Normal Numbers and Uniform Distribution

Bill Mance

"What is…?" Seminar July 02, 2009

Let λ denote Lebesgue measue on $[0, 1)$. λ has the property that for all intervals $I = [a, b)$

$$\lambda(I) = b - a$$

Definition 1 Any measurable $S \subset [0, 1)$ is said to be of **measure 0** if for all $\epsilon > 0$ there exists a cover $\{I_n\}_{n=1}^{\infty}$ of S such that all I_n are intervals and

$$\sum_{n=1}^{\infty} \lambda(I_n) < \epsilon$$

Proposition 1 $S = \mathbb{Q} \cap [0, 1)$ is of measure zero.

Proof : Since S is countable, we can write $S = \{q_1, q_2, \ldots\}$. Let $\epsilon > 0$ be given and define $I_n = [q_n - \frac{\epsilon}{2^{n+2}}, q_n + \frac{\epsilon}{2^{n+2}})$. Then clearly $q_n \in I_n$ and $\lambda(I_n) = \frac{\epsilon}{2^{n+1}}$ so $\{I_n\}$ is a cover of S. Lastly, we note that

$$\sum_{n=1}^{\infty} \lambda(I_n) = \sum_{n=1}^{\infty} \frac{\epsilon}{2^{n+1}} = \epsilon/2 < \epsilon$$

Definition 2 A set $S \subset [0, 1)$ is **nowhere dense** if the interior of the closure of S is empty.

Definition 3 A set $S \subset [0, 1)$ is **meagre** if it is a countable union of nowhere dense sets.

A set of measure zero may be interpreted as being small in a measure theoretic sense while a meagre set may be thought of as being small in a topological sense.

Let $x \in [0, 1)$. Then let

$$x = \sum_{n=1}^{\infty} \frac{d_n}{b^n} = 0.d_1 d_2 \ldots$$

be the unique b-ary expansion of x such that $d_n < b - 1$ infinitely often. We will assume without mention that all b-ary expansions have this property.

Definition 4 A **block of length k in base b** is an ordered k-tuple $(b_1, \ldots, b_k)$ of integers which can assume the values $0, 1, \ldots, b - 1$.

Definition 5 Let $x = .d_1 d_2 \ldots$ be the b-ary expansion of x and let $B = (b_1, \ldots, b_k)$ be any block of length k in base b. Then define

$$N_n(B, x) = |\{m \leq n : (d_m, d_{m+1}, \ldots, d_{m+k-1}) = (b_1, b_2, \ldots, b_k)\}|$$

Definition 6 A number $x \in [0, 1)$ is called a **normal number** if for all k and blocks B of length k

$$\lim_{n\to\infty} \frac{N_n(B, x)}{n} = b^{-k}$$

The set of numbers that are normal in all bases is of full measure. However, the set of normal numbers in some base b is meagre. In fact more is true: the set of non-normal numbers has full hausdorff dimension.

First example due to Sierpinski in 1917. The most familiar example is due to Champernowne in 1933. The number

$$0.1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ldots$$

is normal in base 10. Copeland and Erdos showed in 1946 that

$$0.2\ 3\ 5\ 7\ 11\ 13\ 17\ 19\ 23\ 29\ 31\ldots$$

is normal in base 10.

Let $\{x\}$ denote the fractional part of a real number x. Let $X = \{x_n\}_{n=1}^{\infty}$ denote a sequence of real numbers.

Definition 7 Given a sequence X and an interval $I = [a, b) \subset [0, 1)$, define

$$A_n(I, X) = |\{m \leq n : \{x_m\} \in I\}|$$

Definition 8 Let ν be a probability measure on $[0, 1)$. A sequence X is said to be **ν-uniformly distributed mod 1** if

$$\lim_{n\to\infty} \frac{A_n(I, X)}{n} = \nu(I)$$

for all intervals $I \subset [0, 1)$. If $\nu = \lambda$ then X is said to be **uniformly distributed mod 1**.

Proposition 2 A sequence X is uniformly distributed mod 1 if and only if for all integers $h \neq 0$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n} = 0$$

One may note that if $x = 0.d_1d_2\ldots$ then for all n,

$$x \in \Big[\frac{d_1}{b^1} + \frac{d_2}{b^2} + \ldots + \frac{d_n}{b^n}, \frac{d_1}{b^1} + \frac{d_2}{b^2} + \ldots + \frac{d_n + 1}{b^n}\Big)$$

Proposition 3 If we define the shift operator $T_b = bx \bmod 1$ then x is normal in base b if and only if the sequence $\{T_b^n x\}_{n=0}^{\infty}$ is uniformly distributed mod 1.

The following was proven by Davenport and Erdos in 1952

Proposition 4 Let $p(n)$ be any increasing integer valued polynomial. Then the concatenation $0.p(1)p(2)p(3)\ldots$ is normal to the base in which it is expressed.

Definition 9 For natural numbers $a_j \geq 1$, define

$$[a_1, a_2, \ldots, a_n] = \cfrac{1}{a_1 + \frac{1}{a_2 + \ldots \frac{1}{a_n}}}$$

and $[a_1, a_2, \ldots] = \lim_{n\to\infty}[a_1, a_2, \ldots, a_n]$. Given $x \in [0, 1)$, $x = [a_1, a_2, \ldots]$ or $x = [a_1, a_2, \ldots, a_n]$ is called the **continued fraction representation** of x.

Any rational number will have two finite continued fraction representations. We will always pick the shorter representation. The representation is infinite if and only if x is irrational.

Define the probability measure μ on measurable sets A as follows

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

Definition 10 If we define the shift operator $Tx = \{\frac{1}{x}\}$ then x is **normal** with respect to the continued fraction representation if $\{T^n x\}_{n=0}^{\infty}$ is μ-uniformly distributed mod 1.

Example due to Adler, Keane and Smorodinsky (1981): concatenate the continued fraction representations of the rational numbers $1/2, 1/3, 2/3, 1/4, 2/4, 3/4, \ldots$.