## 3.11 - Walsh functions

Lemma 3.27 - for all $n \in \mathbb{N}$, $\exists$ unique $n_1, n_2 \dots n_k \in \mathbb{Z}$ s.t. $0 \le n_1 < n_2 \dots n_k$, and $n = 2^{n_1} + 2^{n_2} + \dots + 2^{n_k}$.

and, $n$ is odd iff $n_1 = 0$.

Defn. of Walsh function - $r$ = rademacher, $n$ = above $n$

$$W_n = r_{n_1} r_{n_2} \dots r_{n_k}$$

## 3.12 - Normal Numbers & Randomness

Normal Numbers Theorem: any binary sequence of length $r$ will occur in the proportion $2^{-r}$.

- can predict occurences but not exact positions

A measure-zero subset of $\mathbb{Z}$ has the following property:

$$\lim_{n \to \infty} \frac{1}{n} \left| \left\{ 1 \le r \le n,\ d_r(x) = c_1 \dots d_{r+1} = c_2 \dots d_{r+s-1}(x) = c_s \right\} \right| = \frac{1}{2^s}$$

i.e. the proportion of places $r$ where the sequence $d_r(x), d_{r+1}(x), \dots d_{r+s-1}(x)$ matches $c_1, c_2, \dots c_s$

converges to $\frac{1}{2^s}$ as $n \to \infty$.

## 3.13 - notions of "Probability" and "randomness"

1) Kolmogorov's Axiomatic Approach (classic probability)

Probability of an event is assigned in advance, often measure-based

- ex. probability of selecting a number from a subset of $[0,1]$ is the length of the subset.

2) Von Mises Frequency Approach (empirical probability)

Probability should be derived from observed data rather than assumed.

"collective" - infinite sequence of observations where probabilities are computed as relative frequencies.

ex. probability of a sequence for a subset $A$ =

$$r(A) = \lim_{n \to \infty} \frac{1}{n} \left| \left\{ 1 \le j \le n,\ c_j \in A \right\} \right|$$

Borel's theorem is a mix of Kolmogorov & Von Mises

- Frequency based (von Mises) = measure theoretic (Kolmogorov)

Some sequences appear random but aren't. ex. $\frac{2}{3} = 0.101010\dots$ = random at first level ($\frac{1}{2}$ frequency of 0s and 1s) but only 01 or 10 occurs, meaning not random.

"partial randomness"

NNT = outside of some measure zero set, binary expansion of $[0,1)$ has random behavior at <u>all</u> levels of "refinement".

## 3.14 – leading significant digit phenomenon ☆ potential talk topic

Leading digits of naturally occuring numerical data are NOT uniformly distributed –

follows a logarithmic pattern, Benford's Law.

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right)$$

Initially observed by Simon Newcomb in 1881

- numbers appearing in nature often appear as ratios of other numbers.

Weyl's theorem helped formalize this. – Fractional parts of logarithms leads to a uniform distributions

- if data is generated through multiplication, division, or exponential growth, it will naturally converge to Benford's distribution.

Applications – fraud detection, assessing whether datasets are artificial or not.

The existence & persistence of Benford's Law suggests some deeper underlying mathematical principle

## 3.15 – Leading digits in Geometric Sequences

2, 4, 8, 16, ...

leading = 2, 4, 8, 1, 3, 6, 1, 2, 5 ...

What proportion of terms in this sequence begin with each digit?

Gelfand's Theorem (I.M. Gelfand) –

For the geometric sequence $(2^n)$, the proportion of numbers with leading digit $d$ follows the same formula as Benford's Law.

contrast with Borel's – $\frac{1}{2}$ probability of 1 and 0s

Proof – taking logarithms of numbers

$$x = b^{n-1}(c_1 + y)$$

leading digit of $x$ = $\text{int}\left(b^{\log_b(x)}\right)$

extends to all bases not just base 10.

## 3.16 – Distributions of multiple digits (non-leading ones)

Probability that a number starts with a sequence, not just single digits?

Persi Diaconis' theorem – generalized Benford's law to multiple digits, for geometric sequences

$$P(c_1 c_2 \ldots c_r) = \log_b\left(1 + \frac{1}{c_1 c_2 \ldots c_r}\right)$$

Digits are not independent. Appearance of a second digit depends on the first, in a geometric sequence.

Shows a general mathematical principle even when extended to multiple leading digits.

## 3.17 - Dynamical Systems, changes of scale

e.g. river lengths in miles vs. km should both follow Benford's Law.

Scaling numbers - multiplying all values by a constant

- Analyze scaling numbers using Mantissa function: $\operatorname{mant}(x) = \frac{x}{b^n}$, $n$ is chosen so that $\frac{x}{b^n} \in [1, b)$.

  $\Rightarrow x = b^n \times \operatorname{mant}(x)$.

  Mantissa carries significant digits, discards magnitude.

  $0.000314159 \Rightarrow 3.14159$

Map from $(0, \infty)$ to $[1, b)$ using mantissa.

$T_\alpha(x) = \operatorname{mant}(\alpha x)$

Benford's Law remains true after a transformation.

Scale Invariance of Benford's Law: applying a scale transformation to all numbers in a sequence of digits preserves Benford's Law.

## 3.18 - Equivalence of Kronecker & Benford Systems

## Ch. 4 - developing "probability" in a formal sense

Kac's theorem -

$S$ - bounded interval $\quad f: S \to S$ - length-preserving interval

$\mathcal{U}$ - basis (finite union of intervals in $S$)

preimage - if $I$ is an interval, to be length-preserving,

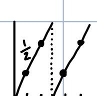
for when $f$ is not injective

Assume $\exists Z$, $\mu(Z) = 0$ s.t.

$\mu(I) = \mu(f^{-1}(I))$

if $x \in S - Z$, $\exists n \in \mathbb{N}$ s.t.

e.g. $x \to 2x \bmod 1$



$\mu(f^{-1}(\frac{1}{4}, \frac{3}{4})) = \mu((\frac{1}{8}, \frac{3}{8}) \cup (\frac{5}{8}, \frac{7}{8}))$

$f^n(x) \in \mathcal{U}$

Define $\Theta_\mathcal{U} : S \to \mathbb{N} \cup \{\infty\}$

Return time map

$$\Theta_\mathcal{U}(x) = \begin{cases} k & \text{if } k = \min\{m \mid f^m(x) \in \mathcal{U}\} \\ \infty & \text{otherwise, if } x \text{ never goes back to } \mathcal{U} \end{cases}$$

$\frac{1}{\mu(U)} \cdot \left( \sum_{n=1}^{\infty} n \cdot \mu\left(\left\{ x \in U \mid \theta_U(x) = n \right\}\right) \right) = \frac{\mu(S)}{\mu(U)}$ averaging the partitions of $U$

flipped; normally $A \subseteq B \to \frac{\mu(A)}{\mu(B)}$

does NOT depend on $f$!

Do more research on Kac's theorem

"Measure":

$\mu$ - subset of $X$

Want to return a "size" (nonnegative $\mathbb{R}$)

Certain sets - very weird, ex. techniques to pick out subsets of $\mathbb{R}$ where you can't measure them.

properties of measure, $\mu$:

if $A \subseteq B$, $\mu(A) \leq \mu(B)$

if $\underbrace{A \cap B = \emptyset}_{\text{disjoint}}$, $\mu(A \cup B) = \mu(A) + \mu(B)$

in $\mathbb{R}$ only, $\mu(x + A) = \mu(x)$ — translation

not in the circle