeed, in any situation where a computational simulation is used, it is suitable to consider and remember variability. In any preuniversitary teaching design, it should be considered that students can be at a stage where they do not accept the continuity of probability. They might also believe that probability only describes outcomes that have already occurred without any future consequences. Furthermore, care must be taken so that students interpret computational representations of probabilistic objects, as trajectories of relative frequencies, in probabilistic terms.

We also conclude that the use of technology was important because students managed to observe that relative frequencies in the plot generated by the applet converged in one value. Although they used geometric language, we think that better responses could be obtained by stressing what the constant and each point represent.

Finally, as a result of the analysis, we observed several ways in which the study can be improved to continue with another research cycle. Particularly, we have seen that the formulation of some questions should be improved, and more questions must be added to obtain further information on some aspects. For instance, it would be suitable to include an additional question to find to what extent do students who make geometric descriptions of the trajectories of relative frequencies understand the probabilistic background of the situation.

Acknowledgements

This work was funded by Project Conacyt 254301 and Project SEP-Cinvestav 188.

References

Batanero, C., Serrano, L., & Garfield, J. B. (1996). Heuristics and biases in secondary school students' reasoning about probability. In L. Puig & A. Gutiérrez (Eds.), Proceedings of the 20th conference of the International Group for the Psychology of Mathematics Education (Vol.2, pp. 51-58). University of Valencia, Spain.

Biehler, R., Ben Zvi, D., Bakker, A., Makar, A. (2013). Technology for Enhancing Statistical Reasoning at the School Level. En M. A. (Ken) Clements et al. (Eds.), Third International Handbook of Mathematics Education, (pp. 643-689) New York, NY: Springer. doi 10.1007/978-1-4614-4684-2_21

Birks, M., Mills, J. (2015). Grounded Theory. A practical guide. Los Angeles: SAGE Publications.

Chernoff, E. J., Sriraman, B. (Eds.) (2014). Probabilistic Thinking. Presenting plural perspectives. New York: Springer.

Chernoff, E. J., Paparistodemou, E., Bakogianni, D., Petocz, P. (2016). Research on learning and teaching probability within statistics. Statistics Education Research Journal 15(2), 6–10.

Cobb, P. & McClain, K. (2004). Principles of Instructional Design for Supporting the Development of Students Statistical Reasoning. En D. Ben-Zvi & J. Garfield (Eds.), The Challenge of Developing Statistical Literacy, Reasoning, and Thinking (pp. 375–396). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ginsburg, H. P., Opper, S. (1988). Piaget's Theory of Intellectual Developlment. Upper Saddle River NJ, USA: Prentice Hall, Inc.

Ireland, S., Watson, J. (2009). Building a connection between experimental and theoretical aspects of probability. International Electronic Journal of Mathematics Education, 4(3). 339-370.

Jones, G. A., Langrall, C. W., Mooney, E. S. (2007). Research in Probability. Responding to classroom realities. In F. K. Lester Jr. (Ed.), Second Hanbook of Research on Mathematics Teaching and Learning (909 – 955). Charlote, NJ: Information Age Publishing & National Council of Teachers of Mathematics.

Langrall, C. W., Makar, K., Nilsson, P., and Shaughnessy, J. M. (2017). Teaching and Learning Probability and Statistics: An Integrated Perspective. In J. Cai (Ed.), Compendium for Research in Mathematics Education (490–525). Reston, VA: National Council of Teachers of Mathematics.

Pfannkuch, M., Budgett, S., Fewster, R., Fitch, M., Pattenwise, S., Wild, C. (2016). Probability modeling and thinking: What can we learn from practice? Statistics Education Research Journal 15(2), 11-37.

Sharma, S. (2016). Probability from a socio-cultural perspective. Statistics Education Research Journal 15(2), 126-144.

Prodromu, T. (2016). Secondary school students' reasoning about conditional probability, samples, and sampling procedures. Statistics Education Research Journal 15(2), 106-125

Saldanha, L. (2016). Conceptual issues in quantifying unusualness and conceiving stochastic experiments: insights from students' experiences in designing sampling simulations. Statistics Education Research Journal 15(2), 81-105.

Sthol, H., y Tarr, J. E. (2002). Developing notions of inference using probability simulation tools. Journal of Mathematical Behavior, 21, 319–337.