Solution

Week 40 (6/16/03)

Staying ahead

Consider a two-dimensional lattice on which a vote for A is signified by a unit step in the positive x-direction, and a vote for B by a unit step in the positive y-direction. Then the counting of the votes until A has a votes and B has b votes corresponds to a path from the origin to the point (a, b), with $a > b$. There are $\binom{a+b}{a}$ such paths (because any a steps of the total $a + b$ steps can be chosen to be the ones in the x-direction). All of these paths from the origin to (a, b) are equally likely, as you can show.[1] The probability that a particular path corresponds to the way the votes are counted is thus $1/\binom{a+b}{a}$.

The problem can therefore be solved by finding the number, N_g, of paths that reach the point (a, b) without passing through the region $y > x$. (We'll call these the "good" paths; hence the subscript "g".)

It will actually be easier to find the number, N_b, of paths that reach the point (a, b) and that do pass through the region $y > x$. (We'll call these the "bad" paths.) The desired probability that A's sub-total is always greater than or equal to B's sub-total equals $1 - N_b/\binom{a+b}{a}$

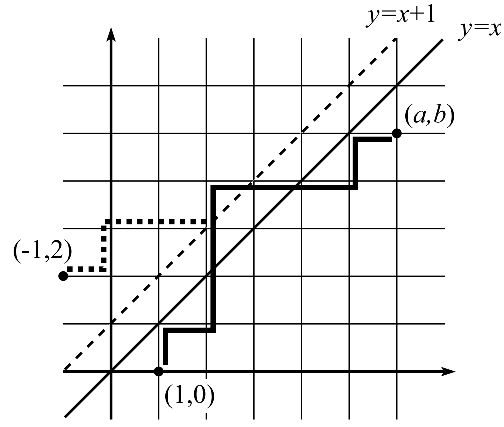
Claim: *The number of "bad" paths from the origin to (a, b) (that is, the number of paths that pass through the region $y > x$) equals $N_b = \binom{a+b}{b-1}$.*

Proof: The number of bad paths from $(0, 0)$ to (a, b) equals the number of bad paths from $(0, 1)$ to (a, b) plus the number of bad paths from $(1, 0)$ to (a, b). Let's look at these two classes of paths.

- The number of bad paths from $(0, 1)$ to (a, b) is simply all of the paths from $(0, 1)$ to (a, b), which equals $\binom{a+b-1}{b-1}$.

- The number of bad paths from $(1, 0)$ to (a, b) equals the number of paths from $(-1, 2)$ to (a, b). This follows from the fact that any bad path from $(1, 0)$ to (a, b) must proceed via a point on the line $y = x + 1$, as shown below. Hence, there is a one-to-one correspondence between the bad paths starting at $(1, 0)$ and all of the paths starting at $(-1, 2)$. This correspondence is obtained by reflecting, in the line $y = x + 1$, the portion of the bad path between $(1, 0)$ and the first point of contact with $y = x + 1$. An example of such a reflection is shown below. The number of paths from $(-1, 2)$ to (a, b) is $\binom{a+b-1}{b-2}$.

[1] The probabilities are all equal to $a!b!/(a + b)!$, as you can see by picking an arbitrary path and writing down the product of the probabilities of each step. This result can be written as $1/\binom{a+b}{a}$, which makes sense, because $\binom{a+b}{a}$ is the total number of paths.



The total number of bad paths from $(0,0)$ to (a,b) is therefore

$$N = \binom{a+b-1}{b-1} + \binom{a+b-1}{b-2} = \binom{a+b}{b-1}. \quad \blacksquare \tag{1}$$

The probability that A's sub-total is always greater than or equal to B's sub-total is therefore

$$P_{A\geq B} = 1 - \binom{a+b}{b-1} \Big/ \binom{a+b}{a} = 1 - \frac{b}{a+1}\,. \tag{2}$$

REMARKS:

1. We may also ask what the probability is that A's sub-total is always strictly greater than B's. The same "reflection" reasoning holds, except that now we must reflect across the line $y = x$, because any path that touches the line $y = x$ is now "bad". The two classes of paths in the above proof are now identical (they're both the ones that go from $(0,1)$ to (a,b)). Therefore, the number of bad paths is twice the number of paths from $(0,1)$ to (a,b), that is, $2\binom{a+b-1}{b-1}$.

   The probability that A's sub-total is always strictly greater than B's sub-total is therefore

$$P_{A>B} = 1 - 2\binom{a+b-1}{b-1} \Big/ \binom{a+b}{b} = \frac{a-b}{a+b}\,. \tag{3}$$

2. Let's consider the special case (of the original problem) where $a = b$, that is, where the election ends in a tie. From eq. (2), we see that that if $a = b \equiv n$, then $P_{A\geq B} = 1/(n+1)$. There is another very nice method of solving this special case. It is the method of generating functions, and it proceeds as follows.

   Let A_n be the total number of "good" paths that go from $(0,0)$ to (n,n), and let B_n be the number of "good" paths that touch the line $y = x$ for the first time on the $2n$th step. (The A_n and B_n here have nothing to do with the people A and B given in the problem.) We will find a relation between A_n and B_n, and then find a recursion relation for A_n.

   The relation between the A_n and B_n may be found as follows. The number, B_n, of good paths touching $y = x$ for the first time after $2n$ steps equals the number of paths which go from $(1,0)$ to $(n,n-1)$ without touching the line $y = x$. If we imagine shifting out coordinate system, we see that this equals the number of paths that go from $(0,0)$ to $(n-1,n-1)$ without going above the line $y = x$. But this is just A_{n-1}. So we have

$$B_n = A_{n-1}. \tag{4}$$

Now let's find a recursion relation for A_n. We may categorize the A_n good paths from $(0,0)$ to (n,n) according to when they first touch the line $y=x$. There will be, for example, some paths that touch this line after 2 steps. The number of these is $B_1 A_{n-1}$. This is true because there are B_1 $(=1)$ paths that touch the line for the first time at the point (1,1), and then there are A_{n-1} good paths that go from $(1,1)$ to (n,n). Likewise, there are $B_2 A_{n-2}$ paths that touch $y=x$ for the first time at $(2,2)$, on the way to (n,n). Continuing in this fashion, we find

$$A_n = B_1 A_{n-1} + B_2 A_{n-2} + \cdots + B_{n-1} A_1 + B_n A_0, \tag{5}$$

where A_0 is defined to be 1.

Using eq. (4), we obtain the recursion relation

$$A_n = A_0 A_{n-1} + A_1 A_{n-2} + \cdots + A_{n-2} A_1 + A_{n-1} A_0, \tag{6}$$

We can calculate the first few of these. They are $A_0 = 1$, $A_1 = 1$, $A_2 = 2$, $A_3 = 5$, $A_4 = 14$, and $A_5 = 42$. A lucky guess can produce the result $A_n = \binom{2n}{n}/(n+1)$, but let's obtain this in a more deductive way.

A nice way to solve recursion relations such as the one in eq. (6) is to use generating functions. By this, we mean consider the polynomial

$$F(x) = A_0 + A_1 x + A_2 x^2 + A_3 x^3 + \cdots. \tag{7}$$

The key in this problem is to note that if we square $F(x)$ and use eq. (6), we obtain

$$\begin{aligned} (F(x))^2 &= A_0 A_0 + (A_0 A_1 + A_1 A_0)x + (A_0 A_2 + A_1 A_1 + A_2 A_0)x^2 \\ &\qquad + (A_0 A_3 + A_1 A_2 + A_2 A_1 + A_3 A_0)x^3 + \cdots \\ &= A_1 + A_2 x + A_3 x^2 + A_4 x^3 + \cdots \\ &= (F(x) - 1)/x. \end{aligned} \tag{8}$$

Therefore, $F(x)$ satisfies the equation,

$$xF^2 - F + 1 = 0. \tag{9}$$

The solution to this is

$$F = \frac{1 \pm \sqrt{1-4x}}{2x}. \tag{10}$$

We must now expand the square root in a Taylor series in order to obtain the coefficients of the various powers of x. It is clear that we must choose the minus sign if we want the coefficient of x^0 (which is A_0) to be 1. By taking a plethora of derivatives, you can show that the Taylor expansion of $\sqrt{1+y}$ can be written as

$$\begin{aligned} \sqrt{1+y} &= 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(\left(\frac{1}{2}\right) \left(\frac{-1}{2}\right) \left(\frac{-3}{2}\right) \left(\frac{-5}{2}\right) \cdots \left(\frac{-(2n-3)}{2}\right) \right) y^n \\ &= 1 + \sum_{n=1}^{\infty} \left(\frac{(-1)^{n-1}(2(n-1))!}{2^{2n-1} n!(n-1)!} \right) y^n. \end{aligned} \tag{11}$$

Therefore, with $y \equiv (-4x)$, eq. (10) gives

$$F = \sum_{m=0}^{\infty} \left(\frac{(2m)!}{m!(m+1)!} \right) x^m. \tag{12}$$

Thus,

$$A_n = \frac{(2n)!}{n!(n+1)!} = \frac{1}{n+1}\binom{2n}{n} \tag{13}$$

is the number of "good" paths that go from $(0,0)$ to (n,n). Since the total number of paths from the origin to (n,n) is $\binom{2n}{n}$, the probability of a good path is $1/(n+1)$, as we wanted to show.