First Occurrences of Particular Pairs in a Sequence of Independent Flips of a Fair Coin [1]

Charles S. Peskin
Courant Institute of Mathematical Sciences, New York University
May, 2025

In a sequence of independent flips of a fair coin, we denote the outcome of each flip by H or T, and we study the statistics of the first occurrence of particular pairs of adjacent outcomes, such as HT and HH, and also the probability that one particular pair comes before another one.

We begin with the particular pair HT. Let $s_n(\mathrm{HT})$ be the number of sequences of length n with HT as the terminal pair, and with no other occurrence of HT in the sequence. Every such sequence is uniquely determined by its first occurrence of H, since every outcome before that first H (if any) has to be T, and every outcome between the first H and the terminal T (if any) has to be H. Since there are $n-1$ possible locations for the first H in a sequence of length n that ends in T, it follows that

$$s_n(\mathrm{HT}) = n - 1. \tag{1}$$

Moreover, since there are 2^n equally likely sequences of length n, and since $n-1$ of them have the first occurrence of HT at the end, the probability of this outcome is given by

$$p_n(\mathrm{HT}) = \frac{n-1}{2^n}. \tag{2}$$

Next we consider the particular pair HH, for which the situation is somewhat more complicated, and we need to use a recursive argument. For $C = \mathrm{H}$ or T, let $s_n(C, \mathrm{HH})$ be the number of sequences of length n that begin with C, end with HH, and have no other instances of HH within the sequence besides the terminal HH. Now it is easy to see that

$$s_{n+1}(\mathrm{H}, \mathrm{HH}) = s_n(\mathrm{T}, \mathrm{HH}), \tag{3}$$

$$s_{n+1}(\mathrm{T}, \mathrm{HH}) = s_n(\mathrm{T}, \mathrm{HH}) + s_n(\mathrm{H}, \mathrm{HH}). \tag{4}$$

This is because sequences of length $n+1$ with the first occurence of HH at the end can be generated by pre-pending an H or a T to the beginning of such a sequence

[1] with thanks to Percy Deift for the lunchtime discussion that led to this work.

of length n, but with the restriction that only a T can be used if the sequence of length n begins with an H. Thus every allowed sequence of length $n+1$ that begins with an H has that H followed by an allowed sequence of length n that begins with a T (hence equation (3)), but every allowed sequence of length $n+1$ that begins with a T can have that T followed by two types of allowed sequences of lengh n: those that begin with T and those that begin with an H (hence equation (4)).

Now let

$$s_n(\mathrm{HH}) = s_n(\mathrm{H}, \mathrm{HH}) + s_n(\mathrm{T}, \mathrm{HH}) \tag{5}$$

so that $s_n(\mathrm{HH})$ is the number of sequences of length n that terminate in HH and have no other occurrence of HH within the sequence. By adding equations (3) and (4) we get

$$s_{n+1}(\mathrm{HH}) = s_n(\mathrm{HH}) + s_n(\mathrm{T}, \mathrm{HH}), \tag{6}$$

and by making use of equation (4) with n lowered by one we can rewrite this as

$$s_{n+1}(\mathrm{HH}) = s_n(\mathrm{HH}) + s_{n-1}(\mathrm{HH}), \tag{7}$$

which is the recursion relation of a Fibonacci sequence!

For initial conditions, we have

$$s_1(\mathrm{HH}) = 0, \;\; s_2(\mathrm{HH}) = 1, \tag{8}$$

and it follows that

$$s_n(\mathrm{HH}) \;=\; F(n-1), \tag{9}$$

$$p_n(\mathrm{HH}) \;=\; \frac{F(n-1)}{2^n}, \tag{10}$$

where F denotes the Fibonacci sequence defined by

$$F(0) = 0, \;\; F(1) = 1, \;\; F(n+2) = F(n+1) + F(n), \tag{11}$$

and where $p_n(\mathrm{HH})$ is the probability that the first occurrence of HH is completed on the n^{th} flip of the coin.

By symmetry, the results derived above for HT are also applicable to TH, and the results derived above for HH are also applicable to TT. Thus, we have found the probability distributions for the positions of the first occurrences of HH, HT, TT, and TH.

The first 10 values of $p_n(\mathrm{HT})$ and of $p_n(\mathrm{HH})$ are shown in the following table:

Probability of First Occurrence at $(n-1, n)$		
n	$p_n(\text{HT}) = \dfrac{n-1}{2^n}$	$p_n(\text{HH}) = \dfrac{F(n-1)}{2^n}$
1	0	0
2	1 / 4	1 / 4
3	2 / 8	1 / 8
4	3 / 16	2 / 16
5	4 / 32	3 / 32
6	5 / 64	5 / 64
7	6 / 128	8 / 128
8	7 / 256	13 / 256
9	8 / 512	21 / 512
10	9 / 1024	34 / 1024

Let N_{HT} be the random index of the position in the sequence of coin flips that completes the first appearance of HT, and similarly let N_{HH} be the random index of the position that completes the first appearance of HH. Let $\mathbb{P}$ denote probability and let $\mathbb{E}$ denote expected value. From the numbers in the table, we see that

$$\frac{1}{2} = \mathbb{P}\left[N_{\text{HT}} \leq 3\right] = \mathbb{P}\left[N_{\text{HH}} \leq 4\right]. \tag{12}$$

This would seem to indicate that HT has a tendency to occur earlier than HH. This impression is confirmed by evaluating the expected values

$$\mathbb{E}\left[N_{\text{HT}}\right] = 4, \tag{13}$$

$$\mathbb{E}\left[N_{\text{HH}}\right] = 6, \tag{14}$$

see Appendix A. Thus, it may come as a shock to learn that

$$\mathbb{P}\left[N_{\text{HT}} < N_{\text{HH}}\right] = \mathbb{P}\left[N_{\text{HH}} < N_{\text{HT}}\right] \tag{15}$$

(Since it is not possible to have $N_{\text{HT}} = N_{\text{HH}}$, equation (15) is equivalent to the statement that either side of (15) is equal to 1/2.)

The proof of (15) is very simple. Which of the two pairs HT or HH comes first is decided by the coin flip that comes immediately after the first occurrence of H.

With probability $\frac{1}{2}$, the outcome is T, and then HT has come before HH. Likewise, with probability $\frac{1}{2}$ the outcome is H, and then HH has come before HT.

Thus, if we only care about which comes first in a given sequence, the foregoing analysis about the separate first occurrences of HT and HH is somehow misleading. The reason for this is not hard to find — it is the very strong interaction (i.e., non-independence) of the random variables N_{HT} and N_{HH}. It has already been mentioned that these random variables cannot be equal, but a more subtle point is that if HH occurs before HT, this sets things up for HT to occur soon thereafter, since the very next occurrence of T is guaranteed to complete a pair HT. If HT occurs first, on the other hand, then the next occurrence of HH requires, but is not guaranteed to follow, the next occurrence of an H. Indeed, there is only a probability $\frac{1}{2}$ that the next occurence of an H will be followed by another H, and if not, then it is necessary to wait for the next H to try again, and so on. It is these late effects that occur after it has been decided whether HT or HH comes first that create the asymmetry in the probability distributions for the first occurrences of HT and HH considered separately.

Now that we know that in direct competition for which comes first HT and HH are equally likely to win, it is natural to consider other contests of that same kind. Overall, there are 6 such competitions to consider: TT vs TH, TT vs HT, TT vs HH, TH vs HT, TH vs HH, and HT vs HH. Note, however, that we can interchange H and T without making any difference, so there are only 4 competitions that require separate consideration. Of these 4, two transform into themselves under the interchange of H and T, namely HH vs TT, and TH vs HT. For these cases it is obvious by symmetry that

$$\mathbb{P}\left[N_{\mathrm{HH}} < N_{\mathrm{TT}}\right] = \mathbb{P}\left[N_{\mathrm{TT}} < N_{\mathrm{HH}}\right], \tag{16}$$
$$\mathbb{P}\left[N_{\mathrm{TH}} < N_{\mathrm{HT}}\right] = \mathbb{P}\left[N_{\mathrm{HT}} < N_{\mathrm{TH}}\right]. \tag{17}$$

Of the two remaining cases, we have already analyzed HT vs HH, and the result of this analysis is equation (15). From (15-17), we are starting to get the impression that in all of these direct competitions, either outcome is as likely as the other. This, however, is false! In direct competition between TH and HH, we have the result

$$\mathbb{P}\left[N_{\mathrm{TH}} < N_{\mathrm{HH}}\right] = 0.75, \tag{18}$$
$$\mathbb{P}\left[N_{\mathrm{HH}} < N_{\mathrm{TH}}\right] = 0.25. \tag{19}$$

To see this, note that the *only* way that HH can come before TH is if HH comes at the very beginning of the sequence. Once a T has occurred, HH cannot happen until the next H has occurred, but the first H after the first T already completes an appearance of TH! Of course, the probability that the sequence starts with HH is $\frac{1}{4}$, and this completes the explanation of the result stated in equations (18-19).

It is truly remarkable that TH defeats HH three times out of four, whereas in HT vs HH, either one is equally likely to win. To see just how remarkable this is, recall that TH and HT have exactly the same statistics in all respects, including the probability distribution of their first occurrence. Nevertheless, TH has a powerful advantage over HH that HT does not share.

Simulation

The following Matlab code simulates the coin tosses and compiles results. The code keeps track only of the latest pair of coin-flip results, and it continues each trial until all four of the pairs TT, TH, HT, and HH have appeared. For each pair, the position in the sequence of its first appearance is noted (with position defined as that of the second member of the pair, so the smallest possible position is 2). Thus, each trial generates four distinct integers and these are stored as one column of the array `N`, which has dimensions $4 \times$ `ntrials`. The row index of the array `N` is called `b_code`. It is assigned to a pair of coin-flip results by thinking of that pair as a binary number with T=0 and H=1, and then adding 1, so that `bcode` = 1,2,3, or 4.

```
ntrials = 10^6
N=zeros(4,ntrials);
for tr=1:ntrials
    b_right = (rand < 0.5);
    n=1;
    while any(N(:,tr)==0)
        b_left = b_right;
        b_right = (rand < 0.5);
        b_code = 1 + 2*b_left + b_right;
        n = n+ 1;
        if N(b_code,tr)==0
            N(b_code,tr) = n;
        end
    end
end
```

The data recorded in the array `N` can be used in three different ways. The first way is to average over trials, and this gives an empirical estimate of the expected value of each of the random variables N_{TT}, N_{TH}, N_{HT}, and N_{HH}. The results are very close to the theoretical values of 6,4,4,6:

	Nbar
TT	5.9976
TH	3.9968
HT	4.0024
HH	6.0006

The second use of the recorded data is to count how many times any particular pair occurs before another particular pair, and then to divide by the number of trials to get an estimate of the corresponding probability. The Matlab code for this is

```
wins = zeros(4,4);
for i=1:4
    for j=1:4
        wins(i,j)=sum(N(i,:) < N(j,:));
    end
end
wins=wins/ntrials;
```

Thus, `wins(i,j)` is the empirical probability that the pair with `b_code = i` will appear before the pair with `b_code = j` . The results are as follows:

	TT	TH	HT	HH
TT	0	0.50038	0.25085	0.50033
TH	0.49962	0	0.50086	0.75064
HT	0.74915	0.49915	0	0.5002
HH	0.49967	0.24936	0.4998	0

As predicted, the only contests of this kind in which the contestants do not have equal probability of winning are TH vs HH, and HT vs TT. In both of these, the heterogeneous pair has a probability 3/4 of winning over the homogeneous pair, but note that this is *not* the case for HT vs HH, and likewise *not* the case for TH vs TT.

Finally, for any chosen set of values of n we can find the empical probability that each of the four random variables N_{TT}, etc., is equal to n by counting how many times this equality occurs and then dividing by the number of trials. The code and results are as follows:

```
nmax=10
p=zeros(4,nmax);
for n = 1:nmax
    p(:,n)=sum(N==n,2);
end
p=p/ntrials
```

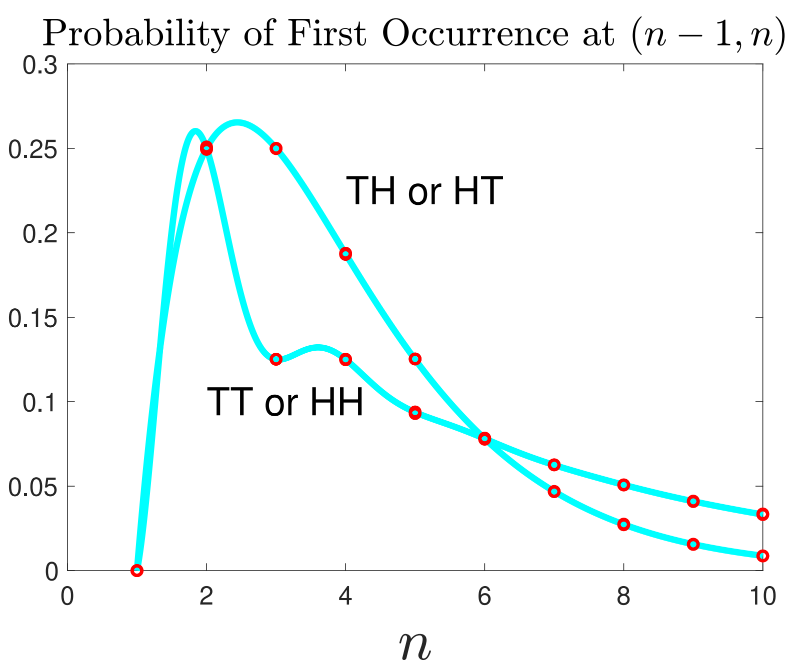

Figure 1: Red dots are the empirical probabilities based on 10^6 trials, and blue curves are the theoretical probabilities, interpolated for visual clarity.

Appendix A

The main purpose of this Appendix is to evaluate $\mathbb{E}[N_{\mathrm{HT}}]$ and $\mathbb{E}[N_{\mathrm{HH}}]$. Of course this can be done directly from the corresponding probability distributions that have been found already, but instead we will use a more interesting recursive method. The same recursive idea leads to another derivation of the probability distributions themselves, and we will give that derivation here as well.

Let N_{H} be the position of the first occurrence of H. Then $\mathbb{P}[N_{\mathrm{H}} = 1] = \frac{1}{2}$, and if $N_{\mathrm{H}} \neq 1$, then $N_{\mathrm{H}} = 1 + N'_{\mathrm{H}}$, where N'_{H} has the same probability distribution as N_{H}. Therefore,

$$\begin{aligned} \mathbb{E}[N_{\mathrm{H}}] &= \frac{1}{2}(1) + \frac{1}{2}\mathbb{E}[1 + N'_{\mathrm{H}}] \\ &= \frac{1}{2}(1) + \frac{1}{2}(1 + \mathbb{E}[N'_{\mathrm{H}}]) \\ &= \frac{1}{2}(1) + \frac{1}{2}(1 + \mathbb{E}[N_{\mathrm{H}}]) . \end{aligned} \tag{20}$$

This is easily solved for $\mathbb{E}[N_{\mathrm{H}}]$ with the result that $\mathbb{E}[N_{\mathrm{H}}] = 2$.

Next, we consider N_{HT}. Once the first H has occurred, the *next* T in the sequence (whenever it occurs) is the T that completes the *first* appearance of HT. Thus, the random variable $N_{\mathrm{HT}} - N_{\mathrm{H}}$ has the same probability distribution as N_{H}. Therefore

$$\mathbb{E}[N_{\mathrm{HT}} - N_{\mathrm{H}}] = \mathbb{E}[N_H] = 2, \tag{21}$$

from which it follows that $\mathbb{E}[N_{\mathrm{HT}}] = 4$.

Finally, we need to evaluate $\mathbb{E}[N_{\mathrm{HH}}]$. After the first H has occurred, there is a probability $\frac{1}{2}$ that the next flip will have the result H. If not, then the wait for the first HH begins anew following the position $N_{\mathrm{H}} + 1$. Thus, with probability $\frac{1}{2}$, the random variable N_{HH} has the value $N_{\mathrm{H}} + 1$, and with probability $\frac{1}{2}$ the random variable $N_{\mathrm{HH}} - (N_{\mathrm{H}} + 1)$ has the same probability distribution as N_{HH}. From these considerations,

$$\begin{aligned} \mathbb{E}[N_{\mathrm{HH}}] &= \frac{1}{2}\mathbb{E}[N_{\mathrm{H}} + 1] + \frac{1}{2}\mathbb{E}[N_{\mathrm{H}} + 1 + N_{\mathrm{HH}}] , \\ &= 1 + \mathbb{E}[N_{\mathrm{H}}] + \frac{1}{2}\mathbb{E}[N_{\mathrm{HH}}] , \end{aligned} \tag{22}$$

and since $\mathbb{E}[N_{\mathrm{H}}] = 2$, this implies $\mathbb{E}[N_{\mathrm{HH}}] = 6$.

In summary, we have shown that

$$\mathbb{E}[N_{\mathrm{H}}] = 2, \ \ \mathbb{E}[N_{\mathrm{HT}}] = 4, \ \ \mathbb{E}[N_{\mathrm{HH}}] = 6. \tag{23}$$

The above recursive considerations can be elaborated into an alternative derivation of the probability distributions of N_{HT} and N_{HH}. To get started, we note that

$$\mathbb{P}\left[N_{\mathrm{H}}=n\right]=2^{-n}, \tag{24}$$

since there is a unique sequence of length n in which the first H occurs at the end of the sequence, namely the sequence in which all of the flips other than the last have outcome T, and the last flip has outcome H. Next, as noted above, $N_{\mathrm{HT}}-N_{\mathrm{H}}$ has the same probability distribution as N_{H}. Therefore,

$$\begin{aligned}\mathbb{P}\left[N_{\mathrm{HT}}=n\right] &= \sum_{n'=1}^{n-1}\mathbb{P}\left[N_{\mathrm{H}}=n'\right)\mathbb{P}\left[N_{\mathrm{HT}}-N_{\mathrm{H}}=n-n'\right] \\ &= \sum_{n'=1}^{n-1}2^{-n'}2^{-(n-n')}=\left(\sum_{n'=1}^{n-1}1\right)2^{-n}=(n-1)2^{-n}. \end{aligned}\tag{25}$$

Finally, for N_{HH}, we have the situation that $N_{\mathrm{HH}}=N_{\mathrm{H}}+1$ with probability $\frac{1}{2}$, otherwise (so again with probability $\frac{1}{2}$) $N_{\mathrm{HH}}-(N_{\mathrm{H}}+1)$ has the same probability distribution as N_{HH}. Therefore,

$$\begin{aligned}\mathbb{P}\left[N_{\mathrm{HH}}=n\right] &= \frac{1}{2}\mathbb{P}\left[N_{\mathrm{H}}=n-1\right]+\frac{1}{2}\sum_{n'=1}^{n-2}\mathbb{P}\left[N_{\mathrm{H}}=n'\right]\mathbb{P}\left[N_{\mathrm{HH}}=n-(n'+1)\right] \\ &= 2^{-n}+\sum_{n'=1}^{n-2}2^{-(n'+1)}\mathbb{P}\left[N_{\mathrm{HH}}=n-(n'+1)\right] \end{aligned}\tag{26}$$

Equation (26) holds for $n\geq 3$ with initial data

$$\mathbb{P}\left[N_{\mathrm{HH}}=1\right]=0,\ \ \mathbb{P}\left[N_{\mathrm{HH}}=2\right]=\frac{1}{4}. \tag{27}$$

In equation (26), we can multiply by 2^n and then make the change of variables $m=n-(n'+1)$ in the sum. Note that m runs through the same set of values as n', just in the opposite order. After these manipulations, we have

$$2^n\mathbb{P}\left[N_{\mathrm{HH}}=n\right]=1+\sum_{m=1}^{n-2}2^m\mathbb{P}\left[N_{\mathrm{HH}}=m\right]. \tag{28}$$

Now let

$$2^n\mathbb{P}\left[N_{\mathrm{HH}}=n\right]=F(n-1). \tag{29}$$

Then (27) and 28) become

$$F(0) = 0, \;\; F(1) = 1, \tag{30}$$

$$\begin{aligned} F(n-1) &= 1 + \sum_{m=1}^{n-2} F(m-1) \\ &= 1 + \sum_{m=0}^{n-3} F(m). \end{aligned} \tag{31}$$

Now by increasing n by 1 in (31), we get

$$\begin{aligned} F(n) &= 1 + \sum_{m=0}^{n-2} F(m), \\ &= 1 + \sum_{m=0}^{n-3} F(m) + F(n-2), \\ &= F(n-1) + F(n-2). \end{aligned} \tag{32}$$

Equations (30) and (32) are the initial conditions and recursion relation of the Fibonacci sequence.

Appendix B

We can also consider the coin-flipping process as a discrete-time Markov chain on the four states TT,TH,HT,HH.[2] The allowed transitions are from C_1C_2 to C_2C_3, with probability 1/2 for $C3 = $ T and 1/2 for $C_3 = $ H.

With the states ordered as above, the transition matrix is

$$P = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \tag{33}$$

[2]This point of view was suggested by Yuri Bakhtin, who also suggested the method used below to determine the probability that one particular state of a Markov chain is visited before another particular state.

Let $P^{C_1C_2}$ be the transition matrix of the Markov chain that has been modified by making C_1C_2 into a hole in the Markov chain, so that the process stops when a transition to C_1C_2 occurs. These modified transition matrices are obtained from P by deleting the row and column corresponding to the state $C1C2$. In particular

$$P^{\mathrm{HT}} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \tag{34}$$

$$P^{\mathrm{HH}} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \tag{35}$$

The powers of the modified transition matrices can be evaluated by making use of the Cayley-Hamilton theorem. The characteristic equation of $2P^{\mathrm{HT}}$ is

$$0 = \lambda(\lambda-1)^2 = \lambda^3 - 2\lambda^2 + \lambda, \tag{36}$$

so it follows from the Cayley-Hamilton tbeorem that

$$0 = \left(2P^{\mathrm{HT}}\right)^3 - 2\left(2P^{\mathrm{HT}}\right)^2 + \left(2P^{\mathrm{HT}}\right). \tag{37}$$

Multiplication by $\left(2P^{\mathrm{HT}}\right)^{n-1}$ for any $n \geq 1$ followed by slight rearrangement gives the result

$$\left(2P^{\mathrm{HT}}\right)^{n+2} - \left(2P^{\mathrm{HT}}\right)^{n+1} = \left(2P^{\mathrm{HT}}\right)^{n+1} - \left(2P^{\mathrm{HT}}\right)^{n}. \tag{38}$$

Thus, the successive powers of $2P^{\mathrm{HT}}$ differ by a constant matrix. To evaluate this constant matrix, we need only consider

$$\left(2P^{\mathrm{HT}}\right)^2 - 2P^{\mathrm{HT}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{39}$$

It follows that

$$\left(2P^{\mathrm{HT}}\right)^n = \begin{pmatrix} 1 & 1 & n-1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad n \geq 1 \tag{40}$$

The characteristic equation of $2P^{\mathrm{HH}}$ is

$$
\begin{aligned}
0 &= \det \begin{pmatrix} \lambda - 1 & -1 & 0 \\ 0 & \lambda & -1 \\ -1 & -1 & \lambda \end{pmatrix} \\
&= \lambda^2(\lambda - 1) - 1 - (\lambda - 1) \\
&= \lambda^3 - \lambda^2 - \lambda. \qquad (41)
\end{aligned}
$$

It follows that

$$\left(2P^{\mathrm{HH}}\right)^3 = \left(2P^{\mathrm{HH}}\right)^2 + \left(2P^{\mathrm{HH}}\right), \qquad (42)$$

and for $n \geq 1$ we can multiply by $\left(2P^{\mathrm{HH}}\right)^{n-1}$ to obtain

$$\left(2P^{\mathrm{HH}}\right)^{n+2} = \left(2P^{\mathrm{HH}}\right)^{n+1} + \left(2P^{\mathrm{HH}}\right)^{n} \qquad (43)$$

Note that this is the recursion relation of the Fibonacci sequence, which is defined by

$$F(0) = 0, \;\; F(1) = 1, \;\; F(n+2) = F(n+1) + F(n); \qquad (44)$$

To see in exactly what way the powers of $2P^{\mathrm{HH}}$ are described by the Fibonacci sequence, we need to evaluate

$$\left(2P^{\mathrm{HH}}\right)^2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}. \qquad (45)$$

Then

$$
\begin{aligned}
2P^{\mathrm{HH}} &= \begin{pmatrix} F(1) & F(1) & F(0) \\ F(0) & F(0) & F(-1) \\ F(1) & F(1) & F(0) \end{pmatrix} \qquad (46) \\
\left(2P^{\mathrm{HH}}\right)^2 &= \begin{pmatrix} F(2) & F(2) & F(1) \\ F(1) & F(1) & F(0) \\ F(2) & F(2) & F(1) \end{pmatrix} \qquad (47)
\end{aligned}
$$

Note that we need to use $F(-1) = 1$ to make the (2,3) element fit the pattern that the argument of F increases by one in each matrix element for every increase of the power of the matrix by 1. This extension of F to negaive argument is consistent with the recursion relation of the Fibonacci sequence.

It follows from (46-47) together with the matrix recursion relation (43) that

$$\left(2P^{\rm HH}\right)^n = \begin{pmatrix} F(n) & F(n) & F(n-1) \\ F(n-1) & F(n-1) & F(n-2) \\ F(n) & F(n) & F(n-1) \end{pmatrix}, \quad n \geq 1. \tag{48}$$

Now that we have the matrix powers determined, we can easily find the probability distributions for the first occurrences of HT and HH. Let $q_{\rm HT}(n)$ be the probability that HT has *not* appeared in the first n coin flips, and similarly let $q_{\rm HH}(n)$ be the probability that HH has *not* appeared in the first n coin flips. In both cases, $q(1) = 1$ and $q(2) = \frac{3}{4}$. For $n > 2$, $n-2$ is the number of steps of the Markov chain that have occurred, and therefore

$$\begin{aligned} q_{\rm HT}(n) &= \frac{1}{4} u \left(P^{\rm HT}\right)^{n-2} u^{\rm t} \\ &= \frac{1}{4}\frac{1}{2^{n-2}}(n+1) = \frac{1}{2^n}(n+1) \end{aligned} \tag{49}$$

$$\begin{aligned} q_{\rm HH}(n) &= \frac{1}{4} u \left(P^{\rm HH}\right)^{n-2} u^{\rm t} \\ &= \frac{1}{4}\frac{1}{2^{n-2}}\left(4F(n-2) + 4F(n-3) + F(n-4)\right) \\ &= \frac{1}{2^n} F(n+2) \end{aligned} \tag{50}$$

where $u = (1,1,1)$ and $u^{\rm t}$ denotes the transpose of u.

Note the need for repeated use of the Fibonacci sequence recursion relation in the foregoing to derive

$$\begin{aligned} 4F(n-2) + 4F(n-3) + F(n-4) &= 5F(n-2) + 3F(n-3) \\ &= 3F(n-1) + 2F(n-2) \\ &= 2F(n) + F(n-1) \\ &= F(n+1) + F(n) \\ &= F(n+2) \end{aligned} \tag{51}$$

Now let $p_{\rm HT}(n)$ be the probability that the first appearance of HT is completed on the $n^{\rm th}$ coin flip, and let $p_{\rm HH}(n)$ have the corresponding meaning for HH. Then for both cases, $p(n) = q(n-1) - q(n)$. Therefore

$$p_{\text{HT}}(n) = \frac{n}{2^{n-1}} - \frac{n+1}{2^n} = \frac{n-1}{2^n} \tag{52}$$

$$\begin{aligned} p_{\text{HH}}(n) &= \frac{F(n+1)}{2^{n-1}} - \frac{F(n+2)}{2^n} \\ &= \frac{2F(n+1) - F(n+2)}{2^n} \\ &= \frac{2F(n+1) - F(n+1) - F(n)}{2^n} \\ &= \frac{F(n+1) - F(n)}{2^n} \\ &= \frac{F(n-1)}{2^n} \end{aligned} \tag{53}$$

These are the same first-occurrence distributions as the ones found earlier by considering the coin flip process directly.

We also would like to evaluate the probability that HT occurs before HH within any sequence of coin flips, and the probability that TH occurs before HH within any sequence of coin flips.

In general, in a Markov chain with transition matrix P, let $r(i, j, k)$ be the probability, starting from the state i, of reaching the state j before k. Then $r(j, j, k) = 1$ and $r(k, j, k) = 0$. After one step of the Markov chain, the search for j or k begins anew from the new starting location i' that was reached in that one step from i to i'. Therefore,

$$\begin{aligned} r(i,j,k) &= \sum_{i'} P(i,i')r(i',j,k) \\ &= P(i,j)(1) + P(i,k)(0) + \sum_{i' \neq j \text{or } k} P(i,i')r(i',j,k). \end{aligned} \tag{54}$$

There is one equation of this kind for each starting location other than j or k, and together they form a linear system in the unknowns $r(i, j, k)$, for i not equal to j or k.

In our case, for competion between HT and HH, we have, for the starting states TT and TH,

$$r(\text{TT,HT,HH}) = \frac{1}{2}r(\text{TT,HT,HH}) + \frac{1}{2}r(\text{TH,HT,HH}) \tag{55}$$

$$r(\text{TH,HT,HH}) = \frac{1}{2}(1) + \frac{1}{2}(0) \tag{56}$$

Substitution of (56) into (55) followed by solution for $r(\text{TT,HT,HH})$ gives the result $r(\text{TT,HT,HH}) = \frac{1}{2}$. Thus, in summary, we have

$$\begin{aligned} r(\text{TT,HT,HH}) &= \frac{1}{2}, & (57)\\ r(\text{TH,HT,HH}) &= \frac{1}{2}, & (58)\\ r(\text{HT,HT,HH}) &= 1, & (59)\\ r(\text{HH,HT,HH}) &= 0, & (60) \end{aligned}$$

and since each of the starting states has probability $\frac{1}{4}$, we have the result that

$$\mathbb{P}\left[N_{\text{HT}} < N_{\text{HH}}\right] = \frac{1}{4}\left(\frac{1}{2} + \frac{1}{2} + 1 + 0\right) = \frac{1}{2}. \qquad (61)$$

In a contest between TH and HH, the starting states that we need to consider are TT and HT. We have

$$\begin{aligned} r(\text{TT,TH,HH}) &= \frac{1}{2} r(\text{TT,TH,HH}) + \frac{1}{2}(1), & (62)\\ r(\text{HT,TH,HH}) &= \frac{1}{2} r(\text{TT,TH,HH}) + \frac{1}{2}(1). & (63) \end{aligned}$$

From (62), $r(\text{TT,TH,HH}) = 1$, and then from (63), $r(\text{HT,TH,HH}) = 1$ as well.

In summary,

$$\begin{aligned} r(\text{TT,TH,HH}) &= 1, & (64)\\ r(\text{HT,TH,HH}) &= 1, & (65)\\ r(\text{TH,TH,HH}) &= 1, & (66)\\ r(\text{HH,TH,HH}) &= 0, & (67) \end{aligned}$$

and it follows that

$$\mathbb{P}\left[N_{\text{TH}} < N_{\text{HH}}\right] = \frac{3}{4}. \qquad (68)$$

Equations (61) and (68) are in agreement with the results found previously concerning the contests HT vs HH and TH vs HH.

In summary, we have replicated all of our previous results by working here within the Markov-chain framework.