INDEPENDENCE OF EVENTS: AN ANALYSIS OF KNOWLEDGE LEVEL IN DIFFERENT GROUPS OF STUDENTS

Verônica Yumi Kataoka[1], Hugo Mael Hernandez Trevethan[2] and Claudia Borim da Silva[3]
[1]Bandeirante University of São Paulo, Brazil
[2]Sciences and Humanities School, National Autonomous University of Mexico, Mexico
[3]University of São Judas Tadeu, Brazil
veronicayumi@terra.com.br

The objective of this work is discussing independent events, which have been the cause of theoretical confusions showed by students and mathematics in-service teachers. The proposed didactical activity was developed with 34 master students and 22 college students in Brazil, and 27 high school students in Mexico. The students, of different levels, didn't answer the problems effectively because they've used only the intuitive idea of independence, (chronological or informal independence). They were asked also to decide if two events were dependent or independent in a dice tossing situation, and the results showed something like "events are independent because there is only one kind of event: throw dice". The outcomes show that misconceptions of independent events and conditional probability persist even in people who had studied formally these concepts, pointing out the need to develop didactic situations to teach more effectively this content at high school and mainly to mathematics undergraduates.

INTRODUCTION

Curricular orientations for mathematics in high school in Brazil and in one Mexican subsystem suggest discussing random experiments, presenting the concept of frequentist probability, and, intuitively, working with conditional probability and independence of events.

The concept of independent events has caused much theoretical confusion among students and teachers. Conceptual errors generally happen because of the single use of common sense for giving an interpretation to independence of events (Nabbout & Maury, 2005; Cordani & Wechsler, 2006). Confusion of the word independence with exclusion may be an example of this, promoting difficulty on understanding two different probabilistic concepts, independence and incompatibility (Cordani & Wechsler, 2006; Tari & Diblasi, 2006).

Another misconception associated to the use of common sense is to consider just the definition of independence to chronological independent events, which, according to Steinbring (1986) is associated to the occurrence of successive experiments. As this author states, the other definition of independence is named as stochastically independent events. It's based on the mathematical formula $P(A \cap B) = P(A) * P(B)$ and its comprehension is restricted to the mathematical demonstration.

This mathematical formula for the independence of events comes from the expression for conditional probability, and for such a reason, a parallel study for both concepts becomes necessary. Furthermore, according to Diaz & Batanero (2009), the importance of constructing knowledge and conceptions related to conditional probability lies in the fact that conditional probability allows us to change our degree of confidence in random events when new information is available.

Even though different researchers have identified such misconceptions already, we think that it will always be an open research question: Could those errors be observed in students with any scholar level? From any country? With any Math background?

The aim of this project was to evaluate the level of comprehension of concepts about independent events among students of different scholar levels, different areas of knowledge and different mathematical backgrounds.

METHOD

Participants

The study was carried on with 27 high school students in Mexico City, Mexico (HS); 22 Math-major students at a state college in Bahia, Brazil (CS); 9 master-degree students in Genetics

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

students at a state college in Bahia, Brazil (G); and 25 master-degree students in Math-education students from a private university in São Paulo, Brazil (ME).

This study with students of different levels, interests and background is justified by the difference in the intensity on the use of that knowledge in the future. High school Mexican students need minimum knowledge, because they must "acquire introductory and propedeutic knowledge on the study of probabilistic and statistical methods, and applications in different areas of knowledge as well" (Dirección General del Colegio de Ciencias y Humanidades, 2004). Students in Mathematics, at the college level, will use this knowledge as a part of their citizenship and mainly by teaching it, so they need a deeper understanding of this knowledge. An even higher level is expected from the Math-Education-master students because most of them are already practicing in Math teachers in service, teaching Probability. For the master-degree students in Genetics, we've considered an important specific characteristic: These students need to know Statistics and Probability. They even have worked with statistics at that level by uploading some data in a computer and getting some feedback, probably meaningless for them, as a part of their area of study. So, it is important for these students to know certain basic concepts in Probability.

Procedures and data analysis

An activity was introduced collectively at the classroom, as a written assessment, without any previous teaching session. It concerned concepts related to independence of events, because it was assumed that students had some previous knowledge from their previous studies (elementary school and/or college), with exception of the high school students. The full instrument has five different sections with 27 questions. In those, topics as cross tabulations, tree diagrams, different kind of events and their probabilities, total probability and Bayes theorem, chronological independence, etc. In this paper we are presenting just the outcomes of six problems concerning independence of events, with two questions related with probabilities calculations presented by Kataoka et al (2008), and four about the concept of independence (Table 1).

Table 1. Presented problems in the instrument

Problems
P.1 Explain in your own words what do you understand as independence of events.
P.2 Consider the following events on tossing a die: Event A – to get a number greater than 3 (number on the face > 3) and Event B – to get a pair. Given that it appeared a number greater than 3, which is the probability of it being an even number?
P.3 In this situation, are the events A and B dependent or independent?
P.4 Consider the following events on tossing a die: Event A – to get a number lesser than 3 (number on the face < 3) and Event B – to get a pair. Given that it appeared a number lesser than 3, which is the probability of it being an even number?
P.5 In this situation, are the events A and B dependent or independent?
P.6. If two events are mutually exclusive, does it mean that those events are also independent? Justify your answer or give a counterexample.

The students' answers for each question were categorized in different hierarchic levels of knowledge, varying from a non-existing knowledge about the question up to a full knowledge level demanded to solve the question, according with Table 2. In order to organize our research, we examined some of the literature about misconceptions about Independence. We've developed our own categorization based on that previously developed by Díaz & Batanero (2009), whom classified a wrong answer as 0, partially wrong as 1 and totally correct as 2. We've also adapted that categorization to fit the needs raised from the nature of the responses in our study.

The students were allowed to use a calculator, but that wasn't a necessary condition. Access to references or some other sort of consulting was not allowed.

RESULTS

It can be observed by the results presented in Table 3 that for problem 1 (P.1) most of the students aimed to conceptualize independent events up to an informal level (code 2). With justifications as "the occurrence of an event doesn't change the occurrence of some other event", they did not consider the idea of probability.

Table 2. Description of categories used to systematize students' answers

Categories	0	1	2	3
P.1 and P.6	a) No answer b) Answers with inconsistent justification (IJ)	Answers with misconceptions (M)	Right answer with informal justification	Right answer with mathematically consistent justification
P.2 and P.4	a) No answer b) Other value without justification c) Other value with inconsistent justification (IJ)	Other value with misconceptions (M)	Right answer	
P.3 and P.5	a) No answer b) Wrong answer without justification or with inconsistent justification (IJ) c) Right answer without justification	a) Wrong answer with misconceptions (M) b) Right answer with inconsistent justification (IJ)	Right answer with informal justification	Right answer with mathematically consistent justification

Table 3. Percentage development of the students for the presented problems in the instrument

Problems	Categories															
	0				1				2				3			
	ME	G	CS	HS	ME	G	CS	HS	ME	G	CS	HS	ME	G	CS	HS
P.1	56	11	36	41	0	44	18	22	44	44	0	37	0	0	45	0
P.2	16	45	18	26	20	22	46	19	64	33	36	56	---	---	---	---
P.3	24	78	18	26	32	11	18	52	36	11	64	15	8	0	0	8
P.4	20	56	18	52	28	11	64	26	52	33	18	22	---	---	---	---
P.5	28	56	18	48	72	44	73	52	0	0	9	0	0	0	0	0
P.6	64	56	55	85	36	44	45	15	0	0	0	0	0	0	0	0
Total	35	50	27	46	31	30	44	31	33	20	12	22	1	0	17	1

In the results of problems 2 to 5, which refer to events stochastically independent, some misconceptions were identified. For example, in P.2 most of the students used information from the conditioning event just to identify the possible results on the conditioned event, but not to calculate the conditional probability. They used the total size of the original sampling space. In P.3 almost all the students just aimed to reach an informal level (code 2). In P.4 the same error of P.2 appeared again. Many students included number 3 into the determination of the conditioned sampling space, which reveals a problem of math reading instead of probability. In P.5, most of the students justified dependence of the events, using the chronological aspect. Some of these misconceptions were identified already by Sanchez (2000), Nabbout and Maury (2005), Tari and Diblasi (2006), and Cordani & Wechsler (2006).

Examples of answers to those four problems are presented in Table 4 for categories from 0 up to 2. For category 3, there were answers just for P.3, such as "They are dependent, because the event determines the probability we have for event B".

Same authors verified that some teachers don't aim to differentiate mutually exclusive events from independent events. This very same misconception was observed at the results obtained for P.6, because none of the students even aimed to present an informal definition for these concepts.

Results on Table 3 show that some high school students, even after calculating accurately the probability on P.2, didn't answer P.3 with a mathematically consistent justification, unlike the

Math-major students. There is a clear difference between the performance of master-in-Math-Education students and master-in-Genetics students in problems 2 to 5, prompting the discussion of the specific and background in both areas of knowledge. Generally, many of the answers were 0 and 1, revealing a lack of preparation on some groups of students.

Table 4. Examples of answers and misconceptions to problems 2 to 5, according with categories previously stated

Problem	Categories		
	0	1	2
P.2	The same probability of getting > 3 (IJ), 50/50.	From event > 3 to get an even number there are 2 values being 4, 6. That means 2/6 or 33% (M).	2/3 because the numbers on a dice greater than 3 are 4, 5 and 6.
P.3	They are independent because the tossing of 1 die gives a number randomly, and that number doesn't depend on the way the die is tossed (IJ).	They are independent, because it can come out any number randomly (M); dependent according to the value you get (M).	Dependent because if we affect A, B will be affected.
P.4	16%.	In the same way there is just the probability of being 2 = 1/6 = 16% (M).	1/2, there's just 2 and 1 left.
P.5	Dependent, because once again one alters the probability "exact" (IJ).	Dependent, to the value of the first event (IJ); they are independent because it is randomly (M).	Independent, because it hasn't the restriction lesser than 3, it will be no influence on the result.

CONCLUSIONS

The results show that, even in persons that studied formally topics related to independence of events and the calculation of their probabilities, misconceptions persist. There's the need to work more with at Math teacher formation courses, at the elementary-school level, explaining the different definition of independence of events and its formal definition in Probability Theory.

REFERENCES

Diaz, C., & Batanero, C. (2009). University students' knowledge and biases in conditional probability reasoning. *International Eletronic Journal Mathematics Education*, *4*, 131-162.

Dirección General del Colegio de Ciencias y Humanidades (2004). Programa Estadística y Probabilidad I y II. CCH-UNAM; México.

Cordani, L. K., & Wechsler, S. (2006). Teaching independence and exchangeability. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. International Association for Statistics Education, Salvador (Brazil). [CDROM].

Kataoka, V. Y., Souza, A. A., Oliveira, A. C. S., Fernandes, F. M. O., Paranaíba, P. F., & Oliveira, M. S. (2008). Probability Teaching in Brazilian Basic Education: Evaluation and Intervention. In *11th International Congress on Mathematical Education*, Monterrey.

Nabbout, M., & Maury, S. (2005). Teacher's representations of independent events: what might an attempt to make sense. In *4o Cerme*. Espanha.

Sanchez, E. (2000). Investigaciones didácticas sobre el concepto de eventos independientes en probabilidad. *Recherches en Didactique des mathématiques*, *20*(3), 305-330.

Steinbring, H. L. (1986). Indépendance stochastique : un exemple de renversement du contenu intuitif d'un concept et de sa définition mathématique formelle. *Recherches en Didactiques des mathématiques*, *7*(3), 5-50.

Tari, A. D'A., & Diblasi, A. (2006). Analysis of didactic suggested distinguishing disjunctive events and independent events. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics.* Intern. Ass. for Stat. Education, Salvador (Brazil). [CDROM].