A Combinatorial Interpretation of $\frac{j}{n}\binom{kn}{n+j}$

David Callan
Department of Statistics
University of Wisconsin-Madison
1300 University Ave
Madison, WI 53706-1532
callan@stat.wisc.edu

April 21, 2006

Abstract

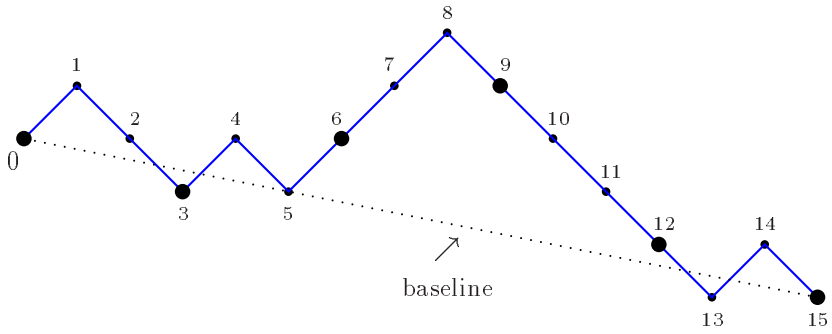
The identity $\frac{j}{n}\binom{kn}{n+j} = (k-1)\binom{kn-1}{n+j-1} - \binom{kn-1}{n+j}$ shows that $\frac{j}{n}\binom{kn}{n+j}$ is always an integer. Here we give a combinatorial interpretation of this integer in terms of lattice paths, using a uniformly distributed statistic. In particular, the case $j = 1, k = 2$ gives yet another manifestation of the Catalan numbers.

1 Introduction

For each pair of integers $j \geq 1$ and $k \geq 2$, the sequence $\left(\frac{j}{n}\binom{kn}{n+j}\right)_{n \geq \frac{j}{k-1}}$ consists of integers since $\frac{j}{n}\binom{kn}{n+j} = (k-1)\binom{kn-1}{n+j-1} - \binom{kn-1}{n+j}$. For $j = 1, k = 2$ this sequence is the Catalan numbers, A000108 in the On-Line Encyclopedia of Integer Sequences ; for $j = 1, k = 3$ it is A007226 and for $j = 1, k = 4$ it is A007228 . In this note, we give a combinatorial interpretation for all j, k in terms of lattice paths. We first treat the case $j = 1$, which is simpler (§2), specialize to $k = 1$ (§3), then generalize to larger j (§4), and end with some remarks (§5).

2 Case $j = 1$

Let $\mathcal{P}_{n,k}$ denote the set of lattice paths of $n+1$ upsteps $U = (1,1)$ and $(k-1)n-1$ downsteps $D = (1,-1)$. Clearly, $|\mathcal{P}_{n,k}| = \binom{kn}{n+1}$: choose locations for the upsteps among the total of kn steps. A path in $\mathcal{P}_{n,k}$ has $kn+1$ vertices or *points*: its initial and terminal points and $kn-1$ interior points. Define the *baseline* of $P \in \mathcal{P}_{n,k}$ to be the line joining its initial and terminal points. For $P \in \mathcal{P}_{n,k}$ label its points $0, 1, 2, \ldots, kn$ left to right and define the *k-divisible* points of P to be those whose label is divisible by k. An example with $n = 5$ and $k = 3$ is illustrated (k-divisible points indicated by a heavy dot).



A path in $\mathcal{P}_{5,3}$

Consider the statistic X on $\mathcal{P}_{n,k}$ defined by $X = \#\text{interior } k\text{-divisible points lying strictly above the baseline.}$ In the illustration $X = 3$ (points 6, 9 and 12).

Theorem. *The statistic X on $\mathcal{P}_{n,k}$ is uniformly distributed over $0, 1, 2, \ldots, n-1$.*

The following count is an immediate consequence of the theorem by considering the paths with $X = n - 1$.

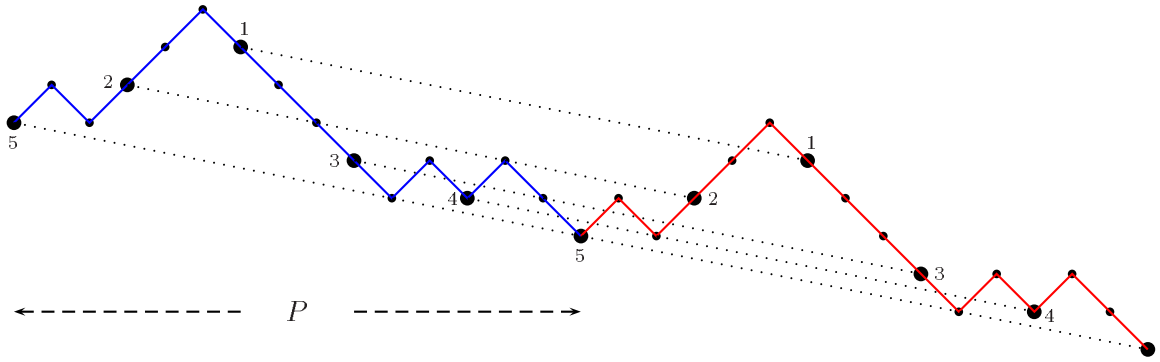
Corollary. $\frac{1}{n}\binom{kn}{n+1}$ *is the number of paths in $\mathcal{P}_{n,k}$ all of whose interior k-divisible points lie strictly above its baseline.*

Proof of Theorem Consider the operation "rotate left k units" on $\mathcal{P}_{n,k}$ defined by transferring the initial k steps of a path in $\mathcal{P}_{n,k}$ to the end. This rotation operation partitions $\mathcal{P}_{n,k}$ into rotation classes. We claim (i) each such rotation class has size n, and (ii) X assumes the values $0, 1, 2, \ldots, n-1$ in turn on the paths of a rotation class. The first claim follows from

Lemma. *Given $P \in \mathcal{P}_{n,k}$, the only k-divisible points lying on its baseline are its initial and terminal points.*

Proof of Lemma Suppose ik, $0 \le i \le n$, is a k-divisible point on the baseline. Since the slope of the baseline is $-\frac{(k-2)n-2}{kn}$, this says that the point with coordinates $(ik, -i\frac{(k-2)n-2}{n})$ lies on P (taking the initial point of P as origin). For each point (x, y) on P, x and y must have the same even/odd parity. Hence $ik \equiv i\frac{(k-2)n-2}{n} \bmod 2$. Simplifying, we find $2i \equiv \frac{2i}{n} \bmod 2 \Rightarrow i \equiv \frac{i}{n} \bmod 1 \Rightarrow n \,|\, i \Rightarrow i = 0$ or n, the last implication because $0 \le i \le n$. □

To prove the second claim, we exhibit a bijection from the paths in $\mathcal{P}_{n,k}$ with $X = n-1$ to those with $X = i$ for each $i \in [0, n-1]$. Given $P \in \mathcal{P}_{n,k}$ with $X = n-1$, draw its baseline L. The entire rotation class of P can be viewed in a single diagram: draw a second contiguous copy of P as illustrated, then join the two occurrences of each interior k-divisible point. This results in n parallel line segments (no two collinear, by the Lemma), each the base line of a path in the rotation class of P. Label the lines (at their endpoints) 0 through $n-1$ from top to bottom.



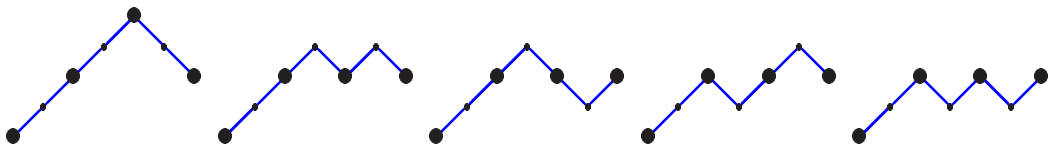
The rotation class of $P = UDU^3D^5UDUD^2 \in \mathcal{P}_{5,3}$

Now the path Q with baseline i has the form BA when P is decomposed as AB with A an initial segment of P. Hence Q is in $\mathcal{P}_{n,k}$ and has $X = i$ since the interior k-divisible points of Q lying (strictly) above its baseline are precisely those labeled $1, 2, \ldots, i-1$. The path B can be retrieved in Q as the initial subpath of Q terminating at its "lowest" k-divisible point where "lowest" is measured relative to the parallel lines, and so the mapping is invertible.

The diagram used in this proof is reminiscent of the one used in *Concrete Mathematics* [1, p. 360] to prove Raney's Lemma, also known as the Cycle Lemma [2, 3].

3 Special Case

The case $j = 1, k = 2$ gives a new interpretation of the Catalan numbers: C_n is the number of lattice paths of $n+1$ upsteps and $n-1$ downsteps such that the interior even-numbered vertices all lie strictly above the line joining the initial and terminal points. The $C_3 = 5$ paths with $n = 3$ are shown.



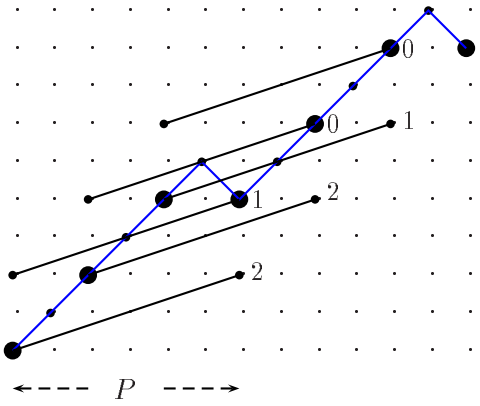
The 5 paths in $\mathcal{P}_{3,2}$

4 General Case

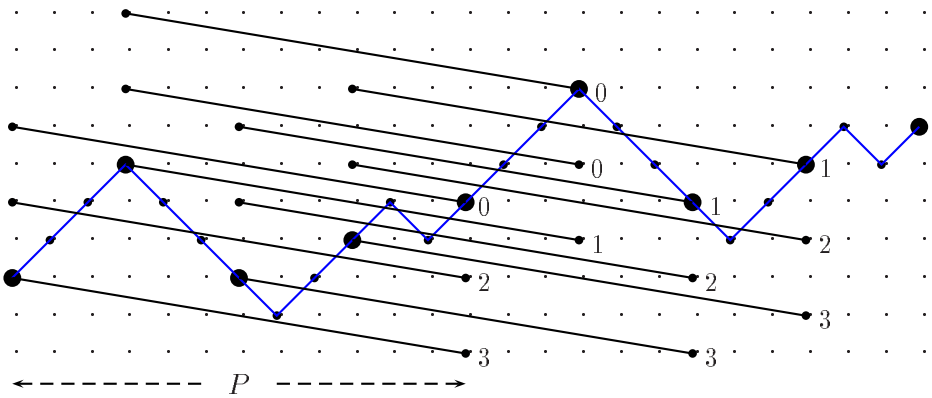
The general case $j \geq 1$ is similar but a little more complicated. Let $\mathcal{P}_{n,k,j}$ denote the set of paths of kn upsteps/downsteps of which $n+j$ are upsteps. Thus $|\mathcal{P}_{n,k,j}| = \binom{kn}{n+j}$. The "$j$" factor in the numerator of $\frac{j}{n}\binom{kn}{n+j}$ requires that we consider the Cartesian product $\mathcal{P}^*_{n,k,j} := \mathcal{P}_{n,k,j} \times [j]$ whose size is $j\binom{kn}{n+j}$. Given $(P, i) \in \mathcal{P}^*_{n,k,j}$, introduce an x-y coordinate system with origin at the initial point of P, identify the parameter i with the line segment joining $(0, 2(i-1))$ and $(kn, -(k-2)n + 2i)$, and call this the baseline for (P, i); it coincides with the previous notion of baseline when $j = 1$, forcing $i = 1$. It is easy to see that, once again, the baseline never contains an interior k-divisible point of P. Define X on $(P, i) \in \mathcal{P}^*_{n,k,j}$ by $X = \#$ interior k-divisible points of P lying strictly above the baseline.

We first show that X is uniformly distributed over $0, 1, 2, \ldots, n-1$. It is no longer true that orbits in $\mathcal{P}_{n,k,j}$ under the "rotate left by k" operator R all have size n but no matter: in general, $P \in \mathcal{P}_{n,k,j}$ uniquely has the form P_1^r with P_1 of length divisible by k and r maximal. Then r necessarily divides n and j, and the orbit of P under R has size n/r. In case $r \geq 1$,

everything will merely be cut down by a factor of r. Declare two elements (P_1, i_1) and (P_2, i_2) to be *rotation-equivalent* if P_1 and P_2 are in the same rotation class under R (regardless of i_1 and i_2). As before, all elements of a rotation-equivalence class can be seen in a single diagram as illustrated.



The rotation-equivalence class in $\mathcal{P}^*_{3,2,2}$ of $P = U^5D$



The rotation-equivalence class in $\mathcal{P}^*_{4,3,3}$ of $P = U^3D^4U^3DU$

Label the baselines (there are jn/r of them; both illustrations have $r = 1$) at their endpoints as follows (each of $0, 1, \ldots, n-1$ will be the label on j/r endpoints). First take the highest endpoint p and consider the set of all endpoints lying weakly to the left of the vertical line

through p. Since there are $j-1$ endpoints directly below p, this set has size at least j. Place label 0 on the j/r highest points in this set, favoring points to the left if a choice must be made between points at the same height. Then take the highest unlabeled endpoint, consider the set of all unlabeled endpoints lying weakly to the left of its vertical line, and place the label 1 on the j/r highest points in this set, again favoring "left". Continue in like manner until all endpoints are labeled.

Then, for each $i = 0, 1, \ldots, n-1$,the j/r objects in the rotation-equivalence class with label i all have $X = i$, and the uniform distribution of X follows. By considering the objects in $\mathcal{P}^*_{n,k,j}$ with $X = n-1$, we obtain our main result.

Main Theorem. *Suppose $j \geq 1$, $k \geq 2$, and $n \geq \frac{j}{k-1}$. Then $\frac{j}{n}\binom{kn}{n+j}$ is the number of lattice paths of $n+j$ upsteps $(1,1)$ and $kn-(n+j)$ downsteps $(1,-1)$ which (i) start at $(0,-2i)$ for some $i \geq 0$, and (ii) have all interior k-divisible points (strictly) above the line through the origin of slope $-\frac{(k-2)n-2}{kn}$.*

5 Concluding Remarks

The main theorem can be generalized somewhat further (essentially the same proof): $\frac{d}{n}\binom{an}{cn+d}$ is the number of lattice paths of $cn+d$ upsteps and $an-(cn+d)$ downsteps which (i) start at $(0,-2i)$ for some $i \geq 0$, and (ii) have all interior k-divisible points (strictly) above the line through the origin of slope $-\frac{(a-2c)n-2}{an}$.

There is also a well known generalization of the Catalan numbers in a different direction: $\frac{j}{kn+j}\binom{kn+j}{n}$ is the number of lattice paths of n steps east (1,0) and $(k-1)n+j-1$ steps north $(0,1)$ that start at the origin and lie weakly above the line $y = (k-1)x$. One way to prove this (slightly generalizing the approach in [4]) is as follows. Consider the set $\mathcal{P}_{n,k,j}$ of paths consisting of n steps east and $(k-1)n+j$ steps north. Measuring "height" of a point above $y = (k-1)x$ as the perpendicular distance to $y = (k-1)x$, define j *high points* for a path $P \in \mathcal{P}_{n,k,j}$: the first high point is the leftmost of the highest points on the path, the second high point is the leftmost of the next highest points of the path, and so on. Note that all j high points necessarily lie strictly above $y = (k-1)x$. Mark any one of the these

high points to obtain the set $\mathcal{P}^*_{n,k,j}$ of marked $\mathcal{P}_{n,k,j}$-paths. Clearly, $|\mathcal{P}^*_{n,k,j}| = j\binom{kn+j}{n}$. Label the $kn+j+1$ points on a marked path $P^* \in \mathcal{P}^*_{n,k,j}$ in order $0, 1, 2, \ldots, kn+j$ starting at the origin. Set $X =$ label of the marked high point. Then X is uniformly distributed over $1, 2, \ldots, kn+j$. The paths with $X = kn+j$ yield the desired paths by deleting the last step (necessarily a north step) and rotating 180°.

All the above generalizations of the Catalan numbers are incorporated in the expression

$$\frac{ad-bc}{an+b}\binom{an+b}{cn+d} = (a-c)\binom{an+b-1}{cn+d-1} - c\binom{an+b-1}{cn+d}$$

and it would be interesting to find a unified combinatorial interpretation for it.

References

[1] R. E. Graham, D. E. Knuth, Oren Patashnik, *Concrete Mathematics* (2nd edition), Addison-Wesley, 1994.

[2] N. Dershowitz and S. Zaks, The cycle lemma and some applications, *European J. of Comb.* **11**, 1990, 35–40.

[3] H. S. Snevily and D. B. West, The bricklayer problem and the strong cycle lemma, *Amer. Math. Monthly* **105**, 1998, 131–143.

[4] Wen-jin Woan, Uniform partitions of lattice paths and Chung-Feller generalizations. *Amer. Math. Monthly* 108 (2001), no. 6, 556–559.

2000 *Mathematics Subject Classification*: 05A15.

Keywords: Catalan, uniformly distributed, k-divisible, baseline, Cycle Lemma.