Why is \( \frac{1}{n+1} \binom{2n}{n} = \frac{1}{2n+1} \binom{2n+1}{n} = \frac{1}{n} \binom{2n}{n-1} \)?

DAVID CALLAN
Department of Statistics
University of Wisconsin-Madison
1210 W. Dayton St
Madison, WI 53706-1693
callan@stat.wisc.edu

December 13, 2004

1. **Introduction** The Catalan number \( C_n \) is the cardinality (number of elements) of the set \( D_n \) of Dyck paths or “mountain ranges” of size \( n \)—the lattice paths of \( n \) upsteps and \( n \) downsteps that never dip below “ground level”. Writing an upstep as \(+1\) and a downstep as \(-1\), a Dyck path can be represented as a sequence of \( n \) 1’s and \( n - 1 \)’s whose partial sums are all nonnegative. We will review three methods of counting Dyck paths of size \( n \) leading, respectively, to the three expressions in the title (each of which is equal to \( C_n \)).

Each of the title expressions consists of a binomial coefficient \( b \) divided by a divisor \( d \). These divisors are \( n + 1 \), \( 2n + 1 \) and \( n \) respectively. In each case there is a class \( \mathcal{L} \) of lattice paths whose cardinality is the binomial coefficient \( b \), and a way of coding these paths as sequences of length \( d \) with a special property: a code sequence cyclically rotated is again a valid code sequence. In each case, there is a parameter \( \nu \) defined on \( \mathcal{L} \) (equivalently, on its code sequences) with the following property: \( \nu \) has \( d \) possible values and on the \( d \) sequences obtained by cyclically rotating a given code sequence, \( \nu \) takes on each of its possible values exactly once. Thus \( \mathcal{L} \) is partitioned into \( \frac{b}{d} \) cyclic-rotation equivalence classes each of size \( d \). The sequences for which \( \nu \) has its maximum possible value are the Dyck paths of size \( n \) (first case) or are in a simple one-to-one correspondence with them (the other two cases). Thus there is one Dyck path per equivalence class and so \( |D_n| = b/d \), the quotient of the binomial coefficient and its divisor.

2. **Case** \( d = n + 1 \) Here \( \mathcal{L} (= \mathcal{L}_1) \) is the set of all \( \binom{2n}{n} \) lattice paths of \( n \) upsteps and \( n \) downsteps. For such a path, the \( n \) downsteps serve to separate \( n + 1 \) ascents each consisting of (zero or more) contiguous upsteps. The path is coded by the lengths of these ascents (see Figure 1).
$n = 4$, ascent length sequence $= (1, 0, 3, 0, 0)$, $\nu = 3$ upsteps above ground level

Figure 1

The parameter $\nu$ is the number of upsteps above the $x$-axis and $\nu$ ranges over $[0, n]$. The paths where $\nu$ has its maximal value $n$ actually are the Dyck paths of size $n$.

3. Case $d = 2n + 1$ Here $\mathcal{L} (= \mathcal{L}_2)$ is the set of all $\binom{2n+1}{n}$ lattice paths of $n + 1$ upsteps and $n$ downsteps. Here a path is coded simply by its sequence of $\pm 1$’s. The parameter $\nu$ is the number of the path vertices strictly above the $x$-axis (or the number of positive partial sums of its $\pm 1$ sequence, see Figure 2).

$$\pm 1 \text{ sequence } = (-1, 1, 1, -1, 1, 1, -1)$$
$$\text{partial sums } = (-1, 0, 1, 0, 1, 1, 1) \text{ with positive entries at positions } 3, 5, 6, 7$$
$$\nu = 4 \text{ vertices above ground level}$$

Figure 2

The terminal point is always above the $x$-axis, so $\nu$ ranges over $[1, 2n + 1]$. The paths with $\nu = 2n + 1$ correspond bijectively to Dyck paths of size $n$ via: delete first step (necessarily an upstep).

4. Case $d = n$ Here $\mathcal{L} (= \mathcal{L}_3)$ is the set of all $\binom{2n}{n-1}$ lattice paths of $n + 1$ upsteps and $n - 1$ downsteps. Here a path is coded by by the lengths of its ascents as in the first case except now, with only $n - 1$ downsteps in a path, the code sequences are of length $n$ The parameter $\nu$ is the number of upsteps above the line $y = 1$ and this time $\nu$ ranges over $[1, n]$. The paths with $\nu = n$ correspond bijectively to Dyck paths of size $n$ via: flip the rightmost upstep between levels $y = 1$ and $y = 2$ (there must be one!) to a downstep.
To reverse this map, locate the last downstep that returns a Dyck path to the $x$-axis and flip it to an upstep (see Figure 3).

![Figure 3]

It is graphically obvious what a rotation does to a path in $\mathcal{L}_2$: just transfer the initial step to the end of the path. In $\mathcal{L}_1$ and $\mathcal{L}_3$, to rotate the code sequence of ascent lengths to the left, cut the path into 3 parts: the initial (possibly empty) ascent, the immediately following downstep, and a terminal segment. Then interchange the initial ascent and terminal segment.

5. **Proofs** In each case, the heart of the matter is showing that $\nu$ takes on different values on each of the $d$ cyclic rotations of a code sequence. See [1] for a picture proof in case $d = 2n + 1$ that $\nu$ takes on its maximal value exactly once in each cyclic rotation class. (The same picture [1, p. 360] actually shows that $\nu$ takes on each of its possible values exactly once.) Exercise: devise similar proofs for the other cases. Note that the lattice paths considered terminate respectively at $y = 0$, $y = 1$ and $y = 2$ in the three cases.

6. **Historical Notes** Kai-Lai Chung and William Feller [2] proved in 1949 using generating functions that $\nu$ is uniformly distributed on $\mathcal{L}_1$ (lattice paths of $n$ upsteps and $n$ downsteps, or coin-tossing games that come out even). This is often referred to as the Chung-Feller theorem, but it was already known in 1908 to Major Percy A. MacMahon [3, p. 168, “a remarkable theorem”] who proved it using formal series (of words on an alphabet). Narayana [4] proved the full assertion for $\mathcal{L}_1$ in 1967. In 1947 Dvoretzky and Motzkin [5] showed $\nu$ takes on its maximum value just once in each rotation class in $\mathcal{L}_2$ (this is the basic Cycle Lemma), and I haven’t seen $\mathcal{L}_3$ in the literature. For further remarks and generalizations of the Cycle Lemma see [6].

7. **Final Remark** Of course, none of the three title expressions makes it immediately obvious that $C_n$ is an integer. A fourth method of counting Dyck paths—the André reflection principle of 1887 [7, Chap. 3.1]—does so, expressing $C_n$ as $\binom{2n}{n} - \binom{2n}{n-1}$. For
more on an automated method of expressing quantities like those in the title as integer linear combinations of binomial coefficients, see [8].

References


