A Taste of Normal Numbers

Xiang Huang [1]

February 6, 2019

[1]Computer Science Department, Iowa State University

Normality: definition

Simply normal

Let's consider an infinite sequence x of zeros and ones, obtained by tossing a fair coin:

0101101001011110101001000101...

When could we say it is random?

Simply normal

Let's consider an infinite sequence x of zeros and ones, obtained by tossing a fair coin:

0101101001011110101001000101...

When could we say it is random?

A minimal requirement: zeros and ones appears "equally often".

Simply Normal

Definition
Let $\Sigma = \{0, 1\}$. An infinite sequence $x \in \Sigma^\infty$ is *simply normal* if for all $b \in \Sigma$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{\Sigma^* b}(x \upharpoonright i) = \frac{1}{2}. \tag{1}$$

Simply Normal

Definition
Let $\Sigma = \{0, 1\}$. An infinite sequence $x \in \Sigma^\infty$ is *simply normal* if for all $b \in \Sigma$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{\Sigma^* b}(x \upharpoonright i) = \frac{1}{2}. \tag{1}$$

- This can be viewed as counting length 1 “substring”.

Simply Normal

Definition
Let $\Sigma = \{0,1\}$. An infinite sequence $x \in \Sigma^\infty$ is *simply normal* if for all $b \in \Sigma$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{\Sigma^* b}(x \upharpoonright i) = \frac{1}{2}. \tag{1}$$

- This can be viewed as counting length 1 “substring”.
- Nothing stops us from counting longer “substring”.

Simply Normal

Definition
Let $\Sigma = \{0, 1\}$. An infinite sequence $x \in \Sigma^\infty$ is *simply normal* if for all $b \in \Sigma$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{\Sigma^* b}(x \upharpoonright i) = \frac{1}{2}. \tag{1}$$

- This can be viewed as counting length 1 “substring”.
- Nothing stops us from counting longer “substring”.
- Two ways of counting: sliding and blocking.

Normality: Sliding

Definition
We say $x \in \Sigma^\infty$ is normal if for all $u \in \Sigma^*$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{\Sigma^* u}(x \upharpoonright i) = \frac{1}{2^{|u|}}. \tag{2}$$

Normality: Blocking

Definition
A sequence $x \in \Sigma^\infty$ is α-normal if for all $u \in \Sigma^*$ such that $|u| = k$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{\Sigma^* u}(x \upharpoonright i \cdot k) = \frac{1}{2^k}. \tag{3}$$

- We can view a block u as a single letter from a giant alphabet Σ^k.
- Then counting blocks in the above way can be viewed as counting letters.
- The above definition says x is normal if x is simply normal to base Σ^k for all k.

Equivalence of the Definitions

Theorem
The sliding definition and the blocking definition are equivalent.

Proof.

- non-trivial.
- by counting.[1]

□

[1] a detailed proof can be found in *Irrational Numbers* by Ivan Niven.

Examples of Normal Numbers

- It is not known numbers as $\sqrt{2}$, e, π are normal to any base.

Examples of Normal Numbers

- It is not known numbers as $\sqrt{2}$, e, π are normal to any base.
- Champernowne number

$$x = 0.1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14 \cdots$$

  is normal to base 10.
- Copeland–Erdös number

$$x = 0.2\ 3\ 5\ 7\ 11\ 13\ 17 \cdots$$

  is also normal.
- Actually, Copeland and Erdös proved that if we concatenate a "dense" subset of natural number, the resulting number will be normal. Here the set of primes is consider to be dense enough.

Selection

Selection

Question: Given a normal number/sequence x, we select a subsequence of x, by **some rule**. What can we say about resulting subsequence? Can it still be normal?

Selection

Question: Given a normal number/sequence x, we select a subsequence of x, by **some rule**. What can we say about resulting subsequence? Can it still be normal?

Definition
(Selection Rule) A selection rule S is a subset of Σ^*.

- Whenever a prefix of x falls in S, we select the next digit.

Selection

Question: Given a normal number/sequence x, we select a subsequence of x, by **some rule**. What can we say about resulting subsequence? Can it still be normal?

Definition
(Selection Rule) A selection rule S is a subset of Σ^*.

- Whenever a prefix of x falls in S, we select the next digit.
- We only discuss normality of the resulting subsequence when it has infinitely many digits.

Selection

Question: Given a normal number/sequence x, we select a subsequence of x, by **some rule**. What can we say about resulting subsequence? Can it still be normal?

Definition
(Selection Rule) A selection rule S is a subset of Σ^*.

- Whenever a prefix of x falls in S, we select the next digit.
- We only discuss normality of the resulting subsequence when it has infinitely many digits.

Selection Rule: Short-term Memory

In a coin-tossing game, a gambler believes that if he sees three heads (1's), then what comes next must be in favor of tails (0's).

We can make this belief into a selection rule

$$S = \Sigma^* 111.$$

For a sequence $x = 0011101111001110\cdots$, we have after the selection

$$x_S = 0\ 1\ 0\cdots$$

We denote the family of selection rules that determine by a "Short-term memory" $u \in \Sigma^*$ as

$$\mathcal{F}_b = \{\Sigma^* u \mid u \in \Sigma^*\}.$$

The subscript "b" refers to "block", because we use a fixed block for the matching.

Preservation of Normality

Theorem
(Postnikova) The family $\mathcal{F}_b$ preserves normality.

- No matter what block u we use, the resulting subsequence will end up to be normal.
- Actually, all the rules/languages in $\mathcal{F}_b$ are just special cases of regular languages, which can be recognized by *finite state machines.*
- Let $\mathcal{F}_r =$ "*the set of all regular languages*".

We then have

Theorem
(Agafonov) The family $\mathcal{F}_r$ preserves normality.

Preservation of Normality: Kamae and Weiss's Approach

Let $S \subseteq \Sigma^*$ be a selection rule. We define an equivalence relation $\sim_S$ on Σ^* as follows:

$$u \sim_S v \quad \text{if} \quad \{w \mid uw \in S\} = \{w \mid vw \in S\}.$$

Consider the quotient set $S/\sim_S$ and $\Sigma^*/\sim_S$, we have

Theorem
(Kamae and Weiss 1975[2]) Let S be a selection rule. If $S/\sim_S$ is a finite set, then for any normal number x, if digits in x get selected "frequently", *then the resulting subsequence, x_S, is also normal.*

Note that if $\Sigma^*/\sim_S$ is also finite, then the whole selecting process can be implemented by a finite state machine.

[2] *Normal Numbers and Selection Rules*

Preservation of Normality: Kamae and Weiss's Approach

Let $S \subseteq \Sigma^*$ be a selection rule. We define an equivalence relation $\sim_S$ on Σ^* as follows:

$$u \sim_S v \quad \text{if} \quad \{w \mid uw \in S\} = \{w \mid vw \in S\}.$$

Consider the quotient set $S/\sim_S$ and $\Sigma^*/\sim_S$, we have

Theorem
(Kamae and Weiss 1975[2]) Let S be a selection rule. If $S/\sim_S$ is a finite set, then for any normal number x, if digits in x get selected "frequently", *then the resulting subsequence, x_S, is also normal.*

Note that if $\Sigma^*/\sim_S$ is also finite, then the whole selecting process can be implemented by a finite state machine.

Note that counterexamples can be found if the "frequent" condition does not meet.

[2] *Normal Numbers and Selection Rules*

Continued Fraction Normality

Continued Fractions

The continued fraction expansion of a real number $x \in [0, 1)$ is

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} = \langle a_1, a_2, a_3 \rangle,$$

where all the a_i's are positive integers.

For irrational numbers, this expansion is unique and infinite. For rational numbers, there are exactly two expansions, both finite, such as

$$\langle 2 \rangle = \frac{1}{2} = \frac{1}{1 + \frac{1}{1}} = \langle 1, 1 \rangle.$$

Normality for Continued Fractions

We say that a real number $x = \langle a_1, a_2, a_3, a_4, \cdots \rangle$ is continued fraction normal (CF-normal) if for every string $s = [d_1, d_2, \cdots, d_k]$ of positive integers we have

Normality for Continued Fractions

We say that a real number $x = \langle a_1, a_2, a_3, a_4, \cdots \rangle$ is continued fraction normal (CF-normal) if for every string $s = [d_1, d_2, \cdots, d_k]$ of positive integers we have

$$\lim_{n \to \infty} \frac{\#\{1 \leq i \leq n : a_{i+j-1} = d_j, 1 \leq j \leq k\}}{n} = \mu(C_s)$$

where

$$\mu(A) = \int_A \frac{\mathrm{d}x}{(1+x)\log 2}$$

and $C_s \subset [0,1)$ is the set of numbers whose first continued fraction digits are the string s.

Normality for Continued Fractions

We say that a real number $x = \langle a_1, a_2, a_3, a_4, \cdots \rangle$ is continued fraction normal (CF-normal) if for every string $s = [d_1, d_2, \cdots, d_k]$ of positive integers we have

$$\lim_{n\to\infty} \frac{\#\{1 \le i \le n : a_{i+j-1} = d_j, 1 \le j \le k\}}{n} = \mu(C_s)$$

where

$$\mu(A) = \int_A \frac{\mathrm{d}x}{(1+x)\log 2}$$

and $C_s \subset [0,1)$ is the set of numbers whose first continued fraction digits are the string s.

Note that C_s can be interpreted as an interval in [0,1). It $\mu(C_s)$ is the probability that s "should" occur. Why don't we use Lebesgue measure, but the above measure, i.e., Gauss measure?

Gauss Map and Gauss Measure

We define Gauss map $T : X \to X$ by

$$T(x) = \begin{cases} \frac{1}{x} \ mod \ 1 & if \ x \neq 0 \\ 0 & if \ x = 0. \end{cases}$$

Let

$$x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cdots}}} = \langle x_0, x_1, x_2, \cdots \rangle$$

We observe that

$$T(x) = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cdots}} = \langle x_1, x_2, \cdots \rangle.$$

Gauss Map and Gauss Measure: cont.

That is, Gauss map is the shift operation for continued fractions. Counting of occurrences of a string s in a continued fraction x can be expressed by shift operations. We can rewrite the CF-normal definition by

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{\chi_{C_s}(T^i(x))}{n} = \mu(C_s)$$

The Gauss map does not preserve Lebesgue measure. However it does preserve Gauss measure. Moreover, T (Gauss map) is ergodic with respect to Gauss measure.

The “paradox” of normality- continued fraction edition

Again, almost all real numbers are CF-normal.

However, again we do not know any commonly used mathematical constants that are CF-normal. (Oddly, e is know to *not* be CF-normal.)

Construction of CF-normal Numbers

Adler, Keane, and Smorodinsky (1981) [3] gave a simple construction of a CF-normal number in the following way.

First start with all the rationals in (0,1) arranged in the following way:

$$\frac{1}{2}, \quad \frac{1}{3}, \frac{2}{3}, \quad \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \quad \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \dots$$

Write out the (shorter) continued fraction expansion for these rationals:

$$\langle 2 \rangle, \quad \langle 3 \rangle, \langle 1, 2 \rangle, \quad \langle 4 \rangle, \langle 2 \rangle, \langle 1, 3 \rangle, \quad \langle 5 \rangle, \langle 2, 2 \rangle, \langle 1, 1, 2 \rangle, \langle 1, 4 \rangle, \cdots$$

Then concatenate all these finite strings into an infinite CF-normal number:

$$\langle 2, 3, 1, 2, 4, 2, 1, 3, 5, 2, 2, 1, 1, 2, 1, 4, \cdots \rangle$$

[3] A Construction of a Normal Number for the Continued Fraction Transformation

Copeland–Erdös's Idea Revisited

If one concatenates the continued fraction expansions of a "sufficiently dense" subsequence of the rationals

$$\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6} \dots$$

then the result is CF-normal.

The restriction to be "sufficiently dense"[4] is enough to cover the subsequence of rationals with square-free numerators and denominators,

$$\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{1}{6}, \frac{5}{6}, \frac{1}{7}, \dots$$

[4] Joseph Vandehey. *New normality constructions for continued fraction expansions, 2016.*

As it turns out, the primes no longer count as sufficiently dense, but tweaks to Adler, Keane, and Smorodinsky's method can still show that one can construct CF-normal numbers out of the sequence of rationals with prime numerator and denominator:

$$\frac{2}{3}, \frac{2}{5}, \frac{3}{5}, \frac{2}{7}, \frac{3}{7}, \frac{5}{7}, \frac{2}{11}, \frac{3}{11}, \frac{5}{11}, \frac{7}{11}, \ldots$$

Arithmetic Progressions does not preserves CF-Normality

A classical result due to Wall says that if $0.a_1a_2a_3\cdots$ is normal, then so is $0.a_k a_{m+k} a_{2m+k} a_{3m+k}\cdots$, for any positive integers k, m.

[5] *Continued fraction normality is not preserved along arithmetic progressions*

Arithmetic Progressions does not preserves CF-Normality

A classical result due to Wall says that if $0.a_1 a_2 a_3 \cdots$ is normal, then so is $0.a_k a_{m+k} a_{2m+k} a_{3m+k} \cdots$, for any positive integers k, m.

However, the same thing is not true for CF-normality.

Heersink and Vandehey[5] show that given a continued fraction expansion $\langle a_1, a_2, a_3, \cdots \rangle$ that is CF-normal, then for any integers $m \geq 2$, $k \geq 1$, the continued fraction $\langle a_k, a_{m+k}, a_{2m+k}, a_{3m+k}, \cdots \rangle$ will never be normal.

[5] *Continued fraction normality is not preserved along arithmetic progressions*

Thank you for your time!