12th International Congress on Mathematical Education
Program Name XX-YY-zz (pp. abcde-fghij)
8 July – 15 July, 2012, COEX, Seoul, Korea (This part is for LOC use only. Please do not change this part.)

TEACHING PROBABILITY IN SECONDARY SCHOOL

Paulo Cezar Pinto Carvalho

IMPA and FGV – Rio de Janeiro, Brazil

pcezar@impa.br

Probability teaching in secondary school many times emphasizes computing probabilities as ratios of favorable to possible cases. Often, however, not enough attention is given to whether the possible cases are equally likely. We argue that being able to identify the adequateness of an equiprobable model for a given situation is a fundamental ability to be developed in secondary school. The role of simulation in understanding the meaning of computed probability values and realizing the power and limitations of probabilistic methods is also discussed.
Probability, modeling, simulation, secondary school.

INTRODUCTION

In most countries, official curricula or recommendations from educational authorities include teaching notions of probability to students since early grades, but with special emphasis in secondary school (for students aged 15 or older). In Brazil, for instance, the National Curriculum Guidelines (Brasil, 1998) indicate, for students in Grade 5 and 6, “construction of sample spaces and indication of the possibility of success of an event by means of a ratio”. For students in secondary school, the standards are quite vague, indicating only that students should be able to “apply notions of probability and statistics to solve problems” (Brasil, 2000). In practice, teachers are guided by the content of textbooks and university admission exams. In the United States, the NCTM standards (NCTM, 2000) set more specific goals, such as “understand the concepts of sample space and probability distribution and construct sample spaces and distributions in simple cases”.

There are good reasons for the early introduction of notions of probability and for its presence in secondary school curricula and in admission exams. Almost all areas of application, nowadays, use probabilistic methods. Understanding the random nature of many processes is fundamental, not only for professional development in many areas, but also to make correct, informed decisions in many situations in daily life. Thus, it is a good idea to expose students to probability as early as possible and to revisit it often, in such a way that the students can achieve an adequate maturity level when they finish secondary school.

In many cases, however, teachers and even text books are not well prepared to fulfill the mission of developing probability thinking in the students. Among other reasons this is due to the fact that most teachers at all levels have usually poor training in probability. Many teachers leave college having taken at most one course in Probability and Statistics, usually emphasizing procedures (such as looking up values in tables of the normal distribution), instead of discussing models.

Secondary school text books, in turn, many times regard probability as an appendix of counting methods, which are used to count the number of favorable and possible cases that appear in the ratio that express a probability value (often without a discussion of whether taking a ratio of number of cases is appropriate for the case at hand).

The purpose of this lecture is to discuss what should be the emphasis to be given in the teaching of probability to students in secondary school, in order to achieve a deeper understanding of the subject and to promote readiness for more advanced studies at college level. The discussion is based on the personal experience of the author with the Brazilian educational system, but most issues should be similar for several other countries.

EQUIPROBABLE MODELS

Since early grades, students are encouraged to express probability of an event as the ratio of the number of favourable cases to the entire number of cases. In secondary school, this gives teachers an opportunity to pose problems where the difficulty lies in counting those cases, which requires applying counting techniques. The fact that this requires that all possibilities be equally likely is not duly stressed. For instance, I came across the following problem when reviewing a text book for 6th graders:

Example 1. Peter and four of his friends are going to run a race. What is the probability that Peter comes out as the winner?

Of course, the intended answer is 1/5, that represents the ratio between the number of outcomes where Peter wins (1) divided by the total number of outcomes (5). However, problems like these are useless and harmful: they contribute for an attitude of defining probabilities as ratios, without questioning whether the outcomes are equally likely. In this case, there is no reason to assume that an equally likely model is appropriate, since real people have different athletical abilities. One could, however, modify the original problem in such a way that using am equiprobable model is appropriate, as follows:

Example 1 (modified). Five friends are going to run a race. Five t-shirts of different colours (red, etc) will be randomly assigned to them. What is the probability that the boy with the red t-shirt is the winner?

Regardless of who the winner is, the probability that he wears a red t-shirt is 1/5, since we stated that the shirts were randomly assigned.

Equiprobable models are the basis for most of the work in probability in elementary and secondary school. Actually, equiprobable models are the only probabilistic models that can be built without statistics. This means that, if there are no given probability values in a problem, then the problem can only be solved if one manages to model the situation using an equiprobable model. Even when such a model is available, sometimes it is not apparent to students (or teachers), such as in the following example.

Example 2. Students in a class have organized a raffle. 15 students bought 1 ticket each; 10 students bought 2 tickets each and 5 students bought 3 tickets each. Which is the most likely: that the student who wins the raffle has bought 1, 2 or 3 tickets?

When I propose this situation, I usually have students choosing each of the responses. People who answer "3 tickets" are solving correctly the wrong problem: of course, a student that bought 3 tickets is more likely to win that those who bought 1 or 2 tickets. However, the set of students that bought 2 tickets is the most likely to contain the winner: together, they bought 20 tickets, whereas the people that 1 or 3 tickets bought a total of 15 each, as seen in Table 1.

Table 1: Number of tickets per group

Number of tickets bought	Number of students	Number of tickets in group
1	15	15
2	10	**20**
3	5	15
Total	30	50

The key to the correct solution is to identify the proper sample space, which is composed by the 50 tickets, each with the same probability of winning, and not by the 30 students (people who answer "1 ticket" are erroneously considering this sample space, where outcomes are not equally likely).

FINDING THE RIGHT EQUIPROBABLE MODEL

As mentioned before, in problems where no probability values are mentioned, solution can only be given by setting the problem in an equiprobable model. Many times, however, student erroneously take as equiprobable the first sample space they can think of.

Example 3. A unbiased coin is tossed three times. What is the probability that heads come up exactly twice?

Consider the following proposed solution. *If one observes the number of heads in the 3 tosses, there are 4 possibilities: 0, 1, 2 or 3 (that is, the sample space is* S = *{0, 1, 2, 3} and the number of possible cases is 4). Among them, only one (2) corresponds to the event "heads are observed twice" (that is, there is only one favourable event). Thus the probability that we observe exactly two heads is* ¼ .

My experience is that many students, when exposed to the reasoning above, do not see anything wrong and are ready to accept its argument: since there are 4 possible outcomes and we are interested in just one of them, the associated probability is ¼ . The problem of course, is that probabilities can be computed as simple ratios only when all outcomes are equally likely. That's not the case here. When asked by students "why not?", my answer is that the burden of arguing for equal probabilities lies with the person who is proposing such a model. Why should the outcomes be equally likely? In fact, it is easy to convince oneself of the inappropriateness of an equiprobable model for this situation, by considering a larger number of tosses: if one tosses a coin 10 times, most people will agree that observing 10 heads is a much rarer event than observing 5 heads.

It is important to stress to students that one can assume an equiprobable model only when there is some kind of symmetry, which renders all possible cases interchangeable. Usually, that occurs in experiments that use devices built exactly for this purpose (such as dice, cards, or identical balls in urns). Coins are an interesting example: although not specifically built for providing a perfect randomizer, they present almost perfect symmetry, which makes a equiprobable model adequate for tossing outcomes. Another important source of equiprobable models is gender in human births. Actually statistics show that more boys are born than girls; on the other hand, infant mortality is higher among boys. All things considered, the ratio of surviving infants is approximately 1:1, which justify using an equiprobable model when analysing offspring gender.

In order to solve Problem 1 appropriately, one must consider each individual toss, where heads and tails are equally likely. Of course, this remains true, regardless of the sequence of results obtained before (I like to stress to my students that coins do not have any memory of past results neither care about keeping them balanced). Therefore, in a sequence of tosses, all possible sequences of outcomes are equally likely. Thus, the appropriate sample space for Example 1 is S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}, in which all outcomes are equally likely. Observing two heads correspond to the subset {HHT, HTH, THH} of S, whose probability is 3/8.

Example 4. Two identical, fair dices are thrown simultaneously. What is the probability that the sum of the upturned faces is 6?

As before, listing the possible sum of the faces (2 from 12) does not lead to a correct solution, because these sums are not equiprobable (more combinations of faces lead to a sum equal do 7 than to a sum equal to 2). The correct equiprobable model is obtained by considering the sample space of the 36 possible pairs of results and the set {(2, 4), (3, 3), (4, 2)} of the favorable cases. Since all 36 cases are equally likely, the probability of getting sum equal to 6 is 3/36 = 1/12. Sometimes, however, a student will object to this solution and present an alternative one. *Since the dice are identical, we cannot distinguish between, say, (2, 4) and (4, 2). Thus, the correct sample space is formed by the 15 unordered pairs of distinct results, plus the 6 cases where the outcome is the same for both dice. Therefore, the correct probability is 2/21.* The solution is incorrect, of course, since the 21 possible results are not equally likely. The lesson to be learned is that, in some situations, producing a good model requires altering conditions, in such a way that chances are not altered. In this case, the fact that the dice are distinguishable or not is irrelevant, but it is much more convenient to consider them as distinguishable, for that leads to a equiprobable space of ordered pairs. In example 8, below, we will see another case where changing the way to view a situation leads to a much easier solution.

In the two previous examples, we used two essential modelling paradigms: *equiprobability* of the individual results and *independence* of the successive tosses or throws (the outcome of one toss or throw has no influence on other outcomes). Although independence is quite obvious in these cases, sometimes people get confused, partly due to a common misconception regarding the Law of Large Numbers, which states that the observed frequency of an event must approach its probability, as the number of independent realizations of the experiment goes to infinity. As discussed by many authors (see, for instance, Tversky & Kanehman (1974)), many people interpret this result as saying that an observed sequence of results with frequency below normal must be compensated, in the near future, by a larger than normal frequency. Students should learn that the Law of Large Numbers works by *swamping* and not by *compensation*; that is, any finite set of observations is irrelevant, in the long run. It is a good idea to test student understanding with examples such as the one below.

Example 5. If one tosses a coin ten times, which of the following sequence of outcomes is more likely: THHTTHTTH or TTTTTTTTTT?

When caught off-guard, many people will answer that the first one is (much) more likely. The argument in Example 3 shows that they (and all other 10-outcome sequences) are equally likely. But this requires, for some students, to go against their first intuition, which says that a sequence where occurrences of heads and tails compensate each other is more likely than a sequence with 10 straight tails.

EQUIPROBABILITY AND PATTERNS

In the previous example, there is another reason for which many people will believe that THHTTHTTH is more likely than TTTTTTTTTT: in the latter there is a very noticeable pattern, which gives the observer a hint that it is a rare event (which is true, since the associated probability is $1/2^{10} = 1/1024$). For the first sequence, there is no such hint: it appears to be a quite common sequence, where 5 heads and 5 tails alternate in a seemingly natural manner. In fact, when estimating the probability of the first sequence, people may estimate, instead, the probability of getting 5 heads and 5 tails, which is much larger than the probability of getting 10 straight tails (252/1024 compared to 1/1024).

So far, we have emphasized developing I students the ability of realizing that a given sample space **is not** equiprobable. But being able to identify equally probable outcomes, despite their appearance, is also important, as in the following example.

Example 6. In a lottery, 6 numbers from 1 to 60 are drawn. Which set of six numbers is more likely: 1-2-3-4-5-6 or 8-19-24-36-42-51?

Of course, the answer is that they are equally likely: both have probability equal to $1/\binom{60}{6} =$ 1/50,063,860. But many will believe that the second set of numbers is more likely to be drawn, since its rarity is hidden by the lack of noticeable patterns.

Being able to analyse situations as these is important to support informed decisions in life. In Brazil, for instance, there are individuals that offer, for a fee, a service that consists in generating sets of numbers (to be bet in a lottery such as the one described before) that avoid "unlikely" patterns (for instance, consecutive numbers). The knowledge that all results are equally likely equips the student to avoid paying for a service that has no value.

THE MEANING OF PROBABILITY VALUES

When teaching probability to secondary school students, it does not suffice to teach them how to solve problems, usually by identifying a proper equally probable model and by computing probabilities as ratios. It is very important that students understand the meaning of the computed numbers.

Probability theory is a late comer in mathematical history. As pointed out by many authors, such as Rubinstein (1996), it was developed to support decision making under uncertainty, initially for gambling but later for much more serious applications. It is important that students understand that using probabilities supports decision making under uncertainty but does not completely remove it completely. In order to do that, probability problems should be posed, as often as possible, as decision problems, as in the Monty Hall Problem.

Example 7. Suppose you're on a game show, and you're given the choice of three doors: behind one door is a car; behind the others, goats. You pick a door, say No. 1 [but the door is not opened], and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to switch to door No. 2?" Is it to your advantage to switch?

The answer is "yes". If one does not switch, one gets the prize if and only if the right door was chosen initially, which occurs with probability 1/3. Therefore, if one switches, the probability of winning is 2/3. The problem involves computing probabilities, but with the purpose of making a decision. It should be mentioned that the Monty Hall Problem became very famous after some heated discussion about the correctness of the solution above. Many people (including several with training in Probability and Statistics) modelled the situation incorrectly, and concluded that it is indifferent to switch doors or not. For a more detailed discussion, see Gilman (1992).

It is important that students realize that, in the previous example, although switching is a rational decision, based on comparing probabilities, success is not guaranteed in a single realization of the game. Also, students should realize that the fact that the probability of winning when switching doors is 2/3 does not mean that exactly two in every three attempts will be successful. A good way of achieving this understanding is by *simulating* the situation. The simulation can be done with computers, and many of them are available in the Internet, as in The Monty Hall page (2012).

One does not need computers, however, to simulate situations such as the one in Monty Hall problem. The simulation can be easily done by students, working in pairs, assuming the roles of the host and the contestant. A significant number of runs can be obtained by pooling the results.

Another situation where simulation can be used to help students understand (and gain deeper trust on) computed probability values is given in the following example.

Example 8. Two equally able players, Arthur and Bernard, are competing in a series of games. The first one to reach 10 victories will be declared the winner and receive a $120,000 prize. If the series has to be interrupted when Arthur has won 8 games and Bernard 7, how should the prize be divided?

Similar versions of this problem are among some of the first probability problems to be studied, by Paccioli, Pascal and Fermat among others (David, 1962). Many people, when confronted with this situation, propose splitting the prize according to the past record of the players in the series: since Arthur has won 8 games and Bernard 7, the prize should be divided into parts proportional to 8 and 7; that is, Albert gets $64,000 and Bernard, $56,000.

Probability thinking allows for a radically different way to look at the situation, focusing the *future* instead of the *past*: The prize is divided according to the *expected* outcome, that is, proportionally to the probabilities of victory of each player. One difficulty to compute such probabilities is that the number of games still to be played is variable: the series may end in additional 2, 3 or 4 games, depending on who wins next. This can be represented by the tree in Figure 2, which shows that the endings of the series that lead to Bernard's victory are BBB, ABBB, BABB and BBAB, where A and B represents Arthur's or Bernard's victory, respectively.

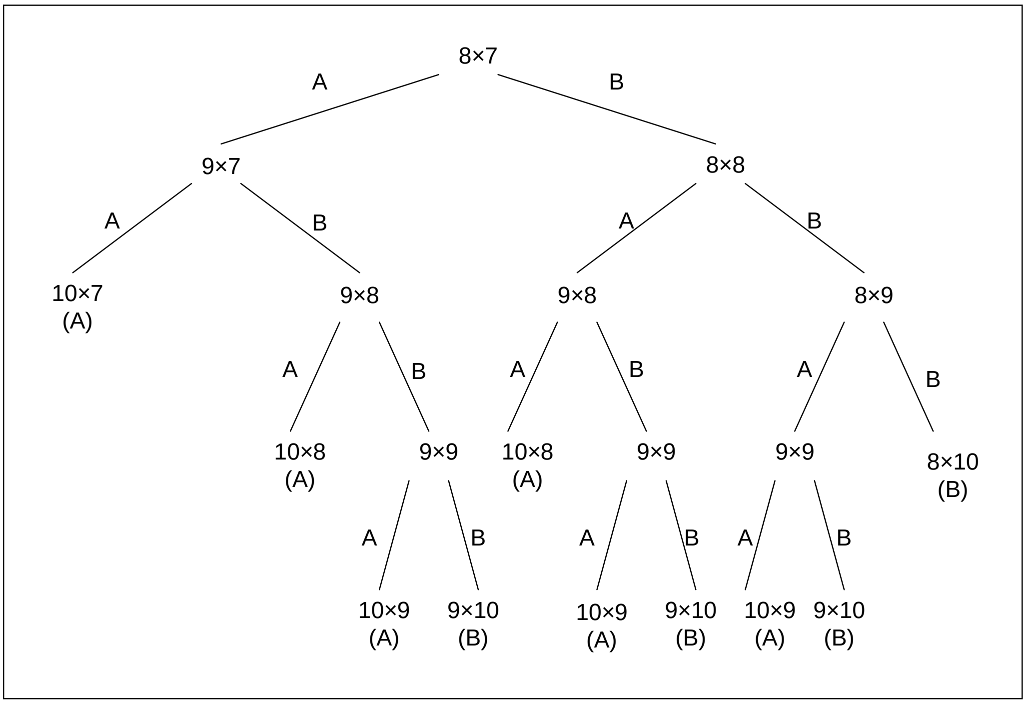

Figure 1. All possible endings for the series.

The fact that Arthur and Bernard are equally able translates into them having the same probability of victory in each game. We will also assume that the outcome in a game does not affect the performance of the players in other games (this is questionable when modelling real competitions, unless they are games of pure chance), so that any sequence of *n* outcomes has probability $1/2^n$. Thus the probability of Berrnard winning the series is $\frac{1}{8}+3\times\frac{1}{16}=\frac{5}{16}$, whereas Arthur's winning probability is $\frac{11}{16}$. Therefore, the prize should be split in parts proportional to 11 and 5, which leads to Arthur getting $82,500 and Bernard $37,500 (which is very different from the first splitting proposal).

An easier way to solve the problem consists in slightly changing the description of the problem. Instead of stopping the series when a player has reached 10 victories, we have them play, in all cases, 19 games (that correspond to the longest possible series) and declare the player who has won most games the winner. Of course, exactly one of the players will have reached 10 victories at the end of the series. Therefore, the winner is still the player to reach 10 victories first – nothing has changed in the final outcome. However, it is much easier to find the probability of final victory for each player. According to the revised rules, Arthur and Bernard have 4 more games to play, with $2^4 = 16$ equally likely possibilities. Bernard will be the winner if and only if Arthur wins not more than one game. There is only one sequence of results where Arthur does not win any game and four where he wins one game (the victory can occur in any of the four games). Thus, the number of favourable cases is 5 and the probability of Bernard winning game is 5/16 = 0.3125, as before.

The situation in Example 8 gives another excellent opportunity to use simulation to promote better understanding of the meaning of probability values. Table 2 shows the results of simulating 10000 runs of the end-game situation. The average earnings of Arthur in the 10,000 runs is $82,812, which is close (but not exactly equal to) the proposed splitting. Again, the simulation can be done in class, with students in pairs playing the roles of Arthur and Bernard and arriving for themselves at the conclusion that probability theory leads to the fair way to split the prize.

Table 2: Simulating the remaining games

Series	Game 1	Game 2	Game 3	Game 4	Outcome	A wins	B wins	A's prize	B's prize
1	B	B	B	--	**B wins**	**0**	**1**	0	120,000
2	B	A	A	--	**A wins**	**1**	**0**	120,000	0
3	A	A	--	--	**A wins**	**1**	**0**	120,000	0
4	A	A	--	--	**A wins**	**1**	**0**	120,000	0
5	B	A	B	A	**A wins**	**1**	**0**	120,000	0
6	B	A	A	--	**A wins**	**1**	**0**	120,000	0
7	B	B	B	--	**B wins**	**0**	**1**	0	120,000
8	B	B	A	A	**A wins**	**1**	**0**	120,000	0
9	B	A	B	A	**A wins**	**1**	**0**	120,000	0
10	B	B	B	--	**B wins**	**0**	**1**	0	120,000
...	**...**	**...**	**...**	**...**	**...**	**...**	**...**	**...**	**...**
9999	A	B	A	--	**A wins**	**1**	**0**	120,000	0
10000	B	B	A	A	**A wins**	**1**	**0**	120,000	0
					Averages	0.6901	0.3099	82,812	37,188

CONCLUSION

Probability teaching in secondary schools should not be restricted to using counting techniques to compute ratios of favourable to possible cases. It is important to thoroughly discuss when such ratios are appropriate to represent probabilities of events. Actually, this discussion does not depend on mastering counting techniques and can be done with younger students.

Probability computations should be accompanied by simulation practices in class, with students playing out the situations. This will help students to gain deeper understanding of the meaning of computed probability values and to realize both the power and the limitations of probabilistic methods.

References

Brasil (1998). *Parâmetros Curriculares Nacionais: Matemática*. Brasília: MEC.

Brasil (2000). *Parâmetros Curriculares Nacionais (Ensino Médio).* Brasília: MEC.

David, Florence Nightingale (1962). *Games, Gods and Gambling*. New York: Hafner.

Gillman, Leonard (1992). "The Car and the Goats," American Mathematical Monthly 99: 3–7.

NCTM (2000). *Principles and standards for school mathematics*. Reston: NTCM.

Rubinstein, Peter L. (1996). *Against the Gods: the remarkable story of risk*. New York: John Wiley.

The Monty Hall page (2012). Retrieved April 9, 2012, from http://math.ucsd.edu/~crypto/Monty/monty.html

Tversky, Amos; Daniel Kahneman (1974). "Judgment under uncertainty: Heuristics and biases." *Science* **185** (4157): 1124–1131.