538 Riddler Classic

Evan Rittner

27 March 2020

Problem

You start with a fair 6-sided die and roll it six times, recording the results of each roll. You then write these numbers on the six faces of *another*, unlabeled fair die. For example, if your six rolls were 3, 5, 3, 6, 1 and 2, then your second die wouldn't have a 4 on it; instead, it would have two 3s.

Next, you roll this second die six times. You take those six numbers and write them on the faces of *yet another* fair die, and you continue this process of generating a new die from the previous one.

Eventually, you'll have a die with the same number on all six faces. What is the average number of rolls it will take to reach this state?

Extra credit: Instead of a standard 6-sided die, suppose you have an N-sided die, whose sides are numbered from 1 to N. What is the average number of rolls it would take until all N sides show the same number?

Solution

For a six-sided die, it will take, on average, $\frac{57484260752921}{5945970126096}$ rolls, which is approximately 9.67. The solutions for one- through six-sided dice are shown in Fig. 1.

Justification

There are a few important initial observations to make about this problem. First, any specific way of numbering the faces of a die can be described more broadly as a member of a set. It is not important which face has which number, nor the specific numbers themselves. The only important piece of information about any numbering is the relative frequencies of the numbers on the die. For example, on a traditional 6-sided die, the relative frequencies are [1,1,1,1,1,1]. On a die with three 3s, one 6, and two 1s, the relative frequencies are [2,0,3,0,0,1], or, more compactly, [3,2,1].

The second important observation to make is that this system can be fully described by the set of all probabilities of one state transitioning to another with a set of n rolls. For example, the state [1,1,1,1,1,1]

n	Exact	Approximate
1	0	0
2	2	2
3	$\frac{22}{7}$	3.86
4	$\frac{838}{145}$	5.78
5	$\frac{468125}{60701}$	7.71
6	$\frac{57484260752921}{5945970126096}$	9.67
n	?	$1.942 \cdot (n-1)$

Figure 1: Solutions for one- through six-sided dice

has a 0.013% chance to transition to the state [6] with one roll ($(\frac{1}{6})^6 \cdot 6!$).

The first observation leads to partition numbers as a way to describe these 'relative probabilities', and the second observation leads nicely to Markov chains, and the math that describes them.

First, partitions. The states that an n-sided die can be in are described by the partitions of n. For $n = 1$, there is only one state ([1]); for $n = 2$, there are two ([1,1], [2]); and for $n = 3$ there are 3 ([1,1,1], [2,1], [3]). For this problem, we always start at the state [1,1, ... 1], and end at state $[n]$.

Each possible partition for each n serves as a state for that Markov chain. Essential to describe this problem is the transition matrix for the Markov chain. Using the transition matrix, and some linear algebra, we can solve for the expected number of steps to the state $[n]$. Take $n = 2$ as an example. There are two states, which I'll call A and B: A is [2], the end, and B is [1,1], the start. The transition matrix is as follows in Fig. **??**. (The states are in reverse alphabetical order, to put the absorbing state, A, at the bottom and on the right.)

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$$

From state B, with die faces numbered 1 and 2, there is a 50% chance two rolls will produce different numbers (1 and 2, or 2 and 1), and a 50% chance the rolls will produce the same number. Once in state A, with both faces the same number, the die will stay there. Now, we use this matrix to find how long it will take to get to state A.

The transition matrix we just described is often notated $\mathbf{P}$. The (square) sub-matrix of $\mathbf{P}$ that only contains transient states (i.e., states that are not necessarily permanent; only B in this case) is $\mathbf{Q}$. The expected number of steps before reaching the an absorbing state, when starting from the ith transient state is the ith entry of

$$\mathbf{t} = (\mathbf{I} - \mathbf{Q})^{-1} \cdot \mathbf{1}$$

where $\mathbf{I}$ is the $(n-1)$ x $(n-1)$ identity matrix and $\mathbf{1}$ is an $n-1$ length column vector. All of these matrices have size $n-1$, since our original transition matrix P is n x n, and we removed the one absorbing state to form $\mathbf{Q}$. For $n=2$, $\mathbf{t}=[2]$, which means that it takes 2 steps, on average, to reach state A from state B.

For n larger than 2, it becomes difficult to calculate the transition probabilities from every single state to another. One pattern emerges, which is that the chance to increase the number of different numbers on the die is zero. For example, the state [3,3] has probability 0 of transitioning to state [2,1,1,1,1], because there is no way to roll five different numbers with a die labelled with only two different numbers. To calculate the rest of the transition probabilities, however, I wrote a Python program.

I will put the equations for the case of $n=6$ below, at the end of this report, since they are a mess. Their evaluation yields the solution of 9.67 steps.

Equations

There are eleven different states for the case of $n=6$ (since there are 11 permutations of 6). They are: A = [6], B = [5,1], C = [4,2], D = [3,3], E = [4,1,1], F = [3,2,1], G = [2,2,2], H = [3,1,1,1], I = [2,2,1,1], J = [2,1,1,1,1], K = [1,1,1,1,1,1]. The goal of this problem is to transition from K to A. The important part of the transition matrix, $\mathbf{Q}$, is below. The rows and columns correspond to the states in reverse alphabetical order, as before, with the top left corner corresponding to K $\to$ K, and the bottom right to B $\to$ B (keep in mind, this is a subset of the full transition matrix; since A is an absorbing state, it is excluded).

$$\mathbf{Q}=\begin{bmatrix}
\frac{5}{324} & \frac{25}{108} & \frac{25}{72} & \frac{25}{162} & \frac{25}{648} & \frac{25}{162} & \frac{25}{648} & \frac{25}{3888} & \frac{25}{2592} & \frac{5}{1296} \\
0 & \frac{5}{54} & \frac{65}{216} & \frac{25}{216} & \frac{35}{648} & \frac{21}{81} & \frac{55}{648} & \frac{95}{5832} & \frac{115}{3888} & \frac{37}{1944} \\
0 & 0 & \frac{65}{324} & \frac{25}{243} & \frac{25}{324} & \frac{175}{486} & \frac{55}{486} & \frac{485}{11664} & \frac{175}{2592} & \frac{113}{3888} \\
0 & 0 & \frac{5}{36} & \frac{5}{54} & \frac{35}{648} & \frac{25}{81} & \frac{55}{324} & \frac{35}{972} & \frac{115}{1296} & \frac{31}{324} \\
0 & 0 & 0 & 0 & \frac{10}{81} & \frac{40}{81} & \frac{10}{81} & \frac{20}{243} & \frac{10}{81} & \frac{4}{81} \\
0 & 0 & 0 & 0 & \frac{5}{72} & \frac{10}{27} & \frac{5}{36} & \frac{1255}{11664} & \frac{1445}{7776} & \frac{481}{3888} \\
0 & 0 & 0 & 0 & \frac{5}{162} & \frac{35}{162} & \frac{55}{324} & \frac{215}{3888} & \frac{455}{2592} & \frac{343}{1296} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{16} & \frac{15}{32} & \frac{3}{16} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{160}{729} & \frac{100}{243} & \frac{68}{243} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{625}{11664} & \frac{1625}{7776} & \frac{1565}{3888}
\end{bmatrix}$$

And the final equation:

$$\mathbf{t}=(\mathbf{I}_{10}-\mathbf{Q})^{-1}\cdot\mathbf{1}=\begin{bmatrix}
\frac{57484260752921}{5945970126096} \\
\frac{27877455989381}{2972985063048} \\
\frac{3229060841509}{364038987312} \\
\frac{2044895497565}{242692658208} \\
\frac{13854655}{1625624} \\
\frac{157088639}{19507488} \\
\frac{22637187}{3251248} \\
\frac{119517}{16588} \\
\frac{81922}{12441} \\
\frac{230059}{49764}
\end{bmatrix}\approx\begin{bmatrix}
9.67 \\ 9.38 \\ 8.87 \\ 8.43 \\ 8.52 \\ 8.05 \\ 6.96 \\ 7.21 \\ 6.58 \\ 4.62
\end{bmatrix}$$