Parabola Volume 57, Issue 3 (2021)

A gem for teaching elementary probability

Henk Tijms[1]

1 Introduction

Probability is the bedrock for data science and statistics. A gem for teaching probability to STEM students is the game of Egg Russian Roulette. This game was played for several years in The Tonight Show with Jimmy Fallon. In this show, Jimmy plays the Egg Russian Roulette game with a guest of the show. The guest was always a celebrity from sports or film: Tom Cruise, Anna Kendrick, Jodie Foster, David Beckham, to name a few. The guest and Jimmy take turns picking an egg from a carton and smashing it on their heads. The carton contains a dozen eggs, four of which are raw and the rest are boiled. Neither Jimmy nor the guest knows which eggs are raw and which are boiled. The first person who has cracked two raw eggs on their head loses the game.

The entertainment value of seeing famous people with raw yolk and albumin running down their hair and faces made the game very popular. Incidentally, the origin of the game has a rich history, dating back to the Middle Ages. In the rural English hamlet of Swaton (184 inhabitants, currently), the throwing of eggs started around 1322 when the new abbot of the town, who owned all of the poultry, handed out eggs to loyal churchgoers as alms. Whenever the church was cut off from the rest of the hamlet by the sometimes overflowing local river, the eggs were chucked to the churchgoers waiting on the other side of this watercourse. Recently, this tradition has been slightly adapted and restored: every year since 2006, this little village hosts a world championship of Russian Egg Roulette, which attracts contestants from all over the world.

Let's go back to Jimmy Fallon's Tonight Show. In the show, the guest is the first to choose an egg. Do you think each player has the same probability of losing the game? Does the guest of the show has an advantage because there are more hard boiled eggs to select from at the start? To answer these questions, the method of absorbing Markov chains will be used. This method essentially boils down to the use of conditional probabilities and matrix calculations, as will be explained in the next section. However, in some intermediate situations arising during the course of the game the probability of the guest of losing the game can be calculated by a simple argument. Suppose that Jimmy and the guest have each smashed one raw egg on their heads. Then, for $i = 2, \ldots, 10$, let g_i be the probability of the guest losing the game when i eggs are left in the carbon. In other words, g_i is the probability that the guest will pick as first a second raw egg when $i - 2$ boiled eggs and 2 raw eggs are left in the carbon. If i is even, then

[1]Henk Tijms Tijms is emeritus professor of operations research and honorary fellow of the Tinbergen Institute in Amsterdam (`h.c.tijms@xs4all.nl`)

the next egg will be picked by the guest; otherwise, the host Jimmy picks the next egg. Therefore,

$$g_2 = 1\,.$$

The other g_i can be recursively computed from

$$g_i = \begin{cases} \frac{i-2}{i}\, g_{i-1} & \text{for } i = 3, 5, 7, 9\,; \\ \frac{2}{i} + \frac{i-2}{i}\, g_{i-1} & \text{for } i = 4, 6, 8, 10\,. \end{cases}$$

These equations are easily explained. For i odd, the guest loses the game only if Jimmy picks a boiled egg from the i eggs left in the carbon and the guest picks as first a raw egg from the remaining $i-1$ eggs. The joint probability of these two events is $\frac{i-2}{i}$ multiplied by g_{i-1}. For i even, the probability of the guest losing the game is the sum of the probability of the event that the guest picks directly a raw egg from the carton with i eggs left and the probability of the event that the guest picks a boiled egg from the carbon with i eggs left and loses in the remainder of the game with $i-1$ eggs left in the carbon. The first event has probability $\frac{2}{i}$ and the second event has probability $\frac{i-2}{i}\, g_{i-1}$. Starting the recursive calculation with $g_2 = 1$, you obtain from the recursion equations the probabilities

$$(g_2, \ldots, g_{10}) = \left(1, \frac{1}{3}, \frac{2}{3}, \frac{2}{5}, \frac{3}{5}, \frac{3}{7}, \frac{4}{7}, \frac{4}{9}, \frac{5}{9}\right).$$

2 Intermezzo: Markov chains

Markov chain analysis is an appealing topic for secondary school students. Using material taken from Tijms [2], this section gives a first impression of the fascinating world of Markov chains. This branch of probability was founded by the Russian mathematician A.A. Markov (1856–1922) at the beginning of the 20th century. Markov's methodology goes beyond situations as coin-flipping and dice-rolling involving independent events to chains of linked events. It is a very powerful probability model that is used today in countless applications in many different areas, such as voice recognition, DNA analysis, stock control, telecommunications and a host of others. Markov chains are everywhere in science today. A nice exposition of the early history of Markov chains is given by Hayes [1].

A *Markov chain* can be seen as a dynamic stochastic process that randomly moves from state to state with the property that only the current state is relevant for the next state. In other words, the memory of the process goes back only to the most recent state. A picturesque illustration of this would show the image of a frog jumping from lily pad to lily pad with appropriate transition probabilities that depend only on the position of the last lily pad visited. In order to plug a specific problem into a Markov chain model, the state variable(s) should be appropriately chosen in order to ensure the characteristic memoryless property of the process. The basic steps of the modeling approach are:

- Choosing the state variable(s) such that the current state summarizes everything about the past that is relevant to the future states.
- The specification of the one-step transition probabilities of moving from state to state in a single step.

Using the concept of state and choosing the state in an appropriate way, surprisingly many probability problems can be solved within the framework of a Markov chain. The set of states is denoted by I and is assumed to be *finite*. The one-step transition probabilities are denoted by

$$p_{ij} = \text{the probability of going from state } i \text{ to state } j \text{ in one step}$$

for $i, j \in I$. The one-step probabilities must satisfy

$$p_{ij} \geq 0 \qquad \text{for all } i, j \in I \quad \text{and} \quad \sum_{j \in I} p_{ij} = 1 \quad \text{for all } i \in I\,.$$

It is convenient to summarize the one-step transition probabilities in a matrix P having p_{ij} as its (i, j)th element.

In Markov chains a key role is played by the n-step transition probabilities. For any $n = 1, 2, \ldots$, these probabilities are defined as

$$p_{ij}^{(n)} = \text{the probability of going from state } i \text{ to state } j \text{ in } n \text{ steps}$$

for all $i, j \in I$. Note that $p_{ij}^{(1)} = p_{ij}$. How to calculate the n-step transition probabilities? It will be seen that they can be calculated by matrix products. This key fact is based on the so-called Chapman–Kolmogorov equations

$$p_{ij}^{(n)} = \sum_{k \in I} p_{ik}^{(n-1)} p_{kj} \quad \text{for all } i, j \in I \text{ and } n = 2, 3, \ldots\,.$$

This recurrence relation can be seen by noting that the probability of going from state i to state j in n steps is obtained by summing the probabilities of the mutually exclusive events of going from state i to some state k in the first $n-1$ steps and then going from state k to state j in the nth step.

An extremely useful observation is that the n-step transition probabilities $p_{ij}^{(n)}$ can be calculated by multiplying the matrix P of one-step transition probabilities by itself n times. Let's verify this for $n = 2$:

$$p_{ij}^{(2)} = \sum_{k \in I} p_{ik} p_{kj}$$

for all $i, j \in I$. This is exactly the definition for the elements of the matrix product $P \times P = P^2$. The argument can be extended to conclude that $p_{ij}^{(n)}$ is the (i, j)th element of the n-fold matrix product P^n. This is a very important conclusion: many computations for finite-state Markov chains can be boiled down to matrix calculations! This is particularly true for so-called absorbing Markov chains. A Markov chain is said to be *absorbing* if there are one or more states i with $p_{ii} = 1$ and thus $p_{ij} = 0$ for $j \neq i$. That is, once the process is in an absorbing state, it always stays there. Absorbing Markov chains are very useful for analyzing success runs.

Let us illustrate the above concepts with two examples.

Example

The first example deals with coin-tossing: what is the probability of getting a run of three or more heads in 10 tosses of a fair coin? An absorbing Markov chain with four states can be used to answer this question. Let state i mean that the last i tosses resulted in heads for $i = 0, 1, 2, 3$. State 3 is taken as an absorbing state, and so $p_{33} = 1$ and $p_{3j} = 0$ for $j = 0, 1, 2$. For $i = 0, 1, 2$, the one-step transition probabilities are $p_{i,i+1} = p_{i0} = 0.5$ and $p_{ij} = 0$ otherwise. To find the probability of a success run of length three or more in 10 coin tosses, the matrix $P = (p_{ij})$ with $i, j = 0, 1, 2, 3$ is multiplied 10 times by itself. The resulting matrix P^{10} is

$$P^{10} = \begin{pmatrix} 0.2676 & 0.1455 & 0.0791 & 0.5078 \\ 0.2246 & 0.1221 & 0.0664 & 0.5869 \\ 0.1455 & 0.0791 & 0.0430 & 0.7324 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Since the process stays in State 3 once it is there, the event of three or more consecutive heads in 10 coin tosses occurs if and only if the process is in state 3 after 10 coin tosses. The probability of this event can be read from the first row of the matrix P^{10} and is equal to $p_{03}^{(10)} = 0.5078$.

Example

Another illustrative example of an absorbing Markov chain is the following random-walk type of problem. Joe Dalton desperately wants to raise his bankroll of $600 to $1,000 in order to pay his debts before midnight. He enters a casino to play European roulette. He decides to bet on red each time using bold play; that is, Joe bets either his entire bankroll or the amount needed to reach the target bankroll, whichever is smaller. Thus, the stake is $200 if his bankroll is $200 or $800 and the stake is $400 if his bankroll is $400 or $600. In European roulette a bet on red is won with probability $\frac{18}{37}$ and is lost with probability $\frac{19}{37}$. What is the probability that Joe will reach his goal? The solution approach to this problem will guide the solution of Egg Russian Roulette. To solve Joe's problem, take a Markov chain with six states $i = 0, 1, \ldots, 5$, where State i means that Joe's bankroll is $i \times 200$ dollars. States 0 and 5 are absorbing and the game is over as soon one of these states is reached. Thus, $p_{00} = p_{55} = 1$. The other p_{ij} are easily found. For example, the only possible one-step transitions from State $i = 2$ are to States 0 and 4, because Joe bets $\$400$ in State 2. Thus, $p_{20} = \frac{19}{37}$ and $p_{24} = \frac{18}{37}$. The other p_{ij} are given by the following matrix P of one-step transition probabilities:

from\to	0	1	2	3	4	5
0	1	0	0	0	0	0
1	$\frac{19}{37}$	0	$\frac{18}{37}$	0	0	0
2	$\frac{19}{37}$	0	0	0	$\frac{18}{37}$	0
3	0	$\frac{19}{37}$	0	0	0	$\frac{18}{37}$
4	0	0	0	$\frac{19}{37}$	0	$\frac{18}{37}$
5	0	0	0	0	0	1

For any starting state, the process will ultimately absorbed in either State 0 or State 5. The absorption probabilities can be obtained by calculating P^n for n sufficiently large. Trying several values of n, it was found that $n = 20$ is large enough to have convergence of all $p_{ij}^{(n)}$ to four or more decimals:

$$P^{20} = P^{21} = \cdots = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.8141 & 0 & 0 & 0 & 0 & 0.1859 \\ 0.6180 & 0 & 0 & 0 & 0 & 0.3820 \\ 0.4181 & 0 & 0 & 0 & 0 & 0.5819 \\ 0.2147 & 0 & 0 & 0 & 0 & 0.7853 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

You read off from row 4 that the probability of Joe reaching his goal when starting with \$600 is equal to $p_{35}^{(20)} = p_{35}^{(21)} = \ldots = 0.5819$. Alternatively, this probability can be calculated by solving four linear equations. To do so, define f_i as the probability of ever getting absorbed in State 0 when the starting state is i. By definition, $f_0 = 1$ and $f_5 = 0$. By conditioning on the next state after State i, you get the four linear equations

$$\begin{aligned} f_1 &= \frac{19}{37} \times 0 + \frac{18}{37} f_2 \\ f_2 &= \frac{19}{37} \times 0 + \frac{18}{37} f_4 \\ f_3 &= \frac{19}{37} f_1 + \frac{18}{37} \times 1 \\ f_4 &= \frac{19}{37} f_3 + \frac{18}{37} \times 1 \,. \end{aligned}$$

The solution to these linear equations is

$$f_1 = 0.1859\,, \quad f_2 = 0.3820\,, \quad f_3 = 0.5819 \quad \text{and} \quad f_4 = 0.7853\,.$$

This is the same solution as found by matrix multiplication. We are now ready to tackle the problem of Egg Russian Roulette.

3 Markov chain analysis for Egg Russian Roulette

An absorbing Markov chain is used to analyze the game of Egg Russian Roulette. The state of the Markov chain is described by the triple (i, r_1, r_2), where i denotes the number of smashed eggs, r_1 is the number of raw eggs picked by the guest and r_2 is the number of raw eggs picked by the host of the game. The states satisfy $0 \leq i \leq 11$ and $r_1 + r_1 \leq 3$. The process starts in State $(0, 0, 0)$ and ends when one of the absorbing states $(i, 2, 0)$, $(i, 2, 1)$, $(i, 0, 2)$, or $(i, 1, 2)$ is reached. The guest loses the game if the game ends in a state $(i, 2, 0)$ or $(i, 2, 1)$ with i odd. In a non-absorbing state (i, r_1, r_2) with i even, the guest picks an egg and the process goes either to state $(i+1, r_1+1, r_2)$ with probability $\frac{4-r_1-r_2}{12-i}$ or to state $(i+1, r_1, r_2)$ with probability $1 - \frac{4-r_1-r_2}{12-i}$. In a non-absorbing state (i, r_1, r_2) with i odd, the host picks an egg and the process goes either to state $(i+1, r_1, r_2+1)$ with probability $\frac{4-r_1-r_2}{12-i}$ or to state $(i+1, r_1, r_2)$ with probability $1 - \frac{4-r_1-r_2}{12-i}$. This sets the matrix P of one-step transition probabilities. The probability that the guest will lose can be computed by calculating P^{11}. This requires that the states are ordered in a one-dimensional array. It is easier to use a recursion to calculate the probability of the guest losing the game. For that, you reason in the same way as in the above gambling problem. For any state (i, r_1, r_2), let $p(i, r_1, r_2)$ be the probability that the guest will lose if the process starts in state (i, r_1, r_2). The goal is to find $p(0, 0, 0)$. This probability can be calculated by the recursion with initial conditions $p(i, 2, 0) = p(i, 2, 1) = 1$ and $p(i+1, 0, 2) = p(i+1, 1, 2) = 0$ for $i = 3, 5, 7, 9, 11$ and recursive identities

$$p(i, r_1, r_2) = \frac{4 - r_1 - r_2}{12 - i} p(i+1, r_1+1, r_2) + \Big(1 - \frac{4 - r_1 - r_2}{12 - i}\Big) p(i+1, r_1, r_2)$$

for $i = 0, 2, 4, 6, 8, 10$ and

$$p(i, r_1, r_2) = \frac{4 - r_1 - r_2}{12 - i} p(i+1, r_1, r_2+1) + \Big(1 - \frac{4 - r_1 - r_2}{12 - i}\Big) p(i+1, r_1, r_2)$$

for $i = 1, 3, 5, 7, 9, 11$. The recursive computations yield the probability $\frac{5}{9}$ that the guest of the show will lose the game. Interestingly, the game turns out to be fair for the case of three raw eggs and nine boiled eggs. These result can also verified by computer simulation – a Python program is easily written. In fact, simulations of the problem are provided by the videos online of episodes of Egg Russian Roulette in The Tonight Show by Jimmy Fallon, with Higgins as unsurpassed sidekick with his characteristically shrill voice, reminiscent of the character Igor from the parody movie Young Frankenstein. In the 18 episodes I found on Internet, the guest lost the game 9 times. Remarkably, the experimental probability of 50% resulting from this very small sample size is not far away from the theoretical probability of 55.6%. It might be an interesting project for students to write a program to simulate the game many times.

References

[1] B. Hayes, First link in the Markov chain, *American Scientist* **101** (2013), 92.

[2] H.C. Tijms, *Basic Probability, What Every Math Student Should Know*, second edition, World Scientific Press, 2021.