Many of us know the fact that, if a coin is thrown repeatedly, the number of times it turns up heads is random and that this randomness can be well described by a Gaussian bell curve. This approximation is valid in many other circumstances and follows a simple heuristic: if we sum up many small random effects, even if they are not entirely independent, that random sum will roughly follow a Gaussian distribution.

What is less known is that other distributions can arise if we consider different aspects of coin tossing. What is the number of times we need to toss the coin until we first observe heads? The answer is the geometric distribution, or its continuous counterpart, the exponential distribution. It is the same randomness we can observe when we step out into the night and wait until we see a shooting star. But what is the waiting time until we observe two consecutive heads? Although the exact behavior is now more difficult to calculate, it is again close to a geometric distribution.

In this article, we give such type of approximations, which arise in all branches of probability theory. It is the first article in the literature that treats exponential approximations in a coherent and unified manner. Not only can we recover classic results, but we can also equip these results with new quantitative bounds.

We are able to prove that the unifying aspect of this kind of randomness is the famous waiting paradox: if we assume that shooting stars arrive with an intensity of one shooting star each hour on average, how long do we have to wait for the next shooting star if we step into the night? It is not half an hour, but again one full hour on average. This seeming paradox is a consequence of the the exponential distribution, and, conversely, this property characterizes the exponential distribution. The closer it the paradox holds, the closer the waiting times are to an exponential distribution.