FiveThirtyEight's June 12, 2020 Riddler

Emma Knight

June 12, 2020

This week's riddler, from Adam Wagner, is about a colony of (hopefully benign) ever-growing bacteria:

Question 1. *You are studying a new strain of bacteria,* Riddlerium classicum *(or* R. classicum*, as the researchers call it). Each* R. classicum *bacterium will do one of two things: split into two copies of itself or die. There is* $4/5^{ths}$ *chance of the former and a* $1/5^{ths}$ *percent chance of the latter.*

If you start with a single R. classicum *bacterium, what is the probability that it will lead to an everlasting colony (i.e., the colony will theoretically persist for an infinite amount of time)?*

Extra credit: Suppose that, instead of $4/5$*, each bacterium divides with probability* p*. Now what's the probability that a single bacterium will lead to an everlasting colony?*

This riddler is a masterstroke in asking the correct question. The answer to this riddler is in the following theorem:

Theorem 2. *Given the setup of the riddler, one has the following:*

- *If* $p \leq \frac{1}{2}$*, then the odds that the colony lasts forever is* 0*.*
- *If* $p \geq \frac{1}{2}$*, then the odds that the colony lasts forever is* $2 - \frac{1}{p}$*.*

Proof. The trick here is to notice that counting the number of bacteria at every time is the wrong way to solve this. Instead, imagine a queue of all of the bacteria. When a bacterium gets to the front of the queue, it either dies with probability $1 - p$ or puts two new bacteria at the end of the queue with probability p. Call that process one iteration of the queue. Then, the only thing that is relevant is the length of the queue at each step as each bacterium behaves identically and is independent from the other bacteria. The queue starts at length 1, and the question is whether the queue runs out.

First, I will compute the odds that the colony dies in exactly $2k+1$ iterations of the queue: in order for this to happen, one needs the bacteria to split k times, and die $k+1$ times. Moreover, there must always be bacteria in the queue after each of the first $2k$ iterations, there must be exactly one bacteria after $2k$ iterations of the queue, and this bacteria must die in the $(2k+1)^{st}$ iteration of the queue. There are $\frac{1}{k+1}\binom{2k}{k}$ (the k^{th} Catalan number) possible ways to order k splits and k deaths such that there is always at least one bacterium in the queue and you end with exactly 1 bacterium in the queue after $2k$ iterations of the queue. Thus, the odds that the colony dies in $2k+1$ iterations of the queue is

$$(1-p)\left((p(1-p))^k \frac{1}{k+1}\binom{2k}{k}\right).$$

Summing this up over all k, one gets that the odds that the colony dies eventually is

$$(1-p)\sum_{k=0}^{\infty}(p(1-p))^k \frac{1}{k+1}\binom{2k}{k}.$$

Define $f(x) = \sum_{k=0}^{\infty} x^k \frac{1}{k+1}\binom{2k}{k}$. Then one has that $\frac{d}{dx}(xf(x)) = \sum_{k=0}^{\infty} x^k \binom{2k}{k} = \frac{1}{\sqrt{1-4x}}$. Thus, $xf(x) = \frac{1-\sqrt{1-4x}}{2}$ and so $f(x) = \frac{1-\sqrt{1-4x}}{2x}$. Now, I just evaluate:

$$\begin{aligned}(1-p)f(p(1-p)) &= (1-p)\frac{1-\sqrt{1-4(p(1-p))}}{2p(1-p)} \\ &= \frac{1-\sqrt{1-4p+4p^2}}{2p} \\ &= \frac{1-|1-2p|}{2p}\end{aligned}$$

If $p \leq \frac{1}{2}$, then this is $\frac{1-(1-2p)}{2p} = \frac{2p}{2p} = 1$, so the odds that the colony dies eventually is 1 and thus the odds that it lives forever is 0. If $p \geq \frac{1}{2}$ then this is $\frac{1+(1-2p)}{2p} = \frac{1-p}{p} = \frac{1}{p} - 1$, so the odds that the colony dies eventually is $\frac{1}{p} - 1$ and thus the odds that it lives forever is $2 - \frac{1}{p}$. $\square$

Thus, one has the exact probability of the colony living forever. In the particular case of the riddler, this is $2 - \frac{5}{4} = \frac{3}{4}$.

But what if there are more bacteria? Assume that the colony starts with k bacteria instead of just 1, and the probability of a bacteria splitting is p and dying is $1-p$. Clearly, if $p \leq \frac{1}{2}$ the probability that the bacteria colony doesn't die off is still 0, so I will only consider the case that $p > \frac{1}{2}$. Write x_i for the probability that the colony lives forever.

Then one has the relation $x_i = px_{i+1} + (1-p)x_{i-1}$. This can be rearranged to $p(x_{i+1} - x_i) = (1-p)(x_i - x_{i-1})$. Write $q = \frac{1-p}{p}$ and $y_i = x_{i+1} - x_0$. Then one has that $x_0 = 0$ and $\lim_{n\to\infty} x_n = 1$ by definition. Rephrasing this in terms of the y_is, one has that $y_{i+1} = qy_i$ and $\sum_{n=0}^{\infty} y_n = 1$. But that says that $y_i = q^i y_0 = q^i x_0$ and so $\frac{x_1}{1-q} = 1$, or $x_1 = 1-q$ (which is what we got before!). Then $x_i = \sum_{n=0}^{i-1} q^n x_1 = \frac{(1-q^i)x_1}{1-q} = 1 - q^i$. Thus, if $p > \frac{1}{2}$, $x_i = 1 - \left(\frac{1-p}{p}\right)^i$.