Why does randomness become predictable?

Randomness is the lack of definite patterns or predictability in information and a random sequence of events often have no order to them. An example of randomness is present in a fair six-sided die. Each side will have a probability of a sixth of being rolled, however this does not guarantee that when you roll a die six times, you will get one of each number. For example, you may get zero 6s, or you may get two sixes, or, if you're extremely lucky, you may have got all 6s - the chances of that happening would be 0.002143347%!

You may be asking yourself how this randomness could be so called "predictable". After all, rolling a 6 does seem very random, and I agree. However, if you roll that same die over and over again 6 million times, how many 6s do you think will be rolled? You may be thinking somewhere around 1 million, right? To that, I'll applaud you, because you are correct. But why is this? Why is it that when we roll a die (significantly) more times we can be more sure of the sixth probability?

```
import random
rolls = 6000000
sixes = 0

for i in range(rolls):
        if random.randint(1, 6) == 6:
                sixes += 1

print("Sixes:", sixes)
print("Non-sixes:", rolls - sixes)
print("Proportion:", sixes / rolls)
```

This python code above highlights how the proportion of sixes to total rolls can differ vastly depending on the number of rolls you do. For example, I simulated 600 rolls three different times, and I got these proportions as the probability of getting a six: 0.177, 0.175, and 0.157. All of these results are similar to each other and are close to the exact sixth value that a fair six-sided die has for each face. However, they still aren't as accurate as we'd hope it to be. If we expand the number of rolls to be 6 million, then we get a whole different story. The three results we get (rounded to the third significant figure) are 0.167, 0.167, and 0.167, which is precise and accurate. This happened due to a simple change in the code - we increased the number of trials.

Let me introduce you to the Law of Large Numbers. The reason why there is so much uncertainty and so much "randomness" involved in the initial example is due to the fact that there is a lot of variation in die rolls. You may get a six, or you may not. And the probability that you get a six is only a sixth, while the probability you don't is five sixths. As a result, with a few rolls it is inevitable that the variation caused by the probability differences leads to a different proportion of sixes than the sixth probability that was promised to us. When we roll the die 6 million times, we find that the proportion of sixes stabilises and is much closer to the true probability, because the variations have less of an impact when the sample size is great. Extreme deviations start mattering less and the averages stabilise. However, earlier we stated that randomness cannot have any patterns or predictability, yet with this example of a large number of rolls, we can clearly see that past a certain point, results start to become predictable, despite it still being random, leading to a possible contradiction.

Unfortunately, in maths, there are no contradictions. So what possibly could explain this facade of predictability that's latched on to a die with a large number of rolls? If you zoom in on a single roll, then yes, it is random, however zooming out, you can see it becomes more and more predictable… in terms of averages. It isn't that the die rolls on six more probabilistically accurately, but it is just that the averages stabilise leading to the number of sixes being approximately one sixth of the total rolls.

Let's now think about what would happen if the number of die rolls tends towards infinity. The proportion of sixes approaches a sixth - the point at which the concept we call randomness becomes entirely predictable. But what is infinity? How can there be such a large enough number of rolls that results in a perfectly accurate representation of a die? None of us can imagine what infinity is like, or even a number remotely close to it, yet it gives very precise results in this context. This suggests that maths allows us to describe behaviour we cannot physically observe.

If we cannot observe such a concept like infinity, then we must have a limit to which we can observe so we use the term "tends to infinity" with limits. This is essential because limits help us define how many trials we've done, in this context, allowing us to make conclusions about the outcome we are testing for and whether its stated probability is true or not. This is used heavily in the financial markets for quantitative firms and hedge funds.

Although we can never reach infinity, we can still predict what happens as we get closer and closer, which is why we use "tends to infinity". Randomness doesn't disappear but patterns and structure emerge when we zoom out and see the entire experiment at scale. When we simulate die rolls, we physically can't reach infinity, resulting in us having to use limits in order to understand how behaviour changes when trials increase.

You may be thinking now that randomness doesn't exist anymore, however it still does. The only reason why it may appear that randomness isn't really random any more is due to structure appearing at scale, especially when at infinity. This gives a reason for limits to be a necessity and highlights how mathematical concepts help us understand impossible terms to the best of our abilities.