A Normal(?) Sequence

Aresh Pourkavoos

January 19, 2022

This idea is about OEIS entry A330731, which I submitted, and most of the information here is also available there. Nonetheless, I wanted to explore the sequence in a bit more detail with less of a wait time before publication, hence the separate source.

The sequence is infinite and binary, i.e. its entries are all either 0 or 1. I designed it to be *normal*, meaning that as the number of terms increases, every subsequence of the same length appears with equal frequency. For example, a randomly chosen subsequence of length 5 has a 1 in 32 chance of being 01001. At the moment, I still have not proven this property, but experimentally, this seems to be the case. However, I hope to prove that it is normal eventually, and in fact, I believe that the frequency of a given substring converges to its final value *faster* than a "random" sequence.

The sequence is defined by...

The naive algorithm to generate the sequence per the definition takes $O(n^3)$ time for the first n terms. This is because to generate each new term, the frequencies of $O(n)$ different tails are checked (from the whole sequence so far down to a relatively short tail), and each counting requires a pass over all terms, which takes $O(n)$ time. In other words, each term takes $O(n^2)$ time to generate. However, with the right data structures, I was able to reduce the time per term to $O(n)$, bringing the time for the first n terms to $O(n^2)$.