PROBABILITY VIA THE *N*TH ORDER FIBONACCI-*T* SEQUENCE

STEPHEN JOHN TURNER
St. Mary's University, Halifax, Nova Scotia B3H3C3

Suppose we repeat a Bernoulli (p) experiment until a success appears twice in a row. What is the probability that it will take exactly four trials when $p = .5$? Answer: There are 2^4 equi-probable sequences of trial outcomes. Of these, there are exactly two with their last two entries labeled success with no other consecutive entries successes. Hence, there is a $1/(2^3)$ chance that the experiment will be repeated exactly four times.

Immediately, questions arise: What is the probability that it takes 5, 6, 7, ..., n trials? What are these probabilities when $p \neq .5$? What answers can be provided when we require N successes in a row?

The answers for the most general case of N successes involve a unique approach. However, it is instructive to treat the case for $N = 2$ first in order to set the framework.

THE CASE FOR $N = 2$

We shall use the idea of "category."

Definition: Category S is the set of all $S + 1$ sequences of trial outcomes (denoted in terms of s and f) such that each has its last two entries as s and no other consecutive entries are s.

Now we have a means for designating those outcome sequences of interest.

Notation: $N(S)$ denotes the number of elements in category S,

$$S = 1, 2, 3, \ldots .$$

There is but one way to observe two successes in two trials so that category one contains the one element (s,s). Also, category two contains one element (f,s,s). The value of $N(3)$ is determined by appending an f to the left of every element in category two and then an s to the left of each element in category two beginning with an f. Thus, category three has two elements:

$$(f,f,s,s) \quad \text{and} \quad (s,f,s,s).$$

Observe that this idea of "left-appending may be continued to construct the elements of category $S + 1$ from the elements of category S by appending an f on the left to each element in category S and an s on the left to each element in category S beginning with an f. There can be no elements in category $S + 1$ exclusive of those accounted for by this "left-appending" method.

A result we can observe is that

$$\begin{aligned} N(S+1) &= N(S) + \text{"the number of } S\text{-category elements that begin with an } f\text{"} \\ &= N(S) + N(S-1). \end{aligned}$$

So we obtain the amazing result that the recursion formula for category size is the same as the recursion formula for the Fibonacci sequence! Since $N(1) = N(2) = 1$, we see that when $p = .5$ the probability that it will take $S + 1$ trials to observe two successes in a row is given by

$$(N(S)/(2^{S+1})) = (F_S)/(2^{S+1})$$

where F_S denotes entry S in the Fibonacci sequence.

If $p \neq .5$, then each category element must be examined in order to count its exact number of f entries (or s entries). Such an examination is not difficult.

Suppose category $S - 1$ has a_i elements which contain exactly i entries that are f, and that category S has b_i elements which contain exactly i entries that are f, $i = 0, 1, 2, \ldots, S - 2$. Then category $S + 1$ contains exactly $a_i + b_i$ elements which contain exactly $i + 1$ entries that are f. Justification for this statement comes quickly as a benefit of the "left-appending" approach to the problem. Hence, we can construct the following partial table:

Category	Number of Elements Containing Exactly i Entries Which Are f							
$i =$	0	1	2	3	4	5	6	...
1	1	0	0	0	0	0	0	
2	0	1	0	0	0	0	0	
3	0	1	1	0	0	0	0	
4	0	0	2	1	0	0	0	
5	0	0	1	3	1	0	0	
6	0	0	0	3	4	1	0	
7	0	0	0	1	6	5	1	
⋮								

Observe that nonzero entries of the successive columns are the successive rows of the familiar Pascal triangle! This observation is particularly useful because the kth entry of the ith row in the Pascal triangle is

$$\binom{i-1}{k-1} = \frac{(i-1)!}{((i-1)-(k-1))!(k-1)!}.$$

Also, since category i contains exactly one element containing $i - 1$ entries which are f, we know the ith row of the Pascal triangle will always begin in row i and column $i - 1$ of the table. Thus, if we move along the nonzero entries of row t of the table (from left to right) we encounter the following successive numbers:

$$\binom{t-1}{0}, \binom{t-2}{1}, \binom{t-3}{2}, \ldots, \binom{a}{b}.$$

To characterize $\binom{a}{b}$, notice that row k of the Pascal triangle ends in row $2k - 1$ of the table. Thus, if $t > 1$ is odd, then $a = b = (t - 1)/2$. And if $t > 1$ is even, then $a = t/2$ and $b = (t/2) - 1$.

Thus, whenever $t > 1$, we know that the probability that "it takes $t + 1$ trials" is given by

$$\sum_{i=0}^{(t-X)/2} \binom{t-(i+1)}{i} (1-p)^{t-(i+1)} p^{(t+1)-(t-(i+1))},$$

where $X = \begin{cases} 1, & \text{if } t \text{ is odd} \\ 2, & \text{if } t \text{ is even} \end{cases}$

THE GENERAL CASE

Now we will be answering the question of the probability that it takes k trials to observe n successes in a row, $k \geq n$. To begin, we generalize the concepts of category, Fibonacci sequence, and Pascal triangle.

Definition: Category x is the set of all $n + (x - 1)$ sequences of f's and s's (denoting failure and success, respectively) such that the last n entries in each sequence are s, and no other n consecutive entries in the sequence are s.

Definition: The nth order Fibonacci-T sequence, denoted f^n, is the sequence $a_1, a_2, a_3, \ldots, a_i, \ldots$, where $a_1 = 1$ and

$$a_i = \begin{cases} \displaystyle\sum_{k=1}^{(i-1)} a_k, & \text{if } 2 \leq i \leq n \\ \displaystyle\sum_{k=i-n}^{(i-1)} a_k, & \text{if } i > n. \end{cases}$$

It is instructive to first define the nth order Pascal-T triangle by example:

(1) If $n = 2$, the Pascal-T triangle is the familiar Pascal triangle;

(2) If $n = 3$, the Pascal-T triangle is of the form

```
                    1
                1   1   1
            1   2   3   2   1
        1   3   6   7   6   3   1
    1   4  10  16  19  16  10   4   1
1   5  15  30  45  51  45  30  15   5   1
```

(3) If $n = 4$, the Pascal-T triangle is of the form

```
                   1
             1   1   1   1
       1   2   3   4   3   2   1
1   3   6  10  12  12  10   6   3   1
```

The nth order Pascal-T triangle has $(j - 1)n - (j - 2)$ entries in the jth row. Letting the first to last of these be denoted by $j_1, j_2, j_3, \ldots, j_{(j-1)n-(j-2)}$, the kth entry in row $j + 1$ is given by

$$\sum_{i=\max(1,k-n+1)}^{\min(k,(j-1)n-(j-2))} j_i \quad \text{for } k = 1, 2, 3, \ldots, jn - (j - 1).$$

We can now proceed by enlisting the "left-appending" procedure outlined earlier. There is but one way to observe n successes in n trials. So $N(1) = 1$. Likewise, there is but one element in category two. To obtain the elements of category three, we append an f to the left of each element in category two and then append an s to the left of each element in category two. So category three contains the two elements

$$(f, f, s, s, \ldots, s) \quad \text{and} \quad (s, f, s, s, \ldots, s)$$

where $s, s, \ldots, s$ signifies that the entry s occurs n times in succession. We may proceed in this manner for each category k, $k \leq n + 1$.

It is clear that category $n + 1$ will contain exactly one element which has the entry s in its first $n - 1$ positions. Thus, category $n + 2$ will have $2(N(n + 1)) - 1$ elements.

Now note that when constructing category $k + n$, we proceed by appending an f to the left of each element in category $(k + n) - 1$ and an s to the left of each element in category $(k + n) - 1$ which does not begin with the entry s in its first $n - 1$ positions. But the number of elements in category $(k + n) - 1$ containing the entry s in their first $n - 1$ positions is the same as the number of elements in category k which begin with an f! Hence,

$$\begin{aligned} N(n + k) &= 2(N(n + k - 1)) - \text{"number of elements in category } k \text{ which begin with an } f\text{"} \\ &= 2N(n + k - 1) - N(k - 1). \end{aligned}$$

We now prove the following useful

Theorem:

$$N(n + k) = \sum_{i=1}^{n} N(n + k - i), \quad k = 1, 2, 3, \ldots$$

Proof: We use simple induction.

(1)
$$\begin{aligned} N(n + 1) = 2^{n-1} &= 1 + \sum_{i=2}^{n} 2^{i-2} \\ &= N(1) + (N(2) + N(3) + N(4) + \cdots + N(n)) \\ &= \sum_{i=1}^{n} N(n + 1 - i). \end{aligned}$$

(2) Supposing truth for the case k, we have

$$\begin{aligned} N(n + k) + 1 = 2N(n + k) - N(k) &= 2\sum_{i=1}^{n} N(n + k - i) - N(k) \\ &= \sum_{i=1}^{n-1} N(n + k - i) + \sum_{i=1}^{n} N(n + k - i) \\ &= \sum_{i=1}^{n-1} N(n + k - i) + N(n + k) \\ &= \sum_{i=1}^{n} N[(n + k) + 1 - i]. \blacksquare \end{aligned}$$

Now note that since $N(1) = 1$, the sequence $N(1)$, $N(2)$, $N(3)$, ... is an nth order Fibonacci-T sequence via the theorem!

Thus, if $p = .5$, then the probability that it will take $n + (k - 1)$ trials to observe n successes in a row, $k \geq 1$, is given by

$$N(k)/(2^{n+k-1}) = (f_k^n)/(2^{n+k-1}),$$

where f_k^n denotes the kth entry in the nth order Fibonacci-T sequence.

We will now determine the probabilities when $p \neq .5$. A foundation is set by observing that if category $k - n + i$ has an element M which has exactly x entries that are f, then the element $(s, s, \ldots, s, f, M)$, beginning with $n - (i + 1)$ entries which are s, is a member of category k and it contains $x + 1$ entries that are f. This is true for $i = 0, 1, 2, \ldots, n - 1$. If we let a_i, $i = 0, 1, 2, \ldots, n - 1$ represent the number of elements in category $k - n + 1$ which have x elements that are f, then category k contains $a_0 + a_1 + a_2 + \cdots + a_{n-1}$ elements which have $x + 1$ entries that are f. This is the recursive building block for the nth order Pascal-T triangle where row i begins in category i and ends in category $(i - 1)n + 2$! The following table partially displays the situation.

Category	Number of Elements Containing Exactly i Entries Which Are f				
	$i =$ 0	1	2	3	...
1	1	0	0	0	
2	0	1	0	0	
3	0	1	1	0	
4	0	1	2	1	
5	0	1	3	3	
⋮	⋮	⋮			
$n + 1$	0	1			
$n + 2$	0	1	⋮		
$n + 3$	0	0			
⋮	⋮	⋮		⋮	
$2n$	0	0	3		
$2n + 1$	0	0	2		
$2n + 2$	0	0	1		
$2n + 3$	0	0	0		
⋮	⋮	⋮	⋮		
$3n + 1$	0	0	0	3	
$3n + 2$	0	0	0	1	
$3n + 3$	0	0	0	0	
⋮	⋮	⋮	⋮	⋮	

Since the number of entries in two successive rows of the nth order Pascal-T triangle always differ by n, then moving from left to right in the table, the ith category row will see its first nonzero entry in column $m - 1$ where $(m - 2)n + 2 \leq i \leq (m - 1)n + 1$, $i > 1$ and $m > 1$.

Let $\begin{bmatrix} i \\ k \end{bmatrix}_n$ denote the kth entry in the ith row of the nth order Pascal-T triangle, $k = 1, 2, 3, \ldots, (i - 1)n - (i - 2)$. Suppose $i \geq 2$. Then the successive nonzero entries in the ith category row, listing from right to left are

$$\begin{bmatrix} i \\ 1 \end{bmatrix}_n, \begin{bmatrix} i-1 \\ 2 \end{bmatrix}_n, \begin{bmatrix} i-2 \\ 3 \end{bmatrix}_n, \ldots, \begin{bmatrix} i-(i-m) \\ (i-m)+1 \end{bmatrix}_n$$

where $(m - 2)n + 2 \leq i \leq (m - 1)n + 1$ for some $m \geq 2$.

Thus, the probability that "it will take $n + (i - 1)$ trials," $i \geq 2$, is given by

$$\sum_{k=0}^{i-m} \begin{bmatrix} i-k \\ k+1 \end{bmatrix} (1 - p)^{(i-k)-1} p^{n+(i-1)-((i-k)-1)}$$

where $(m - 2)n + 2 \leq i \leq (m - 1)n + 1$ for some $m \geq 2$.

AUTHOR'S NOTE

The machinery used in the above solution generates a number of ideas which the reader may wish to explore. A few examples are:

1. If f_k^2 denotes the kth entry in the second order Fibonacci-T sequence, then it can be shown that the sequence $\left\{f_{k+1}^2/f_k^2\right\}$ is a Cauchy sequence and so being, has a limit g_2. From this, it follows that $g_2 = 1 + 1/g_2$ so that g_2 is the golden ratio. This brings up the question of the identity of $g_n = \lim_{k\to\infty} f_{k+1}^n/f_k^n$ when $n \geq 3$. (Here, f_k^n denotes the kth entry in the nth order Fibonacci-T sequence.) It can be argued that $g_n < 2$ for any value of n and $\lim_{n\to\infty} g_n = 2$.
2. It has been shown that
$$f_k^2 = [(g_2)^k - (-g_2)^{-k}]/[g_2 + (g_2)^{-1}].$$
Can we find a similar expression for f_k^n when $n \geq 3$?
3. We can generalize the nth order Fibonacci-T sequence by specifying the first n entries arbitrarily. For instance, the first three cases would be
$$n = 1:\quad a, a, a, a, a, a, \ldots;$$
$$n = 2:\quad a, b, a+b, a+2b, 2a+3b, 3a+5b, \ldots;$$
$$n = 3:\quad a, b, c, a+b+c, a+2(b+c), 2a+3(b+c)+c, \ldots,$$
where a, b, and c are arbitrarily chosen. The investigation of the properties and relationships between these generalized sequences could provide some interesting results.

REFERENCE

H. S. M. Coxeter, *Introduction to Geometry*, 1969, pp. 166, 167.

*****