Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Departamento de Computación

Números Muy Normales

Tesis de Licenciatura en Ciencias de la Computación

Lucas Puterman

Director: Verónica Becher

Codirector: Olivier Carton

Buenos Aires, 2019

NÚMEROS MUY NORMALES

En esta tesis nos proponemos estudiar la noción de supernormalidad definida por Zeev Rudnick hace unos años. Lo poco que se conoce sobre esta noción no está publicado. Benjamin Weiss del Einstein Institute of Math de Hebrew University dio el 16 de Junio de 2010 una una conferencia en el Institute for Advanced Study in Princeton titulada "Random-like behavior in deterministic systems" donde describe la noción de supernormalidad, a la que llama Poisson generic (ver https://video.ias.edu/pseudo2010/weiss). En este video Weiss afirma que la mayoría de los números reales son supernormales y que la supernormalidad es más fuerte que la noción clásica de normalidad, es decir, que si un número es supernormal, entonces es normal pero no al revés. También afirma que el ejemplo más famoso de número normal, el número de Champernowne, no es supernormal. Y deja abierto el problema de dar una construcción explícita de un número supernormal. En esta tesis nos proponemos dar la demostración completa de que el número binario de Champernowne no es supernormal.

Palabras claves: Normalidad, Supernormalidad, Champernowne, Poisson, Pseudoaleatoreidad.

VERY NORMAL NUMBERS

In this thesis we aim to study the notion of supernormality defined by Zeev Rudnick a few years ago. The few things known about supernormality is not published. Benjamin Weiss from the Einstein Institute of Math de Hebrew University gave on June 16th, 2010 a lecture on "Random-like behavior in deterministic systems" where the notion of supernormal sequences is described under the name of Poisson generic sequences. (See https://video.ias.edu/pseudo2010/weiss). In this lecture, Weiss claims that almost every real number is supernormal and that the notion of supernormality is stronger than the classical notion of normaility. Which means that if a number is supernormal then it is normal, but not the other way around. Weiss also states that the most famous example of a normal number, the Champernowne number, is not supernormal. And finally he leaves open the problem of giving an explicit construction of a supernormal number. In this thesis we give the complete proof that the binary Champernowne number is not supernormal.

Keywords: Normality, Supernormality, Champernowne, Poisson, Pseudorandomness.

CONTENTS

1. INTRODUCTION

1.1 Normality

1.2 Supernormality

Let $\lambda > 0$ be a real number.

Let $a_k(\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$ for $k \geq 0$.

Let x be a sequence over a binary alphabet.

For u, w words, $|w|_u$ is the number of occurrences of u in w.

Let $A_{k,n}$ be the frequency of occurrence of words of length n that occur exactly k times in the first $\lfloor \lambda(2^n + n - 1) \rfloor$ symbols of x, or what is to say:

$$A_{k,n}(\lambda) = \frac{\#\{w : |w| = n, |x[1...\lfloor \lambda(2^n + n - 1) \rfloor]|_w = k\}}{2^n}$$

Definition 1.2.1. x is $\lambda-$supernormal if

$$\lim_{n \to \infty} A_{k,n}(\lambda) = a_k(\lambda)$$

for each integer $k \geq 0$.

Definition 1.2.2. x is supernormal if it is $\lambda-$supernormal $\forall \lambda \in \mathbb{R}$

2. SUPERNORMALITY IS STRONGER THAN NORMALITY

2.1 Supernormality implies normality (Proof by Olivier Carton)

Fact 2.1.1.

$$\sum_{k\geq 0} a_k(\lambda) = e^{-\lambda}\sum_{k\geq 0}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$$

Fact 2.1.2.

$$\sum_{k\geq 0} k a_k(\lambda) = e^{-\lambda}\sum_{k\geq 1}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\sum_{k\geq 0}\frac{\lambda^k}{k!} = \lambda$$

Fact 2.1.3.

$$\sum_{k\geq 0} A_{k,n} = 1$$

Fact 2.1.4.

$$A_{k,n} = 0 \ \textit{if}\ k > \lambda 2^n$$

Let $\epsilon > 0$ be chosen.

Let k_0 be chosen such that

$$\frac{1}{\lambda}\sum_{k>k_0} k a_k(\lambda) < \frac{\epsilon}{2}$$

Let n_0 be chosen such that

$$\forall i, n\ s.t\ 0 \leq i \leq k_0 - 1 \wedge n > n_0 \wedge |A_{i,n}(\lambda) - a_i(\lambda)| \leq \frac{\epsilon\lambda}{2k_0(k_0+1)}$$

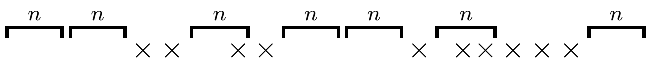
$$Rq? = \sum_{i=0}^{k_0-1} k a_k(\lambda) = \lambda - \sum_{k\geq k_0} k a_k(\lambda) \geq (\lambda - \frac{\epsilon}{2})$$

Consider the positions from 1 to $\lfloor \lambda 2^n \rfloor$ in x. A position is "blamed" if the word of length n starting at that position occurs more than k_0 times in $x\lfloor \lambda(2^n + n - 1)\rfloor$.

The number of non-blamed positions is

$$\begin{aligned}
\sum_{i=0}^{k-1} i A_{i,n}(\lambda) 2^n &\geq 2^n \sum_{i=0}^{k-1} i(a_i(\lambda) - \frac{\epsilon\lambda}{2k_0(k_0+1)}) \\
&\geq 2^n \sum_{i=0}^{k-1} i a_i(\lambda) - \frac{2^n\epsilon\lambda}{2} \\
&\geq 2^n\lambda(1 - \frac{\epsilon}{2}) - \frac{2^n\epsilon\lambda}{2} \\
&= 2^n\lambda(1-\epsilon)
\end{aligned}$$

We cover the positions of $1 \dots \lambda 2^n$ by blocks of length such that blocks do not start at a blamed position. However a block may contain blamed positions.



The number of left positions is less than $\epsilon\lambda 2^n$. For a fixed word w, the number of bad blocks is o(number of possible blocks). Since each block can occur at most $k_0 - 1$ times then it is ok. Hot spot lemma is needed to conclude.

2.2 Normality does not imply supernormality

We aim to prove that normality does not imply supernormality. For this we prove that the Champernowne sequence, proven to be normal in [BC18], fails to be supernormal by proving that it is not 1-supernormal. The strategy followed to achieve this is similar to the ones used in [BCC19] and [PS19].

Let $X(n)$ be the concatenation of all words of length n over the alphabet with two simbols $\mathcal{A} = \{0, 1\}$ in lexicographic order. It is clear that $X(n)$ has length $n2^n$. Then, for example:

$$X(2) = 00\ 01\ 10\ 11$$

Note that spaces were added for reading convenience.

Let the Champernowne sequence be the concatenation of $X(n)$ for $n = 1, 2, \dots$ Then the first symbols of the Champernowne sequence are:

$$champ = 0\ 1\ 00\ 01\ 10\ 11\ 000\ 001\ 010\ 011\ 100\ 101\ 110\ 111\ 0000\ 0001\ \dots$$

By the definition of λ-supernormality, if we take $\lambda = 1$, then we need to prove that

$$\lim_{n\to\infty} A_{k,n}(1) = a_k(1)$$

$$\lim_{n\to\infty} \frac{\#\{w : |w| = n, |x[1...(2^n + n - 1)]|_w = k\}}{2^n} \neq \frac{e^{-1}}{k!}$$

First, let's see that given d, we define $k = 2^d$, if we take $n \geq d + k + 1$ then the whole block $X(k)$ covered in the first 2^n symbols of Champernowne.

Fact 2.2.1. *$\forall n, n \geq 2$ it follows that $\sum_{i=1}^{n} i2^i < 2^{n+log(n)+1}$*

Fact 2.2.2. *$X(k)$ accounts for half of the total amount of symbols in the first 2^{d+k+1} symbols of Chamernowne.*

Proof. $2^{d+k+1} = 2^{d+k}2$ and $X(k)$ has length $k2^k = 2^d 2^k = 2^{d+k}$ □

We prove that Champernowne is not supernormal by showing that the frequence of words of length $k+d+1$ that do not occur in the first 2^{k+d+1} symbols of the Champernowne sequence is higher that the expected frequency if it were supernormal. To accomplish this, we exhaustively look how many different words of length 2^{k+d+1} are there within $X(k)$ and give an upper bound for the amount of different words that can appear in the first 2^{k+d+1} symbols of the Champernowne sequence.

Now, let's take a look at what the words of length $k+d+1$ that occur in $X(k)$ look like. There are four different cases that can happen of how a word x is formed with elements from $X(k)$. In the following analysis, u, v and w are consecutive words of length k in $X(k)$:

2.2.1 Bounding words that appear in Champernowne

If the Champernowne sequence were supernormal then the expected frequency of words that appear at least one time in the first $2^n + n - 1$ symbols would be:

$$\lim_{n\to\infty} \frac{\#\{w : |w| = n, |champ[1...(2^n + n - 1)]|_w > 0\}}{2^n} = 1 - e^{-1}$$

Now, by Fact 2.2.2 we know that $X(k)$ accounts for half of the words of the first 2^{d+k+1} symbols of Champernowne. If we analize what happens with words of length $d+k+1$ using the bounds we have for the occurrences of different words within $X(k)$ and we assume that the remaining $2^{d+k} + d + k - 1$ symbols are all different, then

$$\lim_{d\to\infty} \frac{\#\{w : |w| = d+k+1, |champ[1\ldots 2^{d+k+1} + d + k + 1 - 1]|_w > 0\}}{2^{d+k+1}} =$$

$$\lim_{d\to\infty} \frac{\#\{w : |w| = d+k+1, |champ[1\ldots 2^{d+k+1} + d + k]|_w > 0\}}{2^{d+k+1}} <$$

$$\lim_{d\to\infty} \frac{\text{case 1} + \text{case 2} + \text{case 4} + \text{last 2} + \text{other half} + (d+k)}{2^{d+k+1}} =$$

$$\lim_{d\to\infty} \frac{(2^k - 2\cdot 2^d) + (2\cdot 2^k) + (d2^k + 2\cdot 2^d) + (2d + k + 1) + (2^{d+k}) + (d+k)}{2^{d+k+1}} =$$

$$\lim_{d\to\infty} \frac{(2^k - 2k) + (2\cdot 2^k) + (d2^k + 2k) + (3d + 2k + 1) + (k2^k)}{2k2^k} =$$

$$\lim_{d\to\infty} \frac{3\cdot 2^k + d2^k + 3d + 2k + 1}{2k2^k} + \frac{1}{2} =$$

$$\lim_{d\to\infty} \frac{3}{2k} + \frac{d}{2k} + \frac{d}{k2^k} + \frac{1}{2\cdot 2^k} + \frac{1}{2k2^k} + \frac{1}{2} = \frac{1}{2} < 1 - e^{-1}$$

Finally, if in the original sequence we consider that the n's such that $n = d + 2^d + 1$ are a subsequence of $n = 1, 2, \ldots$ then we can say that if

$$\lim_{n\to\infty} \frac{\#\{w : |w| = n, |champ[1...(2^n + n - 1)]|_w > 0\}}{2^n}$$

exists, then it is not $1 - e^{-1}$. This implies that Champernowne is not 1-supernormal which means it is not supernormal.

Corollary 2.2.0.1. *If x is a normal number it is not implied that x is supernormal.*

3. BIBLIOGRAPHY

BIBLIOGRAPHY

[BC18] Verónica Becher and Olivier Carton. *Normal Numbers and Computer Science*, pages 233–269. Springer International Publishing, Cham, 2018.

[BCC19] Veronica Becher, Olivier Carton, and Ignacio Cunningham. Low discrepancy sequences failing poissonian pair correlations. *Archiv der Mathematik*, 113, 04 2019.

[PS19] Ísabel Pirsic and Wolfgang Stockinger. The champernowne constant is not poissonian. *Funct. Approx. Comment. Math.*, 60(2):253–262, 03 2019.