Fibonacci Distribution and the Expected Number of Coin Flips Until Two Consecutive Heads

Saurabh Joglekar

1 Introduction

In this article, we discuss a famous problem from basic probability theory and link it with the Fibonacci probability distribution. This problem is often presented as a puzzle. It has been discussed multiple times across various forums. The main goal of this article is to identify some gaps in the solution that is typically presented to beginners, and to develop a solution that appeals to readers who seek more mathematical rigor.

> **Problem:** A fair coin is flipped repeatedly until you get two consecutive heads. What is the expected number of coin flips?

A typical and widely presented "solution" goes as follows:

> Let X be the random variable that counts the coin flips until two consecutive heads. Let us denote the outcome of the coin flips by strings of H's and T's. For example, $HTHH$ denotes the outcome where first flip comes up heads, second flip comes up tails, and the third and fourth flips are both heads, at which point the experiment ends. Thus, $X(HTHH) = 4$.
>
> Let $\mathbb{E}[X] = x$.
>
> When $X = 2$, the only possibility is HH. Hence $\mathbb{P}(X = 2) = \frac{1}{4}$.
>
> If the first flip comes up tails, which happens with probability $\frac{1}{2}$, the experiment starts all over again with expected value increasing by 1.
>
> The only possibility left is that the first two flips come up heads and tails respectively, which happens with probability $\frac{1}{4}$. In this case, the experiment starts all over again with expected value increasing by 2.

Considering all three possibilities, we have:

$$x = \frac{1}{4} \times 2 + \frac{1}{2} \cdot (1 + x) + \frac{1}{4} \cdot (2 + x)$$

This equation yields $x = 6$.

Interested readers will find many variations of this argument. However, the central theme of most of the arguments is to write an equation for $\mathbb{E}[X]$ using recursion.

First of all, notice that the main result from probability theory which is used here is the *Law of Total Expectation*, which is often presented as $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$. Although not mentioned explicitly, the random variable Y on which X is conditioned can be defined as:

$$Y(\omega) \quad = \quad \begin{cases} 0 & \text{if } \omega = HH \\ 1 & \text{if } \omega \text{ starts with } T \\ 2 & \text{if } \omega \text{ starts with } HT \end{cases}$$

where ω is any outcome in the sample space Ω (which is represented by a string of Hs and Ts with the string ending at the first occurrence of HH).

Even though the argument presented above is not wrong, there are several points which need to be explained more rigorously. For example, consider the following points:

1. When a T appears, the argument states that the experiment/process starts all over again with the expected value incrementing by the number of coin flips that have already occured. But what does it exactly mean by "starting the experiment/process again"? To answer this, one may introduce the concept of state transitions. But this in turn requires a deeper understanding of Markov Chains and/or Finite State Machines.

2. The argument depends on the identities $\mathbb{E}[X|Y = 1] = 1 + \mathbb{E}[X]$ and $\mathbb{E}[X|Y = 2] = 2 + \mathbb{E}[X]$. Although intuitive, these identities require a rigorous proof.

3. Consider another conditioning random variable Z defined as:

$$Z(\omega) \quad = \quad \begin{cases} 0 & \text{if } \omega = HH \\ 1 & \text{if } \omega = THH \\ 2 & \text{if } \omega \text{ starts with } HT \\ 3 & \text{if } \omega \text{ starts with } TT \\ 4 & \text{if } \omega \text{ starts with } THT \end{cases}$$

Then $\mathbb{E}[\mathbb{E}[X|Z]] = \mathbb{E}[X]$ yields

$$\begin{aligned}
x &= \mathbb{E}[X] \\
&= \mathbb{E}[\mathbb{E}[X|Z]] \\
&= \sum_{k=0}^{4} \mathbb{P}(Z = k) \cdot \mathbb{E}[X|Z = k] \\
&= \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{4} \cdot (2 + x) + \frac{1}{4} \cdot (2 + x) + \frac{1}{8} \cdot (3 + x)
\end{aligned}$$

which again yields $x = 6$.

This shows that the conditioning random variable is not unique. In fact there exist infinitely many random variables on which X can be conditioned. The *Law of Total Expectation* ensures that all such conditioning random variables will yield the same result for $\mathbb{E}[X]$. The reason Y is preferred is that it is possibly the simplest conditioning random variable. Apart from this, there seems to be no reason to prefer Y over Z, or any other conditioning random variable for that matter.

4. Finally, note that the solution outlined above essentially asserts that *if* $\mathbb{E}[X]$ exists and is finite, *then* it is equal to 6. Perhaps the most significant limitation of this solution is the absence of a rigorous proof establishing that $\mathbb{E}[X]$ exists and is finite. It is well known that there are random variables for which the expected value either does not exist or is infinite—the St. Petersburg Paradox being a classic example.

The most straightforward way to address these shortcomings is to find the probability mass function (pmf) of X and use it to prove that $\mathbb{E}[X]$ exists and is finite. This also provides a direct way to calculate $\mathbb{E}[X]$ directly by summing a series.

2 Probability Mass Function (PMF) of X

We begin by formalizing the notations.

Let Ω be the sample space containing all (infinitely many) possible outcomes of the experiment. Each outcome corresponds to a sequence of coin flips, where a single flip results in either H (heads) or T (tails). The experiment terminates upon the first appearance of two consecutive heads. Hence, each outcome $\omega \in \Omega$ is a finite sequence of H's and T's that ends with the substring HH.

Since the coin is fair, it is clear that $\mathbb{P}(H) = \mathbb{P}(T) = \frac{1}{2}$.

Let $X : \Omega \to \mathbb{N}$ be the random variable that assigns to each outcome $\omega \in \Omega$ the total number of symbols in ω. Equivalently, $X(\omega)$ represents the number of coin tosses required until the first occurrence of two consecutive heads. If $X(\omega) = k$, then $\mathbb{P}(\omega) = \frac{1}{2^k}$ for every $k \geq 2$ since each coin flip is independent of the others.

Now let $\Omega_k = \{\omega \mid X(\omega) = k\}$. Then $\Omega_0 = \Omega_1 = \varnothing$, and for every $k \geq 2$, Ω_k contains finitely many elements. It is clear that

$$\Omega = \bigcup_{k=2}^{\infty} \Omega_k$$

Note that $\Omega_2 = \{HH\}$ and $\Omega_3 = \{THH\}$. We now show that $\Omega_k = \{T\} \times \Omega_{k-1} \cup \{HT\} \times \Omega_{k-2}$ for every $k \geq 4$.

We first clarify the meaning of the expression $\{T\} \times \Omega_{k-1}$. By definition, the Cartesian product $A \times B$ is the set of ordered pairs (a, b) such that $a \in A$ and $b \in B$. In the present context, however, there exists a natural bijection between each ordered pair $(a, b) \in A \times B$ and the concatenated string ab, where $a \in A$ and $b \in B$. Accordingly, we adopt the convention of identifying $A \times B$ with the set of such concatenated strings. In particular, we define $\{T\} \times \Omega_{k-1}$ as the set of all finite strings that begin with the symbol T followed by an element of Ω_{k-1}, and we henceforth regard this identification as implicit. $\{HT\} \times \Omega_{k-2}$ is interpreted similarly.

With the above notations and conventions, let $k \geq 4$ be an integer, and let $\omega \in \Omega_k$. For the sake of convenience, we also represent the string ω by $\omega_1\omega_2 \ldots \omega_k$ where ω_j is the result of j-th coin flip. By definition of Ω_k, the string ω contains its first and only occurrence of two consecutive heads at positions $k-1$ and k. In particular, this implies that no other pair of consecutive heads appears in any of the subsequences corresponding to the 2nd through k-th coin flips or the 3rd through k-th coin flips.

We now consider the possible values of the first symbol of ω. If the first coin flip is a tail (i.e., $\omega_1 = T$), then the remaining sequence $\omega_2\omega_3 \ldots \omega_k$ must belong to Ω_{k-1}, since the first occurrence of two heads still occurs at positions $k-1$ and k. This implies $\omega \in \{T\} \times \Omega_{k-1}$.

On the other hand, if the first coin flip is a head (i.e., $\omega_1 = H$), then to avoid an early occurrence of two consecutive heads, it must be that $\omega_2 = T$. In this case, the remaining string $\omega_3\omega_4 \ldots \omega_k$ must belong to Ω_k. Therefore, $\omega \in \{HT\} \times \Omega_{k-2}$.

Combining both cases, we can conclude that

$$\omega \in \Omega_k \implies \omega \in \{T\} \times \Omega_{k-1} \cup \{HT\} \times \Omega_{k-2}$$

and hence

$$\Omega_k \subseteq \{T\} \times \Omega_{k-1} \cup \{HT\} \times \Omega_{k-2}.$$

Conversely, assume that $\omega \in \Omega_{k-1}$. It is clear that prepending ω with a T does not change the fact that the first and only occurrence of two consecutive heads appears at the last two positions in the string. Thus

$$\omega \in \Omega_{k-1} \implies \{T\} \times \{\omega\} \in \Omega_k$$

for every $\omega \in \Omega_{k-1}$. This means that

$$\{T\} \times \Omega_{k-1} \subseteq \Omega_k.$$

A similar argument shows that

$$\{HT\} \times \Omega_{k-2} \subseteq \Omega_k.$$

Thus

$$\{T\} \times \Omega_{k-1} \cup \{HT\} \times \Omega_{k-2} \subseteq \Omega_k.$$

Hence

$$\Omega_k = \{T\} \times \Omega_{k-1} \cup \{HT\} \times \Omega_{k-2} \tag{1}$$

for every integer $k \geq 4$.

We now turn to the matter of PMF of X. We have already shown that for every element $\omega \in \Omega_k$, $\mathbb{P}(\omega) = \frac{1}{2^k}$. Combining this with the fact that every Ω_k contains finite number of elements, it is clear that

$$\mathbb{P}(X = k) = \mathbb{P}(\Omega_k) = \frac{|\Omega_k|}{2^k}$$

Now let $|\Omega_k| = C_k$. Note that $|\{T\} \times \Omega_{k-1}| = |\Omega_{k-1}| = C_{k-1}$ and $|\{HT\} \times \Omega_{k-2}| = |\Omega_{k-2}| = C_{k-2}$.

With this notation, we have $C_1 = 0, C_2 = 1, C_3 = 1$, and equation (1) implies that $C_k = C_{k-1} + C_{k-2}$ for each $k \geq 4$. This is nothing but shifted Fibonacci sequence $\langle F_n \rangle$ where $F_0 = 0, F_1 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for all $n \geq 2$. Thus $C_k = F_{k-1}$ for all $k \geq 1$. Hence

$$\mathbb{P}(X = k) = \frac{F_{k-1}}{2^k}, \qquad \forall k \geq 1 \tag{2}$$

To verify that this is a valid PMF, we need to show that $\displaystyle\sum_{k=1}^{\infty} \frac{F_{k-1}}{2^k} = 1$. To show this, we use Binet's formula which expresses the Fibonacci numbers in closed form:

$$F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}} \tag{3}$$

for all $n \geq 0$ where $\varphi = \frac{1+\sqrt{5}}{2}$ and $\psi = 1 - \varphi = -\frac{1}{\varphi} = \frac{1-\sqrt{5}}{2}$.

Now use equation (3) in the series $\displaystyle\sum_{k=1}^{\infty} \frac{F_{k-1}}{2^k} = \frac{1}{2}\sum_{k=0}^{\infty} \frac{F_k}{2^k}$ and note that this series is absolutely convergent since $\left|\frac{\varphi}{2}\right| < 1$ and $\left|\frac{\psi}{2}\right| < 1$. Thus, $\displaystyle\frac{1}{2}\sum_{k=0}^{\infty} \frac{F_k}{2^k} = \frac{1}{2\sqrt{5}}\left(\sum_{k=0}^{\infty} \frac{\varphi^k}{2^k} - \sum_{k=0}^{\infty} \frac{\psi^k}{2^k}\right) = \frac{1}{2\sqrt{5}} \cdot \left(\frac{1}{1-\frac{\varphi}{2}} - \frac{1}{1-\frac{\psi}{2}}\right) = 1$. This shows that equation (2) represents a valid PMF. This PMF can be appropriately called the ***Fibonacci Distribution***.

3 Expectation of X

Now that we have established the PMF of X, we will turn to its expectation. By definition,

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=2}^{\infty} k \cdot \mathbb{P}(X = k) \\
&= \sum_{k=2}^{\infty} k \cdot \frac{F_{k-1}}{2^k} \\
&= \sum_{k=1}^{\infty} (k+1) \cdot \frac{F_k}{2^{(k+1)}} \qquad (4)
\end{aligned}$$

Applying ratio test to the series in equation (4), we can conclude that the series is absolutely convergent. Thus $\mathbb{E}[X]$ exists and is finite. Hence we can now safely apply the *Law of Total Expectation*. The final result, stated earlier in the article, gives $\mathbb{E}[X] = 6$.

Alternatively, we can compute $\mathbb{E}[X]$ by summing the series. The absolute convergence of series (4) allows us to rearrange the terms in any way without changing the sum. Thus,

$$\sum_{k=1}^{\infty} (k+1) \cdot \frac{F_k}{2^{(k+1)}} = \frac{1}{2}\sum_{k=1}^{\infty} k \cdot \frac{F_k}{2^k} + \frac{1}{2}\sum_{k=1}^{\infty} \frac{F_k}{2^k} \qquad (5)$$

We have already shown that $\displaystyle\frac{1}{2}\sum_{k=1}^{\infty} \frac{F_k}{2^k} = 1$. To compute the other summation in (5), we once again use Binet's formula and appeal to the absolute convergence of the series. Hence,

$$
\begin{aligned}
\frac{1}{2}\sum_{k=1}^{\infty} k \cdot \frac{F_k}{2^k} &= \frac{1}{2\sqrt{5}} \sum_{k=1}^{\infty} k \cdot \frac{\varphi^k - \psi^k}{2^k} \\
&= \frac{1}{2\sqrt{5}} \sum_{k=1}^{\infty} k \cdot \left(\frac{\varphi}{2}\right)^k - \frac{1}{2\sqrt{5}} \sum_{k=1}^{\infty} k \cdot \left(\frac{\psi}{2}\right)^k \\
&= \frac{1}{2\sqrt{5}} \cdot \frac{\varphi}{2} \sum_{k=1}^{\infty} k \cdot \left(\frac{\varphi}{2}\right)^{k-1} - \frac{1}{2\sqrt{5}} \cdot \frac{\psi}{2} \sum_{k=1}^{\infty} k \cdot \left(\frac{\psi}{2}\right)^{k-1} \\
&= \frac{1}{2\sqrt{5}} \cdot \frac{\varphi}{2} \cdot \frac{d}{dx}\frac{1}{1-x}\bigg|_{x=\frac{\varphi}{2}} - \frac{1}{2\sqrt{5}} \cdot \frac{\psi}{2} \cdot \frac{d}{dx}\frac{1}{1-x}\bigg|_{x=\frac{\psi}{2}} \\
&= \frac{1}{2\sqrt{5}} \cdot \frac{\varphi}{2} \cdot \frac{1}{(1-x)^2}\bigg|_{x=\frac{\varphi}{2}} - \frac{1}{2\sqrt{5}} \cdot \frac{\psi}{2} \cdot \frac{1}{(1-x)^2}\bigg|_{x=\frac{\psi}{2}} \\
&= \frac{1}{2\sqrt{5}} \cdot \frac{\varphi}{2} \cdot \frac{1}{(1-\frac{\varphi}{2})^2} - \frac{1}{2\sqrt{5}} \cdot \frac{\psi}{2} \cdot \frac{1}{\left(1-\frac{\psi}{2}\right)^2} \\
&= \frac{1}{4\sqrt{5}}\left[\frac{4\varphi}{(2-\varphi)^2} - \frac{4\psi}{(2-\psi)^2}\right] \\
&= 5
\end{aligned}
$$

Substituting this in equation (5), $\mathbb{E}[X] = 5 + 1 = 6$.