Random walks and even more peregrinations, with solutions

Geoff Galgon, Garrett Ervin

1 Counting and walking preliminaries

For this math circle, it will help to know how to count the number of ways to pick k objects from a group of n objects. This number is denoted $\binom{n}{k}$. For example, the number of ways to choose a group of two people (say, a president and a vice president) from a group of seven people is $\binom{7}{2}$. The number of ways that exactly three heads can appear in a sequence of five coin flips is $\binom{5}{3}$. The number of ways to select a team of five people from a group of eleven is $\binom{11}{5}$. And so on. The formula for calculating $\binom{n}{k}$ is $\frac{n!}{(n-k)!k!}$. We will not prove this formula, but you may take it for granted. For example, we have $\binom{5}{2} = 10$ and $\binom{11}{8} = 165$.

Question 1.1. A tipsy man stands on a number line at the origin. Every time he takes a step, it is to the right with probability 1/2 and to the left with probability 1/2. So, after one step, it is equally likely that he stands at -1 and at $+1$. After two steps, what are his possible positions? Are they equally likely? If not, give the probability that the man stands at each of the possible positions.

Solution 1.2. The possible positions are $-2, 0$, and $+2$ with respective probabilities $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$.

Question 1.3. By a *path* we will mean a particular sequence of steps. So, if the man takes two steps, the possible paths are right-right, right-left, left-right, left-left. We will denote these paths as $(1, 1)$, $(1, -1)$, $(-1, 1)$, and $(-1, -1)$ respectively. Each path *is* equally likely to occur. In the case of two steps, each path occurs with probability 1/4.

If the man walks three steps, how many possible paths are there, and what is the probability that each one occurs? In general, if the man walks N steps, what is the probability that each particular path occurs?

Solution 1.4. Each step left or right has probability $\frac{1}{2}$ of occurring, so a particular path of length N has probability $\frac{1}{2^N}$ of occurring.

Question 1.5. Let us denote a path of N steps taken by the man as $(e_0, e_1, \ldots, e_N)$, where every e_i is either 1 or -1 to indicate either a step to the right or to the left, respectively. For such a path, the mans final position is given by $e_0 + e_1 + \ldots + e_N$. If N is an even number, it is possible that the man ends up back at his starting place (that is, at 0). For $N = 2$, we have seen that there are exactly two paths ending at 0. If $N = 4$, what is the number of paths ending at 0? What about for $N = 6$? In general, for N an even number, what is the number of paths ending at 0? Let T_n denote the number of paths that end at 0 after $N = 2n$ steps.

Solution 1.6. To end at 0, the man must take as many steps left as he takes right, that is, such a path must have as many e_i be -1 as $+1$. There are $\binom{N}{N/2}$ many ways this can happen, that is, there are $\binom{N}{N/2}$ many paths ending at 0.

Question 1.7. Combine your answers from questions 2 and 3 to calculate, for a walk of even length N, the probability that the man ends up at 0.

Solution 1.8. Combining the above, we see that the probability that the man ends up at 0 is $\binom{N}{N/2} \cdot \frac{1}{2^N}$.

Question 1.9. For a walk of length 3, find the probability that the man ends up at 1. What about for a walk of length 4? Of length 5? In general, for a walk of odd length N, calculate the probability that the man finishes his walk at $+1$.

Solution 1.10. Suppose $N = 2n + 1$ is odd. To end up at 1, the man must take $n + 1$ steps right, and n steps left. Reasoning similarly to before, we see that the probability he ends up at 1 is $\binom{N}{n+1} \cdot \frac{1}{2^N}$.

Question 1.11. Generalize your calculations: For a walk of even length N, calculate the probability of ending at m (for m an even number between N and $-N$). What about for odd N, m?

Solution 1.12. To end up at m, the man must take m more steps right than left. Let r, l denote the number of right and left steps respectively. Then we must have $r - l = m$ and $r + l = N$, which gives $2r = N + m$, so that $r = \frac{N+m}{2}$. Thus the probability that the man ends at m is $\binom{N}{(N+m)/2} \frac{1}{2^N}$.

Question 1.13. Suppose now that the probability that our merry drunkard moves to the right at every step is $3/4$ and the probability that he moves to the left is $1/4$. Repeat the above with these new probabilities.

Solution 1.14. The only difference here is the probability of a particular path occurring. A path consisting of r right steps and l left steps now has probability $(\frac{3}{4})^r (\frac{1}{4})^l$ of occurring. So, the probability of ending up at m over a walk of N steps is

$$\binom{N}{(N+m)/2} (\frac{3}{4})^{(N+m)/2} (\frac{1}{4})^{(N-m)/2}.$$

2 Further walking, three sheets to the wind

Question 2.1. A man who's had too much to drink stands at the edge of a cliff. If he walks forward from here he falls and meets his end. At each step, he walks forward with probability 1/3, and backwards with probability 2/3. What is the probability that the man eventually dies by walking off the cliff (assuming he is capable of taking infinitely many steps)?

Hint: Let x_0 denote the probability that he dies, given that he starts at the edge of the cliff. Let x_1 denote the probability that he dies given that he starts one step behind the edge of the cliff. See if you can find two equations which relate x_0 and x_1 and allow you to solve for x_0, the quantity in question.

Solution 2.2. From the edge of the cliff, the man moves forward with probability 1/3 and backwards with probability 2/3. So we have

$$x_0 = 1/3 + (2/3)x_1.$$

On the other hand, the probability of starting one step behind the edge of the cliff and ever reaching the cliff's edge will be the same as the probability of standing at the cliff's edge and ever going over. And because from a position one step behind the cliff's edge our man must step first to the cliff's edge, then step over the cliff's edge, we have with independence

$$x_1 = x_0^2.$$

One can now solve for x_0, yielding two solutions, $x_0 = 1$ and $x_0 = 1/2$. The correct answer is $x_0 = 1/2$. It is, however, not immediately obvious that $x_0 = 1$ isn't a valid solution, and the only clear way (other than using heuristic continuity arguments, or more sophisticated things like drift) we can see this is to evaluate the probability explicitly with the Catalan numbers, e.g.

$$x_0 = \sum_{n \in \mathbb{N}} C_n (1/3)^{n+1} (2/3)^n = 1/2,$$

where C_n is the nth catalan number, e.g. $\frac{1}{n+1}\binom{2n}{n}$. One can evaluate this with a computer.

Question 2.3. Solve the question above more generally: A man who's had too much to drink stands at the edge of a cliff. If if walks forward from the edge he falls and meets his end. He walks forward with probability p, and backwards with probability $1 - p$. What is the probability that the man eventually dies by walking off the cliff (assuming he is capable of taking infinitely many steps)? Draw a graph of these probabilities with the y-axis probability of death and the x-axis probability of moving forward with one step.

Solution 2.4. Following the same method as above, one finds $x_0 = \frac{p}{1-p}$ for $p \leq 1/2$. At 1/2 this is equal to one, and from then on the man will fall off the cliff with probability one. Graphing this function is straightforward.

Question 2.5. In order to fall of the cliff, the man has to appear (back) at the edge of the cliff, and then take a step forward. Another method of computing the probabilities above first involves counting the number of paths that he can take which lead him back to the edge at step N (even), as in the section above, *but only those in which he doesn't at any point step over the cliff.* How many paths are there with two steps allowed? How many paths are there with four steps?

Solution 2.6. $C_1 = 2$. $C_2 = 4$.

Question 2.7. How many such paths (where he never falls off the cliff but returns to the edge at the end) are there with 6 steps? What about with 8 steps? Let C_n denote the number of such paths for an even $N = 2n$.

Solution 2.8. $C_3 = 5$. $C_4 = 14$.

Question 2.9. What proportion of the total number of paths that end up at the starting position do those which never cross over to one side (C_n) fill in these examples? That is, what is $\frac{C_n}{T_n}$?

Solution 2.10. $\frac{C_n}{T_n} = \frac{1}{n+1}$.

Question 2.11. The "C" in the C_n above stands for "Catalan," and these C_n are called the Catalan numbers. There are many different equivalent ways of defining them. We have seen one here, as the number of paths starting and ending at 0 in $2n$ steps never crossing to the other side of 0. As another way of viewing these numbers, consider an $n \times n$ grid, and count the number of possible paths which travel from one corner to the diagonally opposite corner, never crossing the diagonal.

For example, for $n = 1$, the grid includes the vertices $\{(0,0), (0,1), (1,0), (1,1)\}$. If we have to travel from the vertex $(0,0)$ to the vertex $(1,1)$ without crossing the diagonal, there is only one path, $\langle (0,0), (0,1), (1,1) \rangle$. For $n = 2$, the grid includes the 9 vertices, $\{(0,0), \ldots, (2,2)\}$, and there are two paths that don't cross the diagonal. Explain why this characterization for C_n works. That is, why is it equivalent to our previous formulation?

Solution 2.12. Crossing above the diagonal line means exactly taking one more vertical step than horizontal on the grid, which can be directly translated to crossing the origin by taking one more step to the right than to the left, for example.

Discussion Question 2.13. Using the grid characterization given above, can you prove the formula you found for C_n in Question 2.5? This is difficult, we will help you.

Solution 2.14. There are many proofs of the identity $C_n = \frac{1}{n+1}\binom{2n}{n}$. One may check the Wikipedia entry on Catalan numbers to see several. Perhaps the easiest is based on a reflection method, which establishes a bijection between the number of paths which cross the diagonal in a square $n \times n$ grid and the number of monotonic paths in an $(n-1 \times n+1)$ grid. This latter quantity is $\binom{(n-1)+(n+1)}{n-1} = \binom{2n}{n-1}$. Thus the number of paths which don't cross above the diagonal is

$$\binom{2n}{n} - \binom{2n}{n-1},$$

which when expanded using factorial is easily seen to be $\frac{1}{n+1}\binom{2n}{n}$.

Question 2.15. Instead of the traditionally macabre framing adopted above, let's suppose that our man is walking on a number line, starting at the origin. Additionally, suppose that he as a 1/2 chance of moving to the left or to the right. What is the probability that he returns to the origin eventually?

Hint: As done previously, it may be helpful to calculate the probability that he is at the origin after N (N is even) steps. Then the expectation of the number of times he hits the origin will be the infinite sum over all even N of these probabilities (why?). What does this quantity tell you?

Solution 2.16. As indicated in the hint, if X denotes the number of times our man hits the origin, $E(X) = \sum_{n\geq 1} (\frac{1}{2})^{2n}\binom{2n}{n}$. This sum does not converge (using Sterling's approximation, for example). That is, $E(X) = \infty$. One may argue that this implies that, if p denotes the probability of returning to the origin, $p = 1$, for if $p < 1$, then $E(X) \leq \sum_{n\geq 1} Cnp^n < \infty$.

3 Discussion problems

Question 3.1. Suppose now that our man moves up to two dimensions, and on the Cartesian grid he can move up, down, left, or right with probabilities 1/4, respectively. What is the probability that he returns to the origin?

Solution 3.2. In order to return to the origin now, our man must make an equal number of moves to the left and to the right, and an equal number of moves up and down. Each path has probability $(\frac{1}{4})^n$ probability of occurring, so if we take k steps up, k steps down, $n-k$ steps to the left, and $n-k$ steps to the right, summing over all k each of these multinomial possibilities, we will have the probability of being at the origin at step $2n$, p_{2n}:

$$p_{2n} = (\tfrac{1}{4})^{2n} \cdot \sum_{k\leq n} \frac{(2n)!}{(k!)^2((n-k)!)^2}.$$

This expression may be simplified using the identity that $\sum \binom{n}{k}^2 = \binom{2n}{k}$ to show that

$$p_{2n} = (1/4)^{2n} \binom{2n}{n}^2.$$

Using Sterling's approximation again, we can see that this sum $\sum_{n \geq 1} p_{2n} = \infty$. While in the one dimensional case this behaved of order $\sum \frac{1}{n^{1/2}}$, the two dimensional case behaves of order $\sum \frac{1}{n^1}$. The probability of returning to the origin is 1.

Question 3.3. Suppose now that our man steps into three dimensions, and can move any of forward, backwards, left, right, up, or down, with probabilities 1/6, respectively. What is the probability that he returns to the origin?

Note: This requires a computer to get an accurate number.

Solution 3.4. As we might expect, because when going from one dimension to two dimensions we went from $\sum \frac{1}{n^{1/2}}$ to $\sum \frac{1}{n^1}$, in three dimensions we would be considering something like $\sum \frac{1}{n^{3/2}}$, which unlike the other two sums, converges. And our man might never return to the origin. This suspicion is correct. Much as before, we can compute:

$$p_{2n} = (\tfrac{1}{6})^{2n} \cdot \sum_{i+k \leq n} \frac{(2n)!}{(k!)^2 (i!)^2 ((n-i-j)!)^2}.$$

One may use Sterling's approximation to get an upper bound to show that $p_{2n} \leq \frac{C}{n^{3/2}}$ for some constant C, i.e. the sum converges. That is, $E(X) < \infty$, so the probability of returning to the origin is < 1. In order to compute this probability, one may use a computer and show that the probability of returning to the origin is roughly .34. These constants, $1, 1, .34, \ldots$ are called Pólya's Random Walk Constants. Surprisingly, a closed form solution in terms of the gamma function is known for the three dimension case, but not for higher dimensions (though there are solutions in terms of infinite sums or infinite integrals).