Solution to FiveThirtyEight Riddler

Dan Swenson, Black Hills State University

May 17, 2019

1 Intro

This is a solution to the Riddler Classic from May 17, 2019. See problem statement here.

First, we'd like to know the probability of the living army emerging victorious, if there are x living soldiers, and y undead soldiers. Let $V(x, y)$ denote this probability.

Given x living and y undead soldiers, if the next single-combat duel is won by the living soldier, then there will be x living and $(y - 1)$ undead soldiers. Otherwise, we'll have $(living, dead) = (x - 1, y + 1)$. Each duel is a $50 - 50$ affair, so

$$V(x, y) = \left(\frac{1}{2}\right) V(x - 1, y + 1) + \left(\frac{1}{2}\right) V(x, y - 1) \qquad (\star)$$

Also, clearly $V(x, 0) = 1$ if $x > 0$, and $V(0, y) = 0$ if $y > 0$. (We don't need to worry about what happens if x and y are both 0; assuming both armies start out with a positive number of soldiers, there will always be someone left standing on one side or the other.)

We can use these initial conditions and equation $(\star)$ to calculate $V(x, y)$ for all pairs of nonnegative integers (x, y). Determine $V(x, y)$ first for points (x, y) which lie along the diagonal line $x + y = 1$, then for all points on the line $x + y = 2$, and so on. Along a given diagonal $x + y = n$, start at $(1, n - 1)$, then $(2, n - 2)$, and so on until you reach $(n - 1, 1)$. (The probabilities $V(n, 0) = 1$ and $V(0, n) = 0$ were already determined.)

The first few values we get by this procedure are shown here (the x-axis is along the bottom row; the y-axis is along the left-hand column):

0/64							
0/32	1/128						
0/16	1/64	10/256					
0/8	1/32	9/128	56/512				
0/4	1/16	8/64	46/256	(?)			
0/2	1/8	7/32	37/128	176/512			
0/1	1/4	6/16	29/64	130/256	562/1024		
0/0.5	1/2	5/8	22/32	93/128	386/512	1586/2048	
	1/1	4/4	16/16	64/64	256/256	1024/1024	4096/4096

To determine the value labeled (?) from this information, we would take $\frac{1}{2}\left(\frac{56}{512}\right)+\frac{1}{2}\left(\frac{176}{512}\right)$. The result is $\frac{232}{1024}$.

Above, we have made no attempt to reduce the fractions. Thus, the denominator of $V(x,y)$ is easily seen to be $4^{x-1}\cdot 2^{y}$, and we have written the 1's across the bottom in this form, as well as the 0's along the left-hand column. (Note: we get 0/0.5 for the probability when $x=0$ and $y=1$. It's still 0, though.) It remains to find the numerator, which (when x and y are both positive) is simply the sum of the numerator directly below (x,y), and the numerator diagonally above and to the left of (x,y).

2 A formula

Proposition 2.1 *The numerator at (x,y) is equal to*

$$\sum_{i=x+y}^{2x+y-1}\binom{2x+y-1}{i}.$$

That is,

$$V(x,y)=\frac{\sum_{i=x+y}^{2x+y-1}\binom{2x+y-1}{i}}{4^{x-1}2^{y}}.$$

Proof: This is by strong induction. Along the y-axis (where $x=0$), the numerator should be 0, and the formula gives

$$\sum_{i=x+y}^{2x+y-1}\binom{2x+y-1}{i}=\sum_{i=y}^{y-1}\binom{y-1}{i},$$

which is the empty sum, hence 0. At $y = 0$, on the other hand, we get

$$\sum_{i=x}^{2x-1} \binom{2x-1}{i} = \frac{1}{2}\left(\sum_{i=x}^{2x-1} \binom{2x-1}{i} + \sum_{i=x}^{2x-1} \binom{2x-1}{i}\right)$$

$$= \frac{1}{2}\left(\sum_{j=0}^{x-1} \binom{2x-1}{j} + \sum_{i=x}^{2x-1} \binom{2x-1}{i}\right) = \frac{1}{2}(2^{2x-1}) = 4^{x-1} = 4^{x-1}2^y,$$

so the numerator equals the denominator, and the probability is 1, as desired. Here we have used the substitution $j = 2x-1-i$, and the very standard fact that $\binom{N}{k} = \binom{N}{N-k}$, which is easy to see by the factorial formula. We have also used the standard fact that $\sum_{k=0}^{N} \binom{N}{k} = 2^N$.

Now, for a given point (x, y), let $N(x, y)$ denote the desired numerator (so $N(x, y) = V(x, y)4^{x-1}2^y$). For a given point (x, y), with $x, y > 0$, we saw that $N(x, y) = N(x-1, y+1) + N(x, y-1)$. So by our induction hypothesis:

$$N(x,y) = \sum_{i=(x-1)+(y+1)}^{2(x-1)+(y+1)-1} \binom{2(x-1)+(y+1)-1}{i} + \sum_{i=x+(y-1)}^{2x+(y-1)-1} \binom{2x+(y-1)-1}{i}$$

$$= \sum_{i=x+y}^{2x+y-2} \binom{2x+y-2}{i} + \sum_{i=x+y-1}^{2x+y-2} \binom{2x+y-2}{i}$$

Now, we take summands from the left-hand sum, and pair them off with summands from the right-hand sum. However, there is one "extra" summand in the right-hand sum, which won't be paired with anything. The result is

$$N(x,y) = \binom{2x+y-2}{2x+y-2} + \sum_{i=x+y}^{2x+y-2} \left(\binom{2x+y-2}{i} + \binom{2x+y-2}{i-1}\right)$$

The "extra unpaired" binomial coefficient is equal to 1. Now the "Pascal's Triangle" identity, applied to each of the pairs of summands, yields

$$N(x,y) = 1 + \sum_{i=x+y}^{2x+y-2} \binom{2x+y-1}{i} = \binom{2x+y-1}{2x+y-1} + \sum_{i=x+y}^{2x+y-2} \binom{2x+y-1}{i}$$

$$= \sum_{i=x+y}^{2x+y-1} \binom{2x+y-1}{i}. \qquad \square$$

So, we have

$$V(x,y) = \frac{1}{4^{x-1}2^y} \sum_{i=x+y}^{2x+y-1} \binom{2x+y-1}{i}$$

$$= \frac{1}{2^{2x-2}2^y} \left(\sum_{i=0}^{2x+y-1} \binom{2x+y-1}{i} - \sum_{i=0}^{x+y-1} \binom{2x+y-1}{i} \right)$$

$$= \frac{2}{2^{2x-1+y}} \left(2^{2x+y-1} - \sum_{i=0}^{x+y-1} \binom{2x+y-1}{i} \right)$$

$$= 2 \left(1 - \frac{1}{2^{2x+y-1}} \sum_{i=0}^{x+y-1} \binom{2x+y-1}{i} \right)$$

$$= 2 - \frac{1}{2^{2x+y-1}} \sum_{i=0}^{x+y-1} \binom{2x+y-1}{i} - \frac{1}{2^{2x+y-1}} \sum_{i=0}^{x+y-1} \binom{2x+y-1}{2x+y-1-i}$$

$$= 2 - \frac{1}{2^{2x+y-1}} \left(\sum_{i=0}^{x+y-1} \binom{2x+y-1}{i} + \sum_{j=x}^{2x+y-1} \binom{2x+y-1}{j} \right)$$

$$= 2 - \frac{1}{2^{2x+y-1}} \left(\sum_{i=0}^{2x+y-1} \binom{2x+y-1}{i} + \sum_{i=x}^{x+y-1} \binom{2x+y-1}{i} \right)$$

$$= 2 - \frac{1}{2^{2x+y-1}} \left(2^{2x+y-1} + \sum_{i=x}^{x+y-1} \binom{2x+y-1}{i} \right)$$

$$= 1 - \frac{1}{2^{2x+y-1}} \sum_{i=x}^{x+y-1} \binom{2x+y-1}{i}$$

This is the probability that the living army wins. Thus:

Theorem 2.2

$$P(\textit{undead army wins}) = \frac{1}{2^{2x+y-1}} \sum_{i=x}^{x+y-1} \binom{2x+y-1}{i}. \qquad \square$$

Corollary 2.3

$$P(\textit{undead army wins}) =$$

$$P(\textit{between } x \textit{ and } (x+y-1) \textit{ Heads in } (2x+y-1) \textit{ flips of a fair coin}).$$

Proof: Both sides are equal to $\frac{1}{2^{2x+y-1}} \sum_{i=x}^{x+y-1} \binom{2x+y-1}{i}$. $\square$

3 Examination of small values

Our goal is to find values of x and y for which $V(x, y)$ is (approximately) $\frac{1}{2}$. For small values of y, it is easy to find the smallest x-value for which $V(x, y) \geq \frac{1}{2}$. Thus, for a given number of undead soldiers, this corresponding x-value would be the minimum number of living soldiers required to give the living army at least a 50% chance of winning.

The first several such values are

y	x
1	1
2	4
3	9
4	17
5	26
6	38
7	51
8	67
9	85
10	106
...	...
20	431
...	...
30	975
...	...
40	1739
...	...
60	3928
...	...

Conjecturally, it appears that the living army needs *slightly more* than the square of the number of undead soldiers. We can visualize this in the following way: if the undead army is arranged in a single line, then the living army needs to have a solid *square* of soldiers whose side length equals the length of the undead line.

However, this relationship is not exact; starting at $y = 4$, we need strictly more than y^2 living soldiers. So, is this really a quadratic relationship, or something else? What happens when y gets larger? If it is a quadratic relationship, then what are the coefficients?

4 Conclusion

We know that for small values of y, the number x of living soldiers required to give the living an even chance of survival is a little more than y^2. What about if y is large?

By Corollary 2.3, given y, we want to find x such that

$$P(\text{between } x \text{ and } (x+y-1) \text{ Heads in } (2x+y-1) \text{ flips of a fair coin}) \approx \frac{1}{2}$$

Since y is large, we can use the Normal Approximation to the Binomial Distribution on the left-hand side. For large y, the probability of getting between x and $(x+y-1)$ Heads in $(2x+y-1)$ flips of a fair coin is approximately

$$P(x - 0.5 \leq u \leq x + y - 0.5)$$

where u follows a normal distribution with mean $\mu = \frac{1}{2}(2x+y-1)$ and standard deviation $\sigma = \sqrt{\frac{1}{4}(2x+y-1)}$. (We have applied a "continuity correction" here; there are y values between x and $(x+y-1)$, inclusive, so the width of the interval should be y; we therefore "inflate" the interval $[x, x+y-1]$ by 0.5 on both ends to get the desired interval.) This probability is

$$\frac{1}{2}\left(1 + erf\left(\frac{x+y-\frac{1}{2}-\frac{1}{2}(2x+y-1)}{\sqrt{\frac{1}{2}(2x+y-1)}}\right)\right) - \frac{1}{2}\left(1 + erf\left(\frac{x-\frac{1}{2}-\frac{1}{2}(2x+y-1)}{\sqrt{\frac{1}{2}(2x+y-1)}}\right)\right)$$

where erf is the error function. This simplifies to

$$\frac{1}{2}\left(erf\left(\frac{\frac{1}{2}y}{\sqrt{\frac{1}{2}(2x+y-1)}}\right)\right) - \frac{1}{2}\left(erf\left(\frac{-\frac{1}{2}y}{\sqrt{\frac{1}{2}(2x+y-1)}}\right)\right)$$

$$= erf\left(\frac{\frac{1}{2}y}{\sqrt{\frac{1}{2}(2x+y-1)}}\right),$$

since erf is an odd function. So, given y, we want to solve

$$= erf\left(\frac{\frac{1}{2}y}{\sqrt{\frac{1}{2}(2x+y-1)}}\right) = \frac{1}{2}$$

for x. But, the erf function is strictly increasing, so there is a unique value of w such that $erf(w) = \frac{1}{2}$; it is approximately $w = 0.4769362762447$. Let α equal this number; we get

$$\frac{\frac{1}{2}y}{\sqrt{\frac{1}{2}(2x+y-1)}} = \alpha.$$

Equivalently,

$$\frac{1}{2}y = \alpha\sqrt{\frac{1}{2}(2x+y-1)}.$$

Now just solve for x:

$$\frac{y^2}{4} = \frac{\alpha^2}{2}(2x+y-1)$$

$$\frac{y^2}{4\alpha^2} = x + \frac{y-1}{2}$$

$$x = \frac{y^2}{4\alpha^2} - \frac{y}{2} + \frac{1}{2}$$

Thus

$$x \approx 1.099054669y^2 - 0.5y + 0.5.$$

Example: If $y = 60$, we get $x \approx 3927.097$, so we estimate that we need at least 3928 living soldiers to give ourselves at least a 50% chance of winning. This happens to agree exactly with the value shown in the table.

In fact, for $1 \leq y \leq 60$, the formula $\lceil 1.099054669y^2 - 0.5y + 0.5 \rceil$ gives the smallest value of x such that $V(x, y) \geq \frac{1}{2}$, except at $y = 1, y = 9$, and $y = 34$, where it is an overestimate of exactly 1. So it appears that this approximation will be reasonable for large y-values, while either the exact formula given in Theorem 2.2, or the recursive approach from Section 1, can be applied when y is small.