Construction of the democratic sequence

Anne Bertrand Mathis

March 10, 2023

Let $(\lambda_n)_{n\geq 1}$ be a sequence of positive numbers with sum 1. Given a countable alphabet $A = a_1, a_2, ...$ we shall construct a sequence $(y_n)_{n\geq 1}$ on the alphabet A such that every letter a_k appears in the sequence $(y_n)_{n\geq 1}$ with the frequence λ_k in the following way :

We begin by the letter a_i which largest desired frequence λ_i (in case of a tie choose what you want among them).

After choosing $y_1...y_M$ we say that we are at step M ; given a letter a_k we define the default $D_M(a_{k)}$ of the letter a_k at step M as follow:

$$D_M\left(a_k\right) = M\lambda_k - (number\ of\ occurrences\ of\ a_k\ among\ y_1, ..., y_M)$$

Let y_{M+1} be the letter who has the greatest default at step M (there is always a letter with positive default).

We prove that each letter a_k appears in the sequence $(y_n)_{n\geq 1}$ with the expected frequence λ_k so this sequence is called the *democratic sequences associated to the frequencies* $(\lambda_n)_{n\geq 1}$.

Question : is the convergence to the correct frequencies particularly rapid?

Example of democratic sequence : consider the Fibonacci numeration, in basis $\beta = \frac{1+\sqrt{5}}{2}$ and the successive expansions :

$$0, 1, 10, 100, 101, 1000, 1001, 1010, 10000...$$

Suppose that you have written the expansions of the M first numbers $0, 1, 1, 2, ..., M-1$ and that you want to determine the expansion of M. If the expansion of $M-1$ is $\varepsilon_r...\varepsilon_{k+2}\varepsilon_{k+1}\varepsilon_k...\varepsilon_1$, to obtain those of M you will look at the digits on the right of $\varepsilon_r....\varepsilon_{k+2}\varepsilon_{k+1}\varepsilon_k...\varepsilon_1$ and take the largest sequence $\varepsilon_{k+1}\varepsilon_k ... \varepsilon_1$ equal to 00101010..10 or 00101010...1, and the expansion of M is $\varepsilon_r....\varepsilon_{k+2}010000..0$: you keep $\varepsilon_r....\varepsilon_{k+2}\varepsilon_{k+1}$, ε_k is growing from zero to one and the digits to the right of the k^{th}digit ε_k become zeroes (try!).

So we say that "the k^{th}digit ε_k is growing at step M ". It is the only growing digit.

In the Fibonacci system the " k^{th}digit ε_k" is growing β faster that the $(k+1)^{th}$digit ε_{k+1}, hence every h^{th}digit ε_h is growing with probability $\frac{1}{\beta^{h+1}}$.

Now looking at the democratic sequence associated to $(\lambda_n)_{n\geq 1} = \frac{1}{\beta^{n+1}}$, if the letter a_k correspond to "ε_k is growing " you will obtain a sequence saying which digit has to grow at each step.

Conjecture : in the system of numeration in basis $\beta > 1$ related to the β shift, the k^{th} digit is growing with probability $\lambda_k = \frac{1}{\beta^{k-1}} - \frac{1}{\beta^k} = \frac{\beta-1}{\beta^k}$. If the letter a_k corresponds to "ε_k is growing" we conjecture that the democratic sequence is the sequences saying which digit has to grow.

If the answer to both questions were positive it would be an encouraging signal about the following problem : the entropy of the β shift is $log\beta$: if $w_n$design the number of words of length n of the β shift lim $\frac{log w_n}{n} = log\beta$.

Conjecture : we conjecture that if X is a symbolic dynamical system of entropy $log\beta$, then the number x_n of $n-$words of X satisfies $w_n \leq x_n$: the β shift is the system with the smaller number of words (in french we would say that "the β shift is le meilleur rapport qualité-prix").