*Journal of Probability and Statistical Science*
*24(2), 1-15Aug. 2026*

# Using Problems about Probabilities from Social Media for Teaching?

Hans Humenberger
Faculty of Mathematics
University of Vienna, Austria

## ABSTRACT

In this paper we highlight interesting stochastic problems of a young Canadian mathematician on a social media platform (X), especially their potential for being used for teaching at various levels. He managed to be featured in a reputable Austrian daily newspaper by a science journalist, and rightly so!

***Keywords*:** Teaching probability, Social Media

We became aware of the following problems through an article by the science journalist Klaus Taschwer in the Austrian daily newspaper "Der Standard", more precisely in the online version of it, see Taschwer (2024).
There can be read in the first lines (translated from German):

> Coin tosses, urns and balls
> **Mathematician amazes social media with tricky probability puzzles**
> Daniel Litt, a specialist in number theory and algebraic geometry, thinks up harmless-sounding problems that provide counterintuitive answers – and in doing so inspires experts.

That, actually, describes the situation quite well. Daniel Litt teaches and researches as a young assistant professor at the University of Toronto in Canada and apparently has a good sense of stochastic puzzles, even though his research field is not stochastics (who manages to get into renowned newspapers with mathematical puzzles?).
Let us start with a problem which there is called **problem 2**:

> Flip a coin until you get five consecutive tosses of the type HTTTH or HTHTH (where H is heads and T is tails). Then stop. These last five throws were more likely a) HTTTH b) HTHTH c) both alternatives are equally likely?

Really calculating the probabilities of one pattern coming earlier than the other is not that easy, but that's not the point (e.g. using Markoff chains, it would turn out that HTTTH comes earlier than HTHTH with a probability of 5/9 > 1/2). The answer c) "equal probability" can be triggered by the fact that of course every 5-part H-T pattern has the same probability of occurrence in 5 tosses, namely $\left(\frac{1}{2}\right)^5$, and it is always only about the last 5 coin tosses. However, if you look at these two patterns more closely, you can – and should – notice (and that is the main point here): If the first pattern "fails" just before its end, i.e. after HTTT there is not H, but T, then that is equally bad for both patterns, because in both cases you have to wait for the next H (as the first correct value of the respective pattern). However, if the second pattern "fails" just before its end, i.e. after HTHT comes not H, but T, then with the last three throws (HTT) the first three digits of the first pattern are already there, while the second pattern "is again at nothing" (it has to wait for the first H , so to speak for a new start), so the first pattern has a clear advantage in this case and overall. Quantifying this advantage is not part of the task.

We have put this puzzle at the beginning because it practically does not need any calculations, but it is not an easy task. If learners begin to calculate here or argue for equal probability (either from their gut or triggered by the above consideration), then the teacher can,

- Received May 12, 2026, in final form August 2026.
- Hans Humenberger (corresponding author) is affiliated with Faculty of Mathematics University of Vienna, Austria
  Email: hans.humenberger@univie.ac.at

without revealing the solution, give a hint of the following type: Take a closer look at the patterns, especially what happens when a pattern "fails" just before completion!

Now we present the puzzle – there called **problem 3** – for which one needs already a bit more calculations and diagrams (like a tree diagram):

> You have 100 urns, each containing 99 balls. In 99 of the urns, one ball is red, and the others are green. In the last urn, all 99 balls are red. You randomly choose an urn and draw a ball. It is red. You put the ball away. The next ball you draw from the same urn is a) more likely to be red b) more likely to be green c) equally likely red or green?

We denote the set of the first 99 identically equipped urns with $U_{1-99}$ and the last urn by $U_{100}$. We can already see in the tree diagram (Fig. 1) that the probability of the observed event "the ball drawn is red" is in both cases (left and right branch in the tree diagram) $\frac{1}{100}$ (note $\frac{99}{100}\cdot\frac{1}{99}=\frac{1}{100}=\frac{1}{100}\cdot 1$).

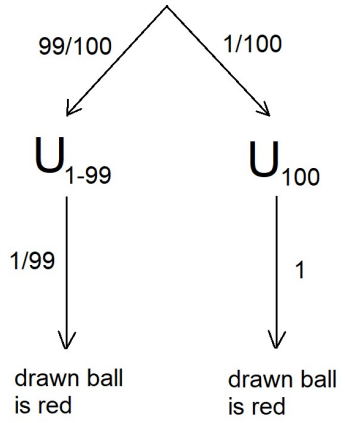



Figure 1: Tree diagram

The value 1/100 is not that important here, what is more important is the equality of the two total values on both branches.

Since in the left case the *next* ball is certainly green and in the right case the *next* ball is certainly red (equal probabilities again!), it is clear that answer c) is correct (equal probability). This task should be easy to do in any class in which the corresponding path rules for tree diagrams have been covered and should not be overwhelming.

Now we come to the formulation of the puzzle, there called **problem 1**, which we will discuss in more detail below (Fig. 2).




Figure 2: Problem 1 of Daniel Litt (with friendly permission from D. Litt)

In Fig. 2 you can also see that after 22,563 votes (end of January 2024), the relative majority chose "more likely to be green" (approx. 37 %). The other alternatives (more likely to be red, equally likely, don't know) were all about 20 % agreement.

First, we want to describe possible ways of thinking, initially without any calculations, that lead to a certain assessment (like this, also students or the people taking part in Daniel Litt's vote could think or argue):

1. Due to the one red ball that was taken from the urn, it is missing from the "equilibrium" (which was probably the case until before, because we don't know anything about $n$), therefore probably "more likely to be green".
2. The one red ball that was drawn is an indication that there are more red balls in the urn than green ones, therefore probably "more likely to be red".
3. The two opposite effects mentioned above will cancel each other out, therefore probably "equally likely".

Only one of these three ways of thinking can be correct, but which one?

## Wrong Approaches to Problem 1

Here, too, the task was not to quantify, but completely without quantitative considerations it will be difficult to decide what is true. Therefore, we present in the following some quantifying approaches that could be made, also by learners, for all three options (*more likely green*, *equally likely*, *more likely red*):

1. The number $n$ takes each of the values $0,1,\ldots,99,100$ with a probability of $\frac{1}{101}$. Then there are $n$ red balls (out of 100) in the urn, and after the first red ball is drawn, there are still $n-1$ red balls (out of 99), whereby the case $n=0$ can be excluded since a red ball was drawn. The probability that the next ball will be red again is:

$$P(2^{nd}\text{ ball is red}) = \frac{1}{101}\left(\frac{0}{99}+\frac{1}{99}+\ldots+\frac{98}{99}+\frac{99}{99}\right) = \frac{1}{101\cdot 99}(1+\ldots+99) = \frac{1}{101\cdot 99}\cdot\frac{99\cdot 100}{2} = \frac{50}{101} < \frac{1}{2}$$

This would suggest *more likely green*.

2. If one excludes the case $n=0$ from the very beginning because a red ball was drawn, one could also replace the above value $\frac{1}{101}$ by $\frac{1}{100}$ (only the values $1,\ldots,100$ are possible for $n$), then the result would be, analogous to the first case, $\mathrm{P}(2^{\text{nd}}\text{ ball is red}) = \frac{1}{100}\cdot\frac{99\cdot 100}{2} = \frac{1}{2}$.
This would rather speak for *equally likely*.

3. Taking this idea a little further, one could limit $n$ to the only values making sense, namely $2,\ldots,100$, since after the first red ball is drawn, another red ball is required. This would then result in the probability $\frac{1}{99}$ for the remaining values of $n$ and finally, analogous to above, $\mathrm{P}(2^{\text{nd}}\text{ ball is red}) = \frac{1}{99}\cdot\frac{99\cdot 100}{2} = \frac{50}{99} > \frac{1}{2}$. That would suggest *more likely red*.

For all three quantifying approaches, with results very near to ½, we *wrongly* considered the probabilities for the number $n$ (the number of red balls in the urn) as constant, namely $\frac{1}{101},\frac{1}{100},\frac{1}{99}$. But these values do change due to drawing the first red ball, for more details see below.

## A Concrete Solution to Problem 1

Before we go into more detail here, we present an ingenious solution (according to the newspaper article by George Lowther; in an interview with the magazine *quanta*[1] Daniel Litt calls this his favorite solution): "Imagine that you start not with 100 balls, but with 101 balls in a row. Select a ball at random. Then color the balls to the left of it green and the ones to the right red. Discard this ball, leaving 100 balls. Then choose a second ball at random. This ball corresponds to the first ball in the original task. The task shows you that you have chosen a red ball, i.e. to the right of the ball that you put away. Now select a third ball. This ball is either to the left of the first ball, between the first and second ball, or to the right of the second ball. In two of the three options, the third ball is red. The probability that the ball is red is 2/3." (Fig. 3)

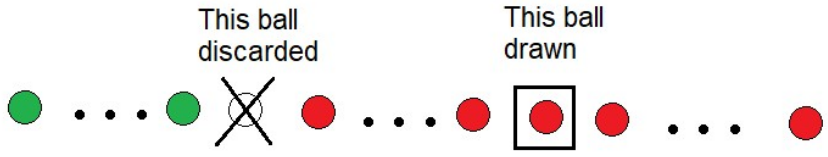



Figure 3: To the solution of George Lowther

This solution assumes that the three options mentioned are equally likely. It is not easy to judge this more precisely, but it is rather plausible because the size of the three mentioned sections depends on the random selections, and the number of balls in these sections can vary from 0 to 99 in all three cases. Since random selection means *no preference*, it may well be that, on average, everything balances out. In Paterson & Stinson (2024) such an approach is called solution by s*ymmetry*.

[1] https://www.quantamagazine.org/perplexing-the-web-one-probability-puzzle-at-a-time-20240829/

## A Solution to Problem 1 Using *Conditional Probabilities*

This result can also be confirmed by more detailed considerations. Hence, it is the correct result. Drawing the first red ball is apparently a pretty strong indication (stronger than intuitively perhaps initially assumed?) that there are many red balls in the urn, because the probability of the next ball being red is not only quite closely above 1/2, but even is 2/3.
Above we used *wrongly* constant probabilities for the number *n*. But this number is a discrete random variable, and its values do not have equal probabilities after drawing the first red ball! Only originally, i.e. a priori or, in other words, before drawing the first red ball, these probabilities were all equal, namely 1/101 („*n* is chosen uniformly at random in [0, 100]"). Altogether, this is a case for the Bayesian formula, with 101 possible cases for calculating $P(1^{st}\text{ drawing is red})$ in the denominator (see below, rule of total probability), furthermore, a famous formula for special arithmetic series – "little Gauss" – is used, yielding a nice connection to another mathematical field:

$$P(n \mid 1^{st}\text{ drawing is red}) = \frac{P(n)\cdot P(1^{st}\text{ drawing is red} \mid n)}{P(1^{st}\text{ drawing is red})} = \frac{\frac{1}{101}\cdot\frac{n}{100}}{\frac{1}{101}\cdot\left(\frac{0}{100}+\frac{1}{100}+\ldots+\frac{100}{100}\right)} =$$

$$= \frac{n}{\frac{100\cdot 101}{2}} = \frac{n}{50\cdot 101} \qquad (*)$$

As sort of test for this calculation, one can add all these values for $0 \le n \le 100$, the result should be 1, which is indeed the case ($\frac{1}{50\cdot 101}\cdot(0+1+\ldots+100) = \frac{1}{50\cdot 101}\cdot\frac{100\cdot 101}{2} = 1$).
Clearly, if there are *n* red balls out of 100 and one red ball has already been drawn, the probability that the next one will be red again is $\frac{n-1}{99}$. Thus, using the above conditional probabilities, one can easily determine the desired probability that the second ball drawn is red (again using the rule of total probability), initially as a term depending on *n*:

$$P(2^{nd}\text{ ball is red}) = \sum_{n=0}^{100}\frac{n}{50\cdot 101}\cdot\frac{n-1}{99} = \frac{1}{50\cdot 101\cdot 99}\cdot\sum_{n=0}^{100} n\cdot(n-1)$$

Nowadays, the sum $\sum_{n=0}^{100} n\cdot(n-1)$ appearing here can also be calculated using a computer algebra system, so to speak by pushing one single button, and you get a total of $P(2^{nd}\text{ ball is red}) = \frac{2}{3}$, which confirms the above solution. Using such means is entirely justifiable if one does not want to leave the area of stochastics.

But it's not that difficult to establish general formulas for the mentioned sum if one wants to link different mathematical fields (stochastics and formulas for some sums). Two short possibilities are suggested below.

- On the one hand, one can use a variable upper bound of the sum, say *k*, and write $\sum_{n=0}^{k} n\cdot(n-1) = \sum_{n=0}^{k} n^2 - \sum_{n=0}^{k} n$, then calculate these two sums separately, both famous sums, yielding altogether $\frac{k\cdot(k+1)\cdot(2k+1)}{6} - \frac{k\cdot(k+1)}{2} = \frac{(k-1)\cdot k\cdot(k+1)}{3}$.

- On the other hand, because of $n\cdot(n-1)=\frac{(n-1)\cdot n\cdot(n+1)}{3}-\frac{(n-2)\cdot(n-1)\cdot n}{3}$ (this is a "trick" students will probably not find autonomously) one can establish a telescoping series: $\sum_{n=0}^{k} n\cdot(n-1)=\sum_{n=0}^{k}\left(\frac{(n-1)\cdot n\cdot(n+1)}{3}-\frac{(n-2)\cdot(n-1)\cdot n}{3}\right)$, from which the result $\frac{(k-1)\cdot k\cdot(k+1)}{3}$ can be read off immediately, since all the other summands vanish.

There is nothing special in this problem concerning the number of balls $m=100$, it also works for all other reasonable values of $m$. For a teaching situation at school, it is probably better to take initially a smaller number than $m=100$ (as in the original problem), e.g. $m=2,3,4,5,6$ and hereby deliberately deviate from the original problem. Below we will discuss the case $m=3$ and see that the result will be 2/3 as well. If subsequently students work in groups with different values ($m=2,3,4,5,6$) yielding in each group the probability of 2/3, then the conjecture for $m=100$ is rather nearby and can be examined also more exactly. After having dealt successfully with smaller values of *m*, such more exact considerations are probably easier for students. Even if students don't consider $m=100$ but stay at considering only small values of *m*, they have achieved much!

**The case of *m* = 3 balls:** A priori, the discrete random variable *n* takes its values $0,1,2,3$ with probability 1/4. Then we have:

$$\mathrm{P}(n\,|\,1^{\text{st}}\text{ drawing is red})=\frac{\mathrm{P}(n)\cdot\mathrm{P}(1^{\text{st}}\text{ drawing is red}\mid n)}{\mathrm{P}(1^{\text{st}}\text{ drawing is red})}=\frac{\frac{1}{4}\cdot\frac{n}{3}}{\frac{1}{4}\cdot\left(\frac{0}{3}+\frac{1}{3}+\frac{2}{3}+\frac{3}{3}\right)}=\frac{n}{6}\qquad(n=0,1,2,3)$$

If there are $n=0,1,2,3$ red balls out of 3 and one red ball has already been drawn, the probability that the next one will be red again is $\frac{n-1}{2}$. Thus, we can conclude:

$$\mathrm{P}(2^{\text{nd}}\text{ ball is red}\mid 1^{\text{st}}\text{ drawing is red})=\sum_{n=0}^{3}\frac{n}{6}\cdot\frac{n-1}{2}=\frac{1}{12}\cdot(0+0+2\cdot1+3\cdot2)=\frac{8}{12}=\frac{2}{3}$$

The other cases $m=2,4,5,6$ work analogously. Even the general case, using the variable *m*, leads to the result 2/3. Here, the algebraic load is only slightly higher than in the original case $m=100$. But with $m=100$ the problem is more interesting and motivating than with a smaller value like $m=4$ or even formulating the general case, in other words using the variable *m*. Daniel Litt had here definitely the right feeling!

## A Solution to Problem 1 Using *Conditional Expected Values*

If the concept of *conditional expected value* is known there is another possibility to argue for the result $\frac{2}{3}$. Using the conditional probabilities **(*)** one can calculate

$$\mathrm{E}(n\,|\,1^{\text{st}}\text{ drawing is red})=\sum_{n=0}^{100}n\cdot\frac{n}{50\cdot101}=\frac{1}{50\cdot101}\cdot\sum_{n=0}^{100}n^2=\frac{1}{50\cdot101}\cdot\frac{100\cdot101\cdot201}{6}=67,$$

telling us that the expected value of *n* (i.e. the average value in the long run), given that the first drawn ball is red, is 67. Now, if one red ball is drawn, we have 66 red, and 33 green balls left. Thus, the probability for the next drawn ball to be red is 2/3. Also, without a formal definition of a *conditional expected value*, or even without this technical term at all, only with the idea "expected value of a discrete random variable = sum of all possible values times the

associated probabilities", and these probabilities have changed because of "1. draw is red", the above calculation is rather obvious.

## Conclusion

We have seen that some of the tasks discussed on social media have high mathematical potential. This cannot be generalized (we cannot assess it further because we have no experience in this regard), but in this case the mathematical – especially stochastic – potential is definitely considerable. The problems also seem to have high potential regarding didactics and practical teaching. Firstly, the tasks are designed without detailed calculations and justifications, so that an initial approach is possible without much fear. Secondly, the first-mentioned task can be used to train people to look at patterns and to discover interesting connections even without detailed calculations. Thirdly, the second-mentioned task can be handled easily with regards to calculations, formal aspects, and reasoning. Fourthly, what is interesting and remarkable about the last-mentioned (admittedly not easy) task is that the relative majority of people who clicked were not correct, so pressure is taken off again. Possibly, it would also suffice, after a phase in which students work autonomously such that they can come to their own (even if only gut) assessment, to discuss that the idea "indication of red" is the right one here.

The ingenious solution by George Lowther surely provides possibilities for discussions, and maybe students see and feel for themselves that one has to care for "why are the three mentioned possibilities equally likely?", because not in every case in which there are three possibilities, these do have probability 1/3 – this is a mistake that often leads to wrong results! The more precise mathematical proof that it is correct is not easy, but with the support of the teacher, especially in advanced courses, should be possible. Maybe there is an even shorter version of the confirmation? The calculation of the sum mentioned can also be transferred to a computer algebra system, yielding 2/3 as numerical result. If stochastics should be linked to formulas for special series, the potential can also be exploited in this regard. Furthermore, the principle of *conditional expected values* can be touched, not necessarily via a formal definition and corresponding theorems, also the idea of an expected value as "sum of values times corresponding *changed* probabilities" suffices.

Also, other numbers than $m = 100$ in problem 1 have high potential. First, didactic potential, because on the one hand cases like[2] $m = 2, 3, 4, 5$ are easier for students and on the other hand such considerations lead to the insight that the solution 2/3 is independent of $m$. Second, also potential regarding mathematics (primarily for pre-service teachers at universities), because the general solution involving the variable $m$ is – algebraically seen – hardly more complex than $m = 100$, but yields final certainty and fundamental insights, without diminishing Daniel Litt's merits!

The mentioned puzzles also have high potential in teacher training; overall, they certainly can enrich teaching at different levels!

[2] The case $m = 3$ structurally corresponds to the famous *Bertrand's Box Paradox*, only the context differs. Or also to the following well-known problem: In a hat there are three cards, one of them on both sides red, one of them on both sides blue, and one is red on one side and blue on the other. Somebody draws per chance one of the cards and lays it on a table; the visible side is red. Then the probability that also the other side of this card is red equals 2/3.

**References**

[1]. Humenberger, H. (2026). Probleme über Wahrscheinlichkeiten auf Social Media für den Unterricht verwenden? *Stochastik in der Schule* 46, 1, 9-14 (in German).

[2]. Paterson, M.B. & Stinson, D.R. (2024). Daniel Litt's probability puzzle. arXiv:2409.08094. URL: https://doi.org/10.48550/arXiv.2409.08094 (latest access 2026, May 11)

[3]. Taschwer, K. (2024). Mathematiker verblüfft auf Social Media mit tückischen Wahrscheinlichkeitsrätseln. DerStandard (in German), URL: https://www.derstandard.de/story/3000000234803/mathematiker-verbluefft-auf-social-media-mit-tueckischen-wahrscheinlichkeitsraetseln (latest access 2026, May 11)