A Needle in a Haystack

Angelo Gnazzo, Leonardo Pagani

1 Introduction

Leo is blindfolded and has to find a needle in a haystack. In the haystack there are 10^{100} straws and 1 needle (yes it's a pretty huge haystack :)). The needle is stuck in between straws and in order to get it all straws must be removed from the haystack first. At every turn, Leo can randomly pick one straw and remove it. The task looks impossible, so Angelo is going to give him a little help. Every straw is labelled with a number from 1 to 10^{100}. If Leo picks a straw, Angelo is going to remove it from the haystack along with all the straws labelled with a bigger number. So, for example, if Leo picks the straw number 123456, at the next turn the haystack will contain all straws from 1 to 123455. On average how many turns it will take Leo to remove all straws and find the needle?

2 Solution

The problem seems complicated, but we can try to work out a solution. Let us define as X_n, the number of turns it takes Leo to find the needle, if there are n straws left in the haystack. We are asked to find $\mathbb{E}[X_{10^{100}}]$. We also know that $\mathbb{E}[0] = 0$. Let's try to find a recursive relation for $\mathbb{E}[X_n]$. We know that, $\forall n = 1...10^{100}$:

$$
\begin{aligned}
\mathrm{E}[\mathrm{X}_n] &= 1 + \frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}[X_k] = 1 + \frac{1}{n}\sum_{k=0}^{n-2}\mathbb{E}[X_k] + \frac{1}{n}\mathbb{E}[X_{n-1}] \\
&= 1 + \frac{n-1}{n(n-1)}\sum_{k=0}^{n-2}\mathbb{E}[X_k] + \frac{1}{n}\mathbb{E}[X_{n-1}] + \frac{n-1}{n} - \frac{n-1}{n} \\
&= 1 - \frac{n-1}{n} + \frac{n-1}{n}(1 + \frac{1}{n-1}\sum_{k=0}^{n-2}\mathbb{E}[X_k]) + \frac{1}{n}\mathbb{E}[X_{n-1}] \\
&= 1 - \frac{n-1}{n} + \frac{n-1}{n}(\mathbb{E}[X_{n-1}]) + \frac{1}{n}\mathbb{E}[X_{n-1}] \\
&= \frac{1}{n} + \mathbb{E}[X_{n-1}]
\end{aligned}
$$

We have found a nice and easy recursive relation for $\mathbb{E}[X_n]$.

This means that: $\mathbb{E}[X_n] = \sum_{k=1}^{n} \frac{1}{k}$.

So the expected number of straws to be lifted before cleaning the haystack is equal to the harmonic sum up to the number of straws in the haystack.

We remember that, for large n, $\sum_{k=1}^{n} \frac{1}{k} \approx \log n + \gamma$, where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

Finally:

$\mathbb{E}[X_{10^{100}}] \approx \log 10^{100} + \gamma = 100 \log 10 + \gamma \approx 230.83.$