Pair Them Up! A Visual Approach to the Chung-Feller Theorem

David Callan


Stable URL:
http://links.jstor.org/sici?sici=0746-8342%28199505%2926%3A3%3C196%3APTUA%3E2.0.CO%3B2-3

*The College Mathematics Journal* is currently published by Mathematical Association of America.
The nicest way to show that two finite sets of objects have the same number of elements is to exhibit a bijection between them. Finding such a bijection may not be easy—usually the historical order is a "counting" proof first and a "bijective" proof later—but the bijection, once found, is often (always?) pretty. A case in point is the celebrated Chung-Feller theorem of 1949. It arose in the context of coin-tossing games.

Suppose Jane tosses a fair coin repeatedly and for each head (H) she wins $1, but for each tail (T) she loses $1. The record of such a game is a sequence of H's and T's, such as HHTHTTHHTT, which can be represented graphically as a zigzag path (Figure 1): each upstep > denotes a win (H), each downstep < a loss (T). Let us suppose that Jane finishes with a net gain of $0, that is, she finishes even (as in Figure 1). If \( X \) is the number of upsteps that lie above the horizontal dashed line AB, then \( 2X \) (the total number of steps that lie above AB) records the number of tosses during which Jane was weakly ahead, that is, ahead or at least even—both before and after the toss. In Figure 1, \( X = 4 \). The Chung-Feller theorem addresses the question (a conditional probability problem): Given that Jane finishes even, what is the distribution of \( X \)? Is Jane more likely to be (weakly) ahead frequently, just occasionally, or very rarely? Of course, finishing even necessitates an even number of tosses, say \( 2n \), and then \( X \) may = 0, 1, 2, \ldots, \( n \).

The remarkable fact is that, given that Jane finishes even, \( X \) is *uniformly* distributed on these \( n + 1 \) values. What this amounts to is that the number of zigzag paths as in Figure 1 (\( 2n \) steps, even finish) with \( k \) upsteps above the dashed line is the same for all \( k = 0, 1, 2, \ldots, n \). Let \( S_k \) denote the set of paths for which \( X = k \). According to the theorem, \( |S_0| = |S_1| = \cdots = |S_n| \). For example, when
Figure 2

For general $n$ there are a total of $\binom{2n}{n}$ paths that finish even (each path has $n$ upsteps, $n$ downsteps and you get to choose where the upsteps go); hence by the theorem, $|S_k| = \binom{2n}{n}/(n+1)$—the ubiquitous Catalan numbers arise again! (As a byproduct, we have a nice combinatorial explanation of why $n+1$ divides $\binom{2n}{n}$.) For a lively discussion of this divisibility relation, see [6].

This theorem was first proved using generating functions in [2], and Feller's classic book [3] contains an interesting account of its relation to Galton's rank-order test. (It shows that Galton's test wasn't much good.) The theorem was subsequently treated by more combinatorial methods in [4] (using inductively defined mappings), and in [1] and [5] (using an explicit mapping—essentially the same one in both papers). Still, none gave explicit bijections between the sets $S_k$—the most direct and combinatorially satisfying way of proving the theorem. Here we will adapt the map of [1] and [5] to describe an easily visualized bijection between any pair $S_i, S_j$ (ordered so that $i < j$). For simplicity of exposition let's take $i = 1, j = 4$, but it will be clear the method is quite general.

So suppose we have a path in $S_1$ (Figure 3). We must produce from this a path in $S_4$ in a "reversible" way. Here's how. First draw horizontal lines dividing the steps below the dashed line into strips as in Figure 3. Now scan the strips right to left and top to bottom, numbering the upsteps. Note that upsteps and downsteps alternate with an upstep first and a downstep last. Mark the first downstep after upstep 3 (in general, after upstep $j-i$) with a heavy line (link) as in Figure 3. (There will be $n-i$ numbered upsteps and certainly $n-i \geq j-i$, so this procedure is always possible.) It divides the path into two parts with a link between them. (The left-hand part might be vacuous, but that's okay.) Now switch the two parts. That is, move the link to the right end of the right-hand part, attach the left-hand part to it, and move the dashed line down to the level of the left end of

Figure 3
the new path. Bingo! That’s it! The path in Figure 4 results. The numbering of the upsteps has been retained and you will see that upsteps in the shaded region of Figure 3 reappear above the dashed line in Figure 4 while no other upstep changes sides of the dashed line. (This holds in general—look separately at each of the two parts to visualize where the new path’s dashed line will lie.)

Conversely, given a path in $S_4$ (say the one in Figure 5), scan the strips above the dashed line, left to right and bottom to top; use the first downstep after upstep $j - i$ as the link; then switch the two parts. This will reverse the original process because it locates the original link. That’s the bijection, and the Chung-Feller theorem follows.


References