Why Bayesians Need Not be Puzzled by Probability Puzzles

Wagenmakers, E.-J.

DOI
10.1080/09332480.2025.2510151

Publication date
2025

Document Version
Final published version

Published in
Chance

Why Bayesians Need Not be Puzzled by Probability Puzzles

Response submitted by Eric-Jan Wagenmakers

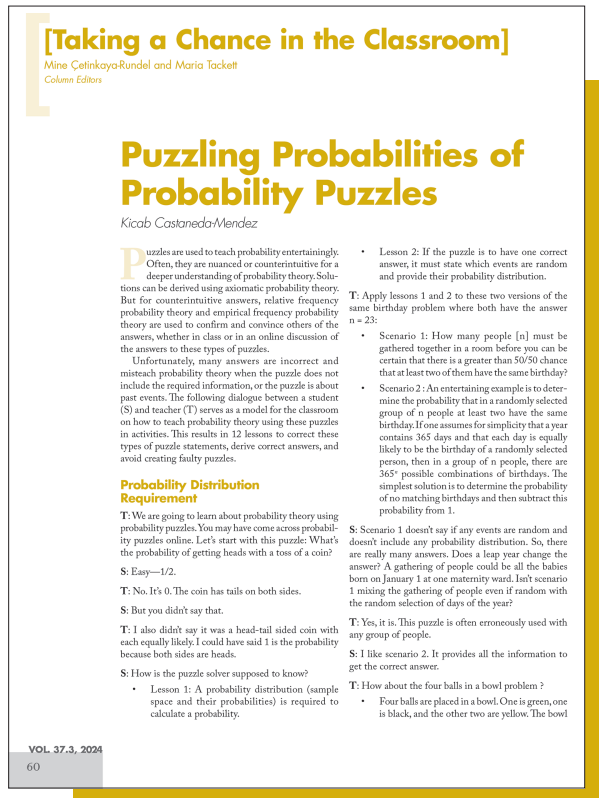
In "Taking a Chance in the Classroom: Puzzling Probabilities of Probability Puzzles" (Vol. 37.3 of *CHANCE*), Kicab Castaneda-Mendez presents a critical analysis of probability puzzles. This analysis hinges on two assumptions: (a) the only correct probabilities for unknown outcomes of past events are either 0 or 1; and (b) there is only a single correct probability. A conversation between a Teacher "T" and a Student "S" results in several lessons, including the following:

> "Lesson 5: Once an event has occurred and its outcome is unknown, do not use the probability distribution before it occurred to calculate probabilities of outcomes after it has occurred."
>
> and
>
> "Lesson 7: Past event sample spaces consist of only one case. The outcome is either favorable (probability 1) or unfavorable (probability 0)."

One of the examples from the paper concerns the probability of a gambler throwing snake eyes:

> "**T**: (…) Assume you, the reader, are the dealer in the snake eyes problem:
>
> - With eyes closed, you roll a pair of dice at a casino table. The dealer, seeing the result, tells you one of the dice is a one. What is the probability of snake eyes (two ones)? Puzzle answer: 1/11.
>
> **S**: As the dealer, I will see whether the outcome is snake-eyes or not. So, I will know which of 0 or 1 is the correct probability, not 1/11. Since there is only one correct probability, it is the probability for the roller too."

I wish to stress that these lessons and examples are diametrically opposed to the Bayesian interpretation of probability as a degree of coherent belief. As stated by Bruno de Finetti in *Theory of Probability, Vol. 1 and 2*:



[Taking a Chance in the Classroom]

Mine Çetinkaya-Rundel and Maria Tackett

Column Editors

Puzzling Probabilities of Probability Puzzles

Kicab Castaneda-Mendez

Puzzles are used to teach probability entertainingly. Often, they are nuanced or counterintuitive for a deeper understanding of probability theory. Solutions can be derived using axiomatic probability theory. But for counterintuitive answers, relative frequency probability theory and empirical frequency probability theory are used to confirm and convince others of the answers, whether in class or in an online discussion of the answers to these types of puzzles.

Unfortunately, many answers are incorrect and misteach probability theory when the puzzle does not include the required information, or the puzzle is about past events. The following dialogue between a student (S) and teacher (T) serves as a model for the classroom on how to teach probability theory using these puzzles in activities. This results in 12 lessons to correct these types of puzzle statements, derive correct answers, and avoid creating faulty puzzles.

Probability Distribution Requirement

T: We are going to learn about probability theory using probability puzzles. You may have come across probability puzzles online. Let's start with this puzzle: What's the probability of getting heads with a toss of a coin?

S: Easy—1/2.

T: No. It's 0. The coin has tails on both sides.

S: But you didn't say that.

T: I also didn't say it was a head-tail sided coin with each equally likely. I could have said 1 is the probability because both sides are heads.

S: How is the puzzle solver supposed to know?

- Lesson 1: A probability distribution (sample space and their probabilities) is required to calculate a probability.
- Lesson 2: If the puzzle is to have one correct answer, it must state which events are random and provide their probability distribution.

T: Apply lessons 1 and 2 to these two versions of the same birthday problem where both have the answer n = 23:

- Scenario 1: How many people [n] must be gathered together in a room before you can be certain that there is a greater than 50/50 chance that at least two of them have the same birthday?
- Scenario 2 : An entertaining example is to determine the probability that in a randomly selected group of n people at least two have the same birthday. If one assumes for simplicity that a year contains 365 days and that each day is equally likely to be the birthday of a randomly selected person, then in a group of n people, there are 365^n possible combinations of birthdays. The simplest solution is to determine the probability of no matching birthdays and then subtract this probability from 1.

S: Scenario 1 doesn't say if any events are random and doesn't include any probability distribution. So, there are really many answers. Does a leap year change the answer? A gathering of people could be all the babies born on January 1 at one maternity ward. Isn't scenario 1 mixing the gathering of people even if random with the random selection of days of the year?

T: Yes, it is. This puzzle is often erroneously used with any group of people.

S: I like scenario 2. It provides all the information to get the correct answer.

T: How about the four balls in a bowl problem ?

- Four balls are placed in a bowl. One is green, one is black, and the other two are yellow. The bowl

VOL. 37.3, 2024

60

> "The only relevant thing is uncertainty–the extent of our own knowledge and ignorance. The actual fact of whether or not the events considered are in some sense determined, or known by other people, and so on, is of no consequence." (xi)
>
> and
>
> "There is a prejudice that uncertainty and probability can only refer to future matters, since these are not 'determined'—in some metaphysical sense attributed to the facts themselves instead of to

> the ignorance of the person judging them. In this connection, it is useful to recall the following observation of E. Borel: 'One can bet on Heads or Tails while the coin, already tossed, is in the air, and its movement is completely determined; and one can also bet after the coin has fallen, with the sole proviso that one has not seen on which side it has come to rest.'" (p. 40)

Much earlier, this Bayesian viewpoint was also advocated by Augustus De Morgan. In the 1838 (!) book *An Essay on Probabilities and on Their Application to Life Contingencies and Insurance Offices*, De Morgan first points out that "Probability is the feeling of the mind, not the inherent property of a set of circumstances" (7) and later discusses uncertainty for unknown outcomes for past events:

> If I am before an urn which contains a black and a white ball, which is all I know, I am then disposed to say the chances are even. The ball which I am to draw is undetermined (by me), and that which we call chance appears to exist. But suppose I draw the ball, and without looking at it hold it in my hand. That which we call chance has ceased to exist–the ball is actually determined, and I am clearly and physically placed in the same position as I should have had before the drawing, if a superintending power, capable of guiding my thoughts and actions without my perceiving it, had predetermined which I should draw. *But my position with respect to knowledge of the ball is not in any way changed*, either by the predetermination of the superintending power before the drawing or by my own act of drawing, as long as I do not know what I am to draw or have drawn. It tells me nothing, if I hear that the drawing is settled, unless it be in a manner by which I can form some guess as to the nature of the settlement." (116; italics added for emphasis)

Consider then an example Bayesian analysis of the snake eye puzzle. Before rolling, the probability of snake eyes is 1/36, both for the roller and for the dealer. Immediately after rolling, the probability of snake eyes is very close to 0 or to 1 for the dealer (who may wish to take into account the unlikely possibility of a visual hallucination). For the roller, however, the probability is still 1/36. The dealer then informs the roller that one of the dice shows a "1." If the roller believes the strategy of the dealer is to issue this report whenever at least one of the dice shows a "1," then the roller's probability increases from 1/36 to 1/11. However, if the roller believes the strategy of the dealer is to report the outcome of the particular die that just happens to land closest to them, then the probability increases from 1/36 to 1/6. The roller ought to have a probability distribution over the dealer's reporting strategy, and averaging over this yields the final outcome. For instance, when the roller deems both of the above reporting strategies exhaustive and equally likely, the roller's probability of having just thrown snake eyes equals $[1/11 \times 1/2] + [1/6 \times 1/2] = 17/132 \approx 0.13$. Crucially, the roller and the dealer have dramatically different probabilities concerning an event that has already transpired—which is desirable and apt, at least in the Bayesian world, since the roller and dealer have access to very different information and hence have very different degrees of belief.

I conclude that a Bayesian interpretation of probability as degree of coherent belief salvages the probability puzzles that are critiqued by Castaneda-Mendez.

Further Reading

de Finetti, B. 1974. *Theory of Probability, Vol. 1 and 2*. New York: John Wiley & Sons.

De Morgan, A. 1838. *An essay on probabilities and on their application to life contingencies and insurance offices*. London: Longman.

Marks, S., Smith, G. 2011. The two-child paradox reborn? *CHANCE* 24:54–59.