A Practical Approach to Probability in the Context of a Science Fair

Hugo M. Hernández Trevethan
Maria del Refugio Muciño Reyes
Universidad Nacional Autónoma de México
animal_estocastico@yahoo.com.mx

Sánchez, Cuauhtémoc Vargas
Kataoka, Verônica Yumi
de Oliveira, Marcelo Silva
Universidade Federal de Lavras, Minas Gerais, Brasil

Summary

In a society that generates information in a fast way, schools have to manage fulfilling the programs in shorts terms. So, extracurricular activities may be helpful for the students to acquire wider knowledge than they may get within the classrooms. On the other hand, since randomness is present in almost all everyday decisions, mainly based on prior *information, it is important to have at least an idea on how certain events may affect the chances of other events to happen. We explore here both ideas in the context of a science fair, in which two high-school senior students conducted an investigation about conditional probability using a game named "Shut the box". We also want to state that after their participation in the science fair, these students have reached higher levels in probabilistic reasoning than usual, and they have acquired knowledge about concepts beyond the official curriculum.*

Introduction

Although the classroom is primarily the place where students acquire new knowledge, learning should not be confined to it. Actually, learning may be traced in many situations in everyday life. Moreover, there are good arguments to take these other occasions of learning into account. Also, it can be observed that there exists certain thinking which states that the knowledge that students acquire should be put into their daily life context, thereby having practical applications as well. Unfortunately, restrictions upon educational programs do not allow students to transfer the acquired knowledge to the desirable levels of contextualization and application: in this case, extra-curricular activities may get an important role in students' education.

One of these possible extra-curricular activities is a science fair, that according to Lima (retrieved from Brasil, 2006, 20), "is shown to be an opening for the curiosity and interest of the student, the creativity and mobilization of the teacher, and for the academic and social life of the school". Mills (2002) states that "researchers in education and psychology support the theory that students learn by actively building or constructing their own knowledge and making sense out of this knowledge", and that individuals may construct new knowledge internally by transforming, organizing, and reorganizing previous knowledge as well as externally, through environmental and social factors that are influenced by culture, language, and interactions with others (see e. g., Cobb, 1994; Greeno, Collins, and Resnick 1996; Bruning, Schraw, and Ronning 1999).

On the other hand, it is well known that randomness is found in almost every daily activity and consequently has a great impact on our life. Hence, the teaching of probability and the development of scientific research in this area are of major importance. Furthermore, both teaching of and research in probability will eventually demand contextualization and application.

In this paper, these ideas will be illustrated by analyzing an extra-curricular activity developed by two high school senior students which was presented at such a science fair. The activity involved research about some probabilistic concepts by adapting a game named "shut the box", which aimed to show the importance of measuring randomness, on the basis of prior information, in order to make a decision in a game, which could well be part of our everyday life's experience.

Science Fair: Students' Educational Process

More recent developments in schools have integrated a lot of extra-curricular activities. These programs try to support students' development through encouraging activities, which may improve their mental and physical health; other programs focus on the acquisition of social skills and skills, which are suitable for research and study.

One of these activities is the science fair, which, according to Ormastroni (1990, 7), "is a public exposition of scientific and cultural work developed by students. The students give demonstrations, provide oral explanations, and answer questions about the methods and their conditions. There is an exchange of knowledge and information between students and the visiting public". Therefore, a science fair is not just intended to support the student's education as an individual, but also to give students the opportunity to find a social framework on which to apply the knowledge that they have acquired in the classroom in by private studies.

The social aspects within scientific education that a science fair may offer to participating students is noted in the definition given in the document written by the Science Teachers in the Training Centre of Rio Grande do Sul, CECIRS (1970, 2): "It is a cultural activity carried out by students, shown through the demonstrations plan and present of their work, their knowledge and understanding in a technical-scientific area. This creates cultural development as well as a better cohesion between the school and the local community."

According to the ideas mentioned above, it is important not only that teachers encourage their students to participate in such a science fair but also evaluate their performance at the fair afterwards.

In this paper we are dealing particularly with the content of a research developed within a project at the science fair "Feria de las Ciencias del Museo de las Ciencias Universum", which is an already established event for high school students in Mexico. In this project, a game dealing with concepts of conditional probability and Bayes' theorem was designed.

The intention of the science fair is to encourage participating students to become familiar with scientific research by developing an investigative project and presenting their results to the public at the fair. The construction of suitable instruments for problem solving is also encouraged by the organizing board. In earlier years, projects at the science fair included research mainly under the topics of physics, chemistry, biology, health sciences, and related areas; more recently, however, this event has expanded to mathematics as well.

The participation process is organized in two stages: In the first phase the students have to present to the organization committee their research project plan, which should follow some pre-fixed guidelines. The project must be supported by a teacher. If the work is selected, then the students will present their project at the science fair. The second stage involves the development of an exposition and the defense of their work before a board of selected researchers specialized in different areas of science followed by presenting the project to the general public in an open session. At the end, the best projects in each discipline receive a prize.

This whole procedure involves the students in the real process of scientific research at a level, which is adequate to their age and knowledge. And if the teacher oversees the project development correctly, the student has a rich educational experience, enabling him or her to handle college-level work (where greater responsibility is placed on one's own learning) with better skills for acquiring knowledge. And this process also allows fulfilling the idea that learning occurs through hands-on interaction rather than through direct instruction in which the student passively receives knowledge (Cobb, 1994; Mills, 2003).

The two students participating in the project outlined above had the same level of understanding in mathematics as their classmates prior to undertaking the project, but they clearly showed a better understanding of statistics and probability as well as a greater interest for these subjects after the project. This is of particular interest if we consider that their whole group of the course of senior year of high school had encountered the subject for the first time, with the exception of some very basic concepts acquired earlier.

For their participation in the science fair, the teacher suggested the students, some activities and lines of research; the students then decided to work with these suggestions on the development of their project. Even more, as stated by von Glasersfeld (retrieved from Mills 2002) "Regardless of how clearly a teacher explains a concept, students will understand the material only after they have constructed their own meaning for the new concepts, which may require restructuring and reorganizing new knowledge and linking it to prior or previous knowledge". This does not mean that following these ideas guarantees complete and immediate easy to reach learning result; the students faced a lot of difficulties, the main of them being the calculations for the conditional probability tables, the idea of the use of prior information. and the development of Bayes' theorem.

Conditional Probability and Bayes' Theorem: Theoretical Aspects

During the science fair, the students presented to the public the formulas and concepts in three stages: *Prior* knowledge, conditional probability, and Bayes' theorem.

In every stage the students used the following method: initially they presented some situations and/or questions, with the idea of allowing the public to develop some probabilistic reasoning, and then they explained their formal theories.

First Stage: Prior Knowledge

The students presented to the public the following situation:

> We know, from experience, that every morning we must leave home at a certain time in order to get to school or to our job on time; and we know as well that, eventually, something unusual may happen and cause a delay on our way. Then we can ask the following questions:
>
> 'How can randomness influence our lives? What decision would we make if we knew that there was a construction work on our way during our trip to school or job? Or if we hear on the radio before leaving home that a car accident has just occurred precisely at some point of our usual route?'

This stage of methodological orientation was intended to show how *prior* information or knowledge may help someone in making an affirmation, or estimate a value, or guess the probability of an event occurring, or even further, how an unpredictable random situation can modify our decision-making in order to take an alternate strategy leading an advantageous outcome. This means that many times it is necessary to consider a series of situations that may change the probability of an event in which we are interested.

They completed this stage by pointing out that:

> "In our example, the fact that a car accident happens on our route will make the probability of traffic jams more likely, or will reduce our probability of getting on time to our destination".

And then they introduced the next question:

> "Is it possible to check out how specific events modify the probability of some other events happening? That is, how can we asses if two events can be named dependent or independent of each other?"

Second Stage: Conditional Probability

To discuss the concepts of dependent and independent events, and the concept of conditional probability, a two-way table was presented, in which the events A and B, from the same sample space, were considered along with their complements, denoted as $\overline{A}$ and $\overline{B}$:

Table 1. Cardinality of the events A and B from the same sampling space.

	A	$\overline{A}$	sum
B	$n(A \cap B)$	$n(\overline{A} \cap B)$	$n(B)$
$\overline{B}$	$n(A \cap \overline{B})$	$n(\overline{A} \cap \overline{B})$	$n(\overline{B})$
sum	$n(A)$	$n(\overline{A})$	N

Hereby, N denoted the number of elements in our sampling space, $n(A)$ the number of elements in A, and so on. After the presentation of the details in table 1, the students posed new questions to the public:

> "Based just on this information, is it possible to calculate the probability of the event A to happen, for instance? In this context, in which cases could events A and B be considered as dependent or independent?"

After a short time of analysis, the students showed that table 1 may be used to calculate the probabilities; for example, $P(A) = n(A)/N$; In the same way, the other probabilities are calculated. To answer the second question (independence of events), the students suggested that it was necessary to introduce the concept of conditional probability first, which was made in the following way:

1. How does it affect the probability of A if you know that B had already (fictionally) happened? If B occurs, it reduces our sampling space just to event B, which has $n(B)$ elements.
2. Therefore, the probability we are looking for has changed to $n(A \cap B)/n(A)$, since we take now only those cases for A, which lie also in B. We denote this conditional probability of A given B as $P(A|B)$.

If event A were independent from B, then $P(A|B)$ will be the same as $P(A)$. In this case, it holds $P(A \cap B) = P(A) \cdot P(B)$, which signifies the relation of independence of events as being symmetric. If this rule does not apply for two events then they are "dependent".

The students then put another question to the public:

"If we considerer now not only two events, but three or more, what would happen with the formula for conditional probabilities?"

This way they started with the third stage, stating that the Bayes' theorem is an application of conditional probability and its formula can be deduced easily from the formula of conditional probability.

Third Stage: Bayes' Theorem

It was discussed with the public, in an informal manner, that Bayes' theorem is valid in many applications of the theory of probabilities. Nevertheless, there is a controversy about the nature of probabilities that the theorem uses. Essentially, adherents of traditional statistics, the Frequentists, admit probabilities only if they are based on experiments which can be repeated and may thus be empirically confirmed, while those named as Bayesian statisticians allow also using subjective probabilities. The theorem is particularly useful to indicate how we should modify our subjective probabilities when we receive additional information for an experiment, although it also can be used when no "prior" information is available. Bayesian Statistics proves useful for estimating unknown probabilities based on subjective prior knowledge; this approach enables ways to construct knowledge where the traditional approach fails – however, at the cost of using subjective information as the Frequentists would object.

By explaining all this to the audience, the students' purpose was to establish the link between the ideas brought out in the first stage about prior knowledge and a new formula, Bayes' theorem.

The next step for the students was to present the whole development of the formula based on the previous discussion of the concepts. After that conceptual presentation, the students began to explore the game with the public, making comparisons between theory and practice, thereby completing the methodological cycle of the project.

The Game

Introduction

The game "Shut the box" consists of a series of small boards, labeled from one to nine, which lie on an axis inside a box, in such a way that each individual board may work like the box's cover (Figure 1).

Figure 1. Game Shut the Box, used as "Bayesian Box".

The game begins with the boards up, or opened, and the goal is closing them all. In order to do this, two six-sided dice must be thrown. Which boards can be closed is decided by the numbers shown either by the single dice, or by the sum of the two numbers on each die. For example, if we get 6 and 3 we could close the board labeled as 9 (if there still is a board labeled by 9 as the sum of the dice yields 9) or the boards labeled 6 and 3 (the single numbers of the dice shown).

These rules give the player a problem of decision-making right from the start. The player should formulate a strategy based on the convenience of closing the boards by sums or by the two single values observed, as long as this alternative ”may be applied” An example of outcome in which the player does not have the chance to decide would be a six-four result, since there is no board labeled 10 and the player must close the four and the six, or a tossing in which the sum presents a number on a board already closed.

An analysis of the probabilities of getting a result between two and nine inclusive, by sums, and the probabilities of getting a certain combination of values on the dice, will allow the player to plan, or at least attempt a strategy.

A further variant in the game was included in this research. By playing with two six-faced dice, a total of twelve faces may be counted. Under that idea, before starting playing, the player is given the option of playing either with the two six-faced dice, or with a die with eight faces and another with four faces, or with one ten-faced and one two-faced die, or, the last option, with a single twelve-faced die. All these options confront the player with an additional decision to choose a combination of the dice first with the target to increase the chances of winning. Also, depending on the selection of dice, the final strategy for the game could be different: the player has to decide on the combination of dice and then on how to play with them, by points or by sums in each throw, even when the player may switch this strategy from toss to toss.

Development

With the purpose of helping the public on the reasoning, some tables with the probabilities for the outcomes in the different combinations of dice, as well as the probability distributions for each die, were presented. Tables 2 and 3 are presented here as an example for the two six-faced dice:

Table 2. Probabilities for the outcomes with two six-faced dice.

	1	**2**	**3**	**4**	**5**	**6**
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

Table 3. Probability distributions for each face in a six-faced die.

Face	**1**	**2**	**3**	**4**	**5**	**6**
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

With the distributions constructed, the students presented the calculations for the probabilities of each result in tossing one or several dice, directly or by sums, as well as the calculations for the conditional probabilities of the sums given that the value on one of the dice is known, with the aim of observing the viability of the application of Bayes’ theorem. Table 4 shows the probabilities for the sums, while Table 5 shows the conditional probabilities In both cases, two six-faced dice are used.

It was pointed out as well that in the ideal world playing a series of games with different combinations of dice, with the aim of checking whether the best strategy gained from experience was the same as the best derived from the calculation of probabilities. At the beginning it was confusing for the public to understand the idea of the activity, since the initial explanation given to them was not bound to the context. As they started with the game, the most common errors were:

- If the player lost two or three games at the very beginning, he or she concluded that the game was impossible to win.
- The player assumed that any combinations of dice, excluding the single twelve-faced die, offers the same chances of winning.
- At the beginning the player did not find a difference between closing the boards by points or by sums in terms of probabilities. Only after a few throws they noticed that this element could make the difference between winning and losing the game.

Table 4. Probability of different sums of two six-faced dice

Sum X	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
P(X)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Table 5. Conditional probabilities for the sums

Sum	**Conditioning value of "second" die**					
	1	**2**	**3**	**4**	**5**	**6**
2	$\frac{1}{11}$	—	—	—	—	—
3	$\frac{2}{11}$	$\frac{2}{11}$	—	—	—	—
4	$\frac{2}{11}$	$\frac{1}{11}$	$\frac{2}{11}$	—	—	—
5	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$	—	—
6	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{1}{11}$	$\frac{2}{11}$	$\frac{2}{11}$	—
7	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$
8	—	2/11	2/11	$\frac{1}{11}$	$\frac{2}{11}$	$\frac{2}{11}$
9	—	—	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$	$\frac{2}{11}$
10	—	—	—	$\frac{2}{11}$	$\frac{1}{11}$	$\frac{2}{11}$
11	—	—	—	—	$\frac{2}{11}$	$\frac{2}{11}$
12	—	—	—	—	—	$\frac{1}{11}$
Total	**1**	**1**	**1**	**1**	**1**	**1**

After direct experience with the game, and once again with an explanation from the students, it was clearer to the public how to make use of conditional probabilities, and that probabilities can play a decisive role for making decisions, and the importance of prior information to improve one's decisions.

Results, Analysis and Interpretation

With the analysis of the tables, it was possible for the students and the public to draw certain conclusions through discussion. Some examples of this are:

1. In the "Total" line of each of the tables of conditional probabilities for the sums, it can be noticed that $\sum_{i=1}^{k} P(A_i \mid B) = 1$.

2. If we observe the tables for the combinations of each possibility of the pair of dice, and we calculate the probability of the possible sums, particularly those in which we could be interested for the game, we find that there is no coincidence since they were calculated as non-conditioned results. But, as we have seen, the sum can be conditioned to the result observed in one of the dice. This can be taken in consideration as a part of the game strategy.

3. It can be noted that the combination of a ten-faced die with a two-faced one allows getting values from two to nine both by single points and by sums. In the combination of the eight-faced die with the four-faced one, just the values from two to eight have that property. And in the combination of the two six-faced dice, only the numbers from two to six allows one to do so.

4. We can also observe in the combination of a ten-faced die with a two-faced one that certain sums have very high probabilities, due to the conditioning event, but there are only few alternatives. In the combination of the two six-faced dice matters are in the opposite way: there are very many ways of getting different sums from a conditional event, but that reduces the probabilities for each one of these sums. And in the combination of the eight-faced die with the four-faced one it seems to be a certain balance between both former cases. This leads us to assume that the best strategy is playing with this pair of dice, and by finishing first the values obtained by sums, particularly for values of five and higher.

Final Considerations

It was pointed out that the sole work in the classroom may limit wider knowledge. Both participants in the science fair were able to acquire more advanced knowledge in statistics and probability compared to their classmates, which indicates a certain advantage of encouraging students to participate in extra-curricular activities. This could be measured not just because of the way they tackled the problems in the project or by the way they presented it to the researchers the project and the general public who assessed, but also because they were able to develop Bayes' theorem, which is beyond the scope of the syllabus included in the institutional program of the school these students attend. This means that they were not just able to learn about Bayes' theorem but they were also able to explain it to an audience.

Moreover, the activity gave the opportunity to the public to think about randomness affecting their lives and the importance of taking this concept into consideration when making decisions. The activity also encouraged further participation in fairs to come with projects in probability. The inclusion of students in these events allows them turning into active agents of their own learning processes, and developing a capacity of critical analysis and investigative skills which qualifies them better for their later profession but also as emancipated citizens. Finally, the teacher must be conscious that encouraging students in these activities demands a high level of pedagogical knowledge and know how alongside good skills from the subject matter, here from probability.

References

Brasil, Ministério da Educação. Secretaria da Educação Básica-SEB. Programa Nacional de apoio às Feiras de Ciências da Educação Básica FENACEB, Brasília, 2006. http://portal.mec.gov.br/

Bruning, R., Schraw, G., & Ronning, R.: 1999, *Cognitive psychology and instruction* (3rd ed.), Prentice Hall, Upper Saddle River, NJ.

CECIRS: 1970, Boletim Porto Alegre, 5, 1–20.

Cobb, P.: 1994, Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher* 23, 13–20.

Greeno, J., Collins, A., & Resnick, L.: 1996, Cognition and learning. In Berliner, D. C. & Calfee, R. C. (Eds.) *Handbook of educational psychology* (15–46), MacMillan, New York.

Lima, M. C. E.: 2004, Fair Science: the school production propagated and the desire to know the pupil. Recife: Space Science. Retrieved December 20, 2007, from http://www.espacociencia.pr.gov.br.

Mills, J. D.: 2002, Using computer simulation methods to teach statistics: A review of the literature. Journal of Statistics Education, 10 (1). http://www.amstat.org/publications/jse/v10n1/mills.html.

Mills, J. D.: 2003, A theoretical framework for teaching statistics. *Teaching Statistics* 25, 56–58.

Ormastroni, M. J. S.: 1990, Manual da Feira de Ciências, Brasília: CNPq, AED, 30.

von Glasersfeld, E.: 1981, An attentional model for the conceptual construction of units and numbers. *Journal for Research in Mathematics Education*, 12, 83–94.