

What Does Random Look Like?

- By Pratyush Joshi, 08D, Wilson's School

When I asked a friend, they largely underestimated streaks and came up with the coin pattern of:

HHTHHHTTHTHTHHHTHHHTT

When I used the coin simulator from the website, I got:

HHTHHHTHTHHHHHHHTTTTHH

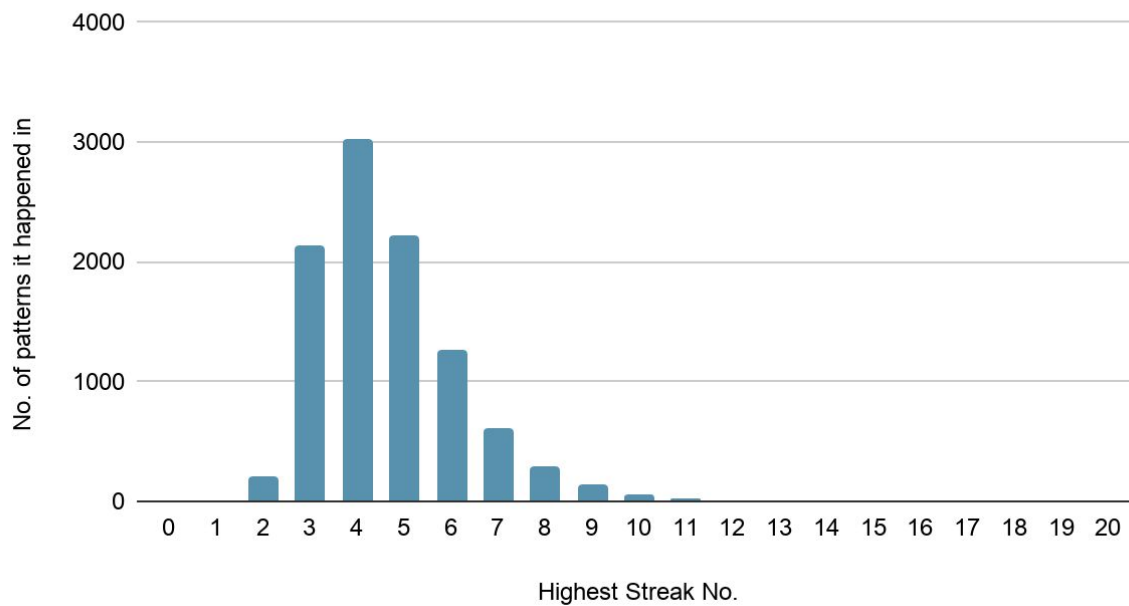
The above figure's highest streak is 7 which is very high compared to 3 being the highest streak when my friend did it.

The main difference between randomness and fakes in this scenario is streaks. Streaks are when there are heads or tails in a row multiple times. For example, HHH would be a streak of 3, while TTTTTT would be a streak of 6. The difference between almost all artificial and computer generated ones, was that the commonness of high streaks was largely underestimated. Most people would expect a streak of 2 or 3 and even a streak of 4 would lead people to believe that it is too much of a stretch and that they will be found out quite easily as fakes but the idea that streaks could be as high as 6 or 7 regularly is still foreign to the human mind.

In reality, what happens when the coins are rolled, as I saw from the website, is that when the coins are tossed, there were many streaks, often streaks of 4, 5, or 6. This differed from what I asked my friends and parents to come up with as the highest streak which I got from that was a streak of 4. When I looked at the computer generated tosses and the real tosses, the missing streaks actually became quite clear. The reason for these streaks is that when a coin is rolled and it lands on Heads, if you roll the same coin again, it *still* has the same chance of landing on Heads. When you roll two coins again after another, there is $\frac{1}{2}$ chance of heads coming the first time and $\frac{1}{2}$ chance of Heads coming the second time again. When we say that there is a $\frac{1}{4}$ chance of the two coins rolling HH, we mean that out of all the possibilities of HT, TH, TT, HH, the streak of Heads will come out $\frac{1}{4}$ of the time. If a coin is rolled and there is 1 more chance to roll the coin there is still a $\frac{1}{2}$ chance that you will roll a streak of 2 heads. If we count streaks of Tails as well, the chance of getting a streak, Heads or Tails is $\frac{1}{2}$, which is actually quite a lot. This can alternatively be written as $2/2^2$ which translates to $1/2^1$. If we roll 20 coins at one time, the chance of getting a streak of 20 is $2/2^{20}$ or more simply $1/2^{19}$. This means that getting a streak of either Heads or Tails is $1/524288$. This means that it is very hard to get a streak of 20, as it means that if you toss the coins 524288 times, you are likely to get it once. But this is the same if you wanted to get a pattern of HTHTHTHTHTHTHTHTHTHT or THTHTHTHTHTHTHTHTHTH. It is likely to be the same as the streak of 20.

To find out how high the streaks got in each pattern of coins, I decided to class the patterns by the highest streak that happened in the pattern, real and fake. After finding out some more fakes and reals, I decided that the fakes would rarely go above 3 and if they did, they would never go past 5. The computer generated ones normally were in the category of 4, 5, 6 and 7. To test out my theory, I wrote a computer program that would simulate rolling 20 coins 10000 times and noting the highest streak found in the pattern. Then it would put it in the list of how high the highest streaks were, and from that data, I made a graph:

No. of patterns it happened in vs Highest Streak No.



As it is clearly seen from this graph, 4 had the highest amount of patterns then 5, then 3. After that, it was 6, 7, 8 and then 2. After that it was 9, 10, 11, 12, 14, 13 and 1 pattern even had a streak of 18 (some may be a little bit difficult to see on this graph). This shows that the most optimal streak for a pattern is 4 then 5 with 3 not too far behind, which I did not predict. Streaks of 6 are still quite high in frequency, however from then they get progressively less, however with some anomalous data at 14 and 13 with 14 having a frequency of 3 and 13 having a frequency of 1. One thing that would surprise people is how far 2 is down on the list. A high streak of 2 would mean a pattern like THTHHTTHTHTHHTHTTHHT for example. The reason for this is that there is a higher chance of a streak of 3 being rolled the next time or 4 rather than only streaks of 1s or 2s. Still, I expected the chance of 2 coming to be higher. I ran the simulation more times and got very similar results. In all the times that I ran the simulation, 4 had the highest frequency (at the beginning, I had expected 5) and then 5 came in second with 3 not far behind. 2 was always 1 and a half and 12 was usually the highest streak that anyone got. The mean streak of all the data was 4.661 which tells that if you were to toss 20 coins, the highest streak would fall between 4 and 5.

In conclusion, to analyse whether your friend tossed a real coin or made up a fake pattern depends on what the largest streak they kept in there was. If it was 2 or 3 then it is probably a fake but if it is somewhere near 4 or 5 then they had probably tossed the 20 coins and did it honestly. But in reality, streaks of 2 did happen, it was just that they happened *much* less in frequency than all the others. The main reason why the human prediction was much less than the actual one was because sometimes it is difficult to understand that when a coin is tossed, what happened before does not actually have any effect on the coin. It can land Heads or Tails. When we say that the pattern HHHH or TTTT has a $\frac{1}{8}$ chance of happening (with 4 coins) we mean that out of all the possibilities 2 of them will be a streak but looking at the data above in the chart, a streak of 4 might happen more often (somewhere around $\frac{3}{10}$ counting the data above).

The whole data I collected here is listed down below:

Streaks	Frequency
0	0
1	0
2	202
3	2135
4	3033
5	2227
6	1257
7	606
8	294
9	139
10	62
11	26
12	14
13	1
14	3
15	0
16	0
17	0
18	1
19	0
20	0