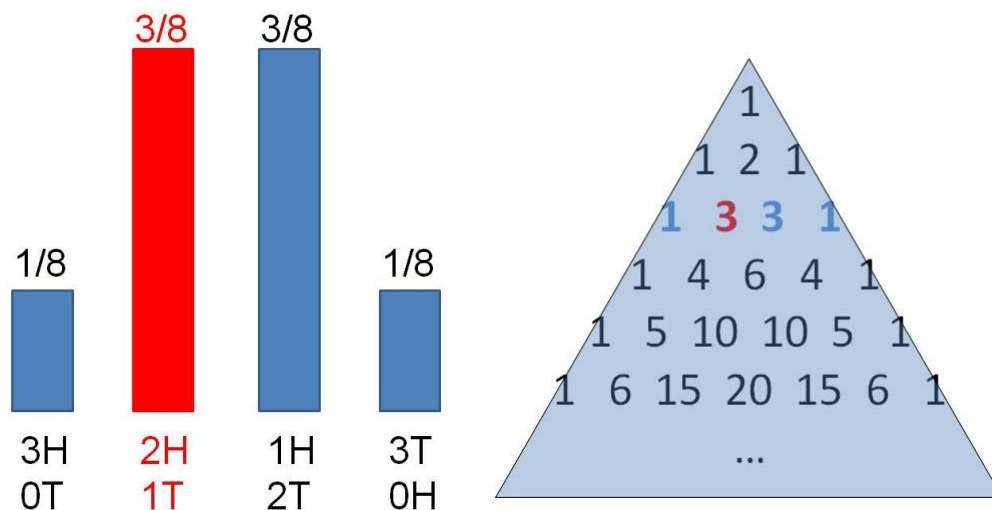**Patterns in Probability: How to see binomial statistics**

International Year of Statistics, June 10, 2013, John K. White, author of *Do The Math!*

We've all come across *either-or* decisions in our lives, like the proverbial fork in the road: left-right, head-tails, red-black. The likelihood of a series of such decisions underpins the basics of probability, from gambling odds to stock markets, from weather simulations to a host of new social networking interactions. But as we celebrate the International Year of Statistics, can we make the statistics more visual? Can we see the underlying patterns in such *binomial* probabilities and in so doing make the math less daunting?
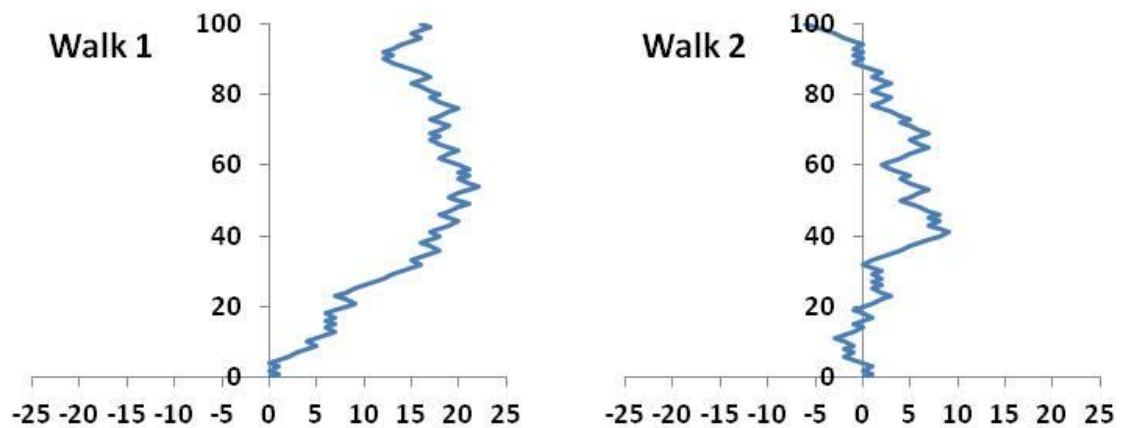
For small multiples, we can construct a sample set, *e.g.*, for two heads in three coin flips, where there are eight possible outcomes (**HHH**, **HHT**, **HTH**, **HTT**, **THH**, **THT**, **TTH**, **TTT**), three of which have two heads and thus a 3/8 chance. A picture (or frequency distribution) of the possible outcomes shows the odds.

Two heads and one tail (2H, 1T) in three coin tosses (frequency distribution and Pascal's triangle)

With larger multiples, however, this simple calculation method can become unwieldy. Another pictorial method, Pascal's triangle, helps, named for the French mathematician Blaise Pascal who was asked to work out payouts for a friend after a gambling night unexpectedly ended, and from which the famous normal curve was later devised to work out higher multiples. But of course one still has to do the math, which although a bit hairy does the job.*
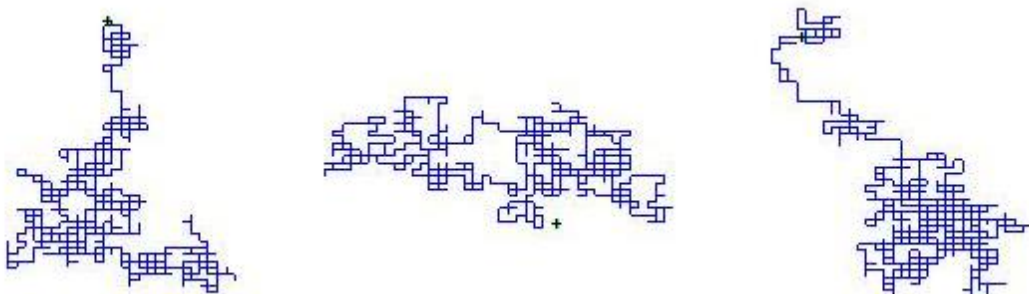
Random walks help to visualize the statistics. For example, a random walker takes 100 steps left or right. What are the odds that the walker is back at the start, that is, in the middle of the distribution? This is the same as asking the odds of fifty heads in one hundred flips or fifty stock decreases in one hundred ticks. With such pictures, we can see that the walker will most likely be back at home, but not always.
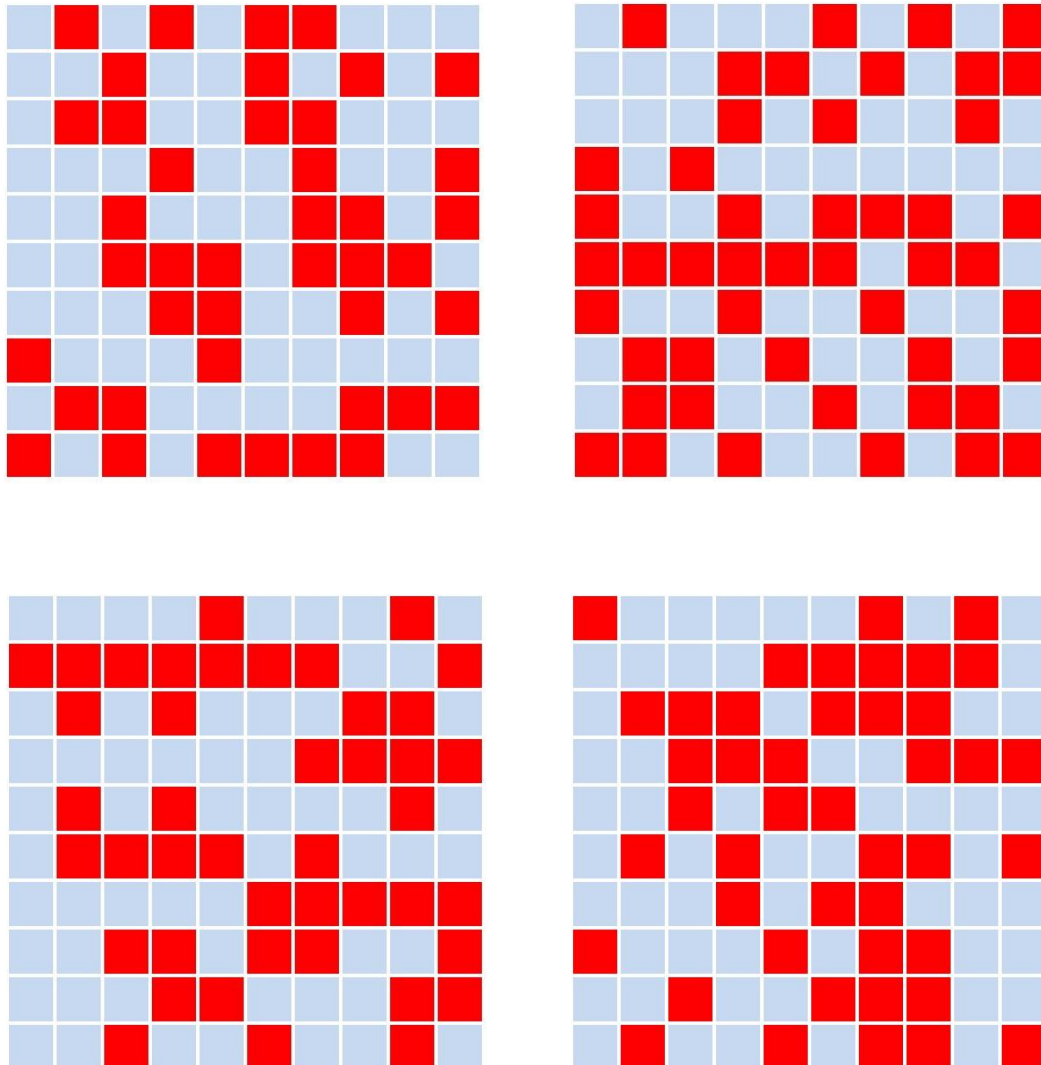

100-step 1D random walks

The same kind of analysis helped to determine that would-be French soldiers lied about their heights to escape Napoleon's army since there were more short men than expected. Similar analysis revealed a point-spread shaving scam in U.S. college basketball games, all because the variation wasn't varied enough. The stock market collapse of 2008 was also tied to repeated losses in an over-leveraged derivatives market, a supposed statistical blip that was meant to happen only once in a blue moon.

A 2D version of the random walk illustrates many real-world scenes, such as clustering, urban growth, animal foraging, and even weather simulations.

1000-step 2D random walks

But can we simplify the stats even more to help understand the likelihood of multiple binomial events? Here, we turn to abstract art, using two colours randomly selected in a 10 x 10 square grid.

Random 2-colour 10 x 10 grids

As we can see, patterns appear, yet there are no 10-0 rows or columns, which we would expect about once every fifty pictures. Count how many 5-5 splits across any row or down any column, the expected maximum or mode. Or count the 4-6, 6-4, or any *m-n* (*m* + *n* = 10) splits. In effect, you are counting the sample set, working out the binomial odds, *seeing* the stats.

We must ever guard against mistaking unlikely events with probable events. Seeing the statistics helps.

John K. White is an adjunct lecturer in the School of Physics, University College Dublin and author of [Do The Math!: On Growth, Greed, and Strategic Thinking](#) (Sage, 2013)

To make your own abstract binomial art [Do The Math! Math and Randomness](#).

* Binomial and normal distributions:

$$(p + q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}$$

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$