Two Uniformly Distributed Parameters Defining Catalan Numbers

David Callan

A *path* is a finite sequence of ± 1's with a graphical representation as a sequence of contiguous steps of slope $+1$ (upsteps) and -1 (downsteps). For example, the path $w = (1, -1, -1, 1, -1, 1, 1, -1)$ is pictured in Figure 1.

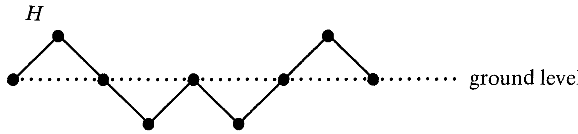

Figure 1

Let $\mathscr{P}_n$ denote the set of $\binom{2n}{n}$ paths consisting of n upsteps and n downsteps. Each path in $\mathscr{P}_n$ starts and terminates at "ground level" as in Figure 1. There is a well known parameter (statistic) on $\mathscr{P}_n$ that we will call *northcnt* (to suggest a count north of a baseline). For $w \in \mathscr{P}_n$, northcnt(w) is the number of w's n upsteps that lie above ground level. Thus northcnt = 2 in Figure 1, and as w ranges over $\mathscr{P}_n$ northcnt has possible values 0 through n. The paths for which northcnt = n—that is, the paths that lie entirely at or above ground level—we call *Catalan* paths. Dually, we call the paths with northcnt = 0 *inverted Catalan* paths: reflection in ground level gives a bijection between the two classes. It is a famous fact that exactly $1/(n+1)$ of the paths in $\mathscr{P}_n$ are Catalan: they are counted by the Catalan number $\frac{1}{n+1}\binom{2n}{n}$. A combinatorially satisfying way to see this is via the Chung-Feller Theorem, which asserts that the parameter northcnt is in fact *uniformly* distributed on $[0, n]$. This partitions $\mathscr{P}_n$ into $n+1$ equal-size classes, one of which consists of the Catalan paths. For combinatorial proofs of the Chung-Feller Theorem, see **[1]**, **[2]**, **[3]**, or **[4]**.

Curiously, there is another parameter on $\mathscr{P}_n$, *westcnt*, that serves the same purpose: it is also uniformly distributed on $[0, n]$ and it has a constant value on the set of inverted Catalan paths. To define westcnt(w), let H denote the highest point of w, taking the leftmost one if there is more than one highest point as in Figure 1. Then westcnt(w) is the number of w's n upsteps that lie to the left (west) of H. Thus the path in Figure 1 has westcnt = 1, and westcnt = 0 precisely for the inverted Catalan paths. The parameter westcnt is implicit in **[5]**.

One could show directly that westcnt is uniformly distributed on $[0, n]$. This is essentially done in **[5]**, modulo translation from bracket sequences to lattice paths. But that still leaves open the question, why? Can one "explain" why northcnt and westcnt are equidistributed? A satisfactory answer would consist of a "nice" bijection $\phi : \mathscr{P}_n \to \mathscr{P}_n$ such that westcnt(w) = northcnt($\phi(w)$) for all $w \in \mathscr{P}_n$. Here we give a simple such bijection.

To define ϕ, first observe that every path in $\mathscr{P}_n$ can be uniquely decomposed as in Figure 2 where the C_i and D_i are inverted Catalan paths (possibly empty), lying

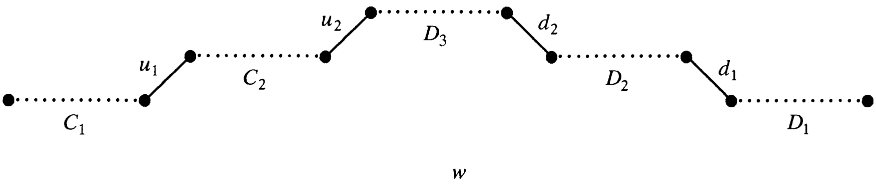

Figure 2

below the dotted segments. Each u_i is an upstep and each d_i is a downstep. There will be k C's and $k+1$ D's for some $k \geq 0$; in the illustration, $k = 2$. To see uniqueness, imagine the space above ground level divided into horizontal strips as indicated by the dotted lines (extended) in Figure 2. Then u_i, d_i are respectively the leftmost upstep and rightmost downstep in the ith strip above ground level.

The path $\phi(w)$ is given by flipping over each C_i path so it becomes a Catalan path C_i' and then rearranging components as in Figure 3. Note that since H (the

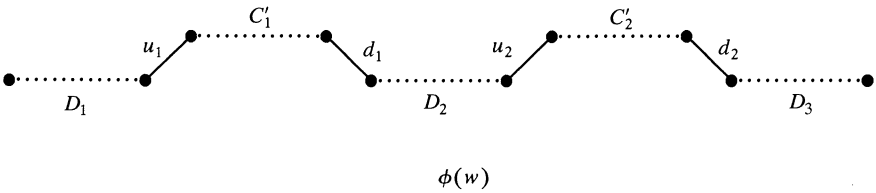

Figure 3

leftmost high point) is the northeast tip of of u_k (of u_2 in Figure 2)

$$\text{westcnt}(w) = \#\ u\text{'s} + \text{total}\ \#\ \text{upsteps in the}\ C_i.$$

Also,

$$\text{northcnt}(\phi(w)) = \#\ u\text{'s} + \text{total}\ \#\ \text{upsteps in the}\ C_i'.$$

However, for each i, # upsteps in C_i = # downsteps in C_i = # upsteps in C_i', and hence westcnt(w) = northcnt($\phi(w)$), as desired.

Finally, to show ϕ is a bijection, we must check reversibility: can the u_i, d_i, C_i', D_i as in Figure 3 be retrieved uniquely from each path in $\mathscr{P}_n$? Yes: consider the first horizontal strip above ground level. Traversing this strip left to right, upsteps and downsteps are encountered alternately. These determine the u_i and d_i (if any). The connecting paths (possibly empty) determine the C_i' and D_i in order. We are done.

REFERENCES

1. M. D. Atkinson and J.-R. Sack, Generating binary trees at random, *Inform. Process Lett.* **41** (1992) 21–23.
2. D. Callan, Pair them up!: A visual approach to the Chung-Feller theorem, *College Math. J.* **26** (1995) 196–198.
3. H. M. Finucan, *Proc. Fourth Australian Conf., Univ. Adelaide, Adelaide, Some elementary aspects of the Catalan numbers, Lecture Notes in Math., Vol. 560*, Springer, Berlin 1976, 41–45.
4. T. V. Narayana, Cyclic permutation of lattice paths and the Chung-Feller theorem, *Skandinavisk Aktuarietidskrift* **50** (1967) 23–30.
5. D. Rubinstein, Catalan numbers revisited, *J. Combin. Theory Ser. A* **68** (1994) 486–490.

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