

Quantifying Equidistribution

A Survey of Finite-State Dimensions

Jack Lutz

Iowa State University

Equidistribution : Arithmetic, Computational, and Probabilistic Aspects
Institute for Mathematical Sciences
National University of Singapore
2019

Very Big Picture

Very Big Picture

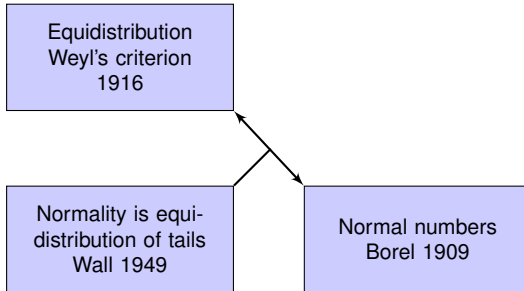
Equidistribution
Weyl's criterion
1916

Very Big Picture

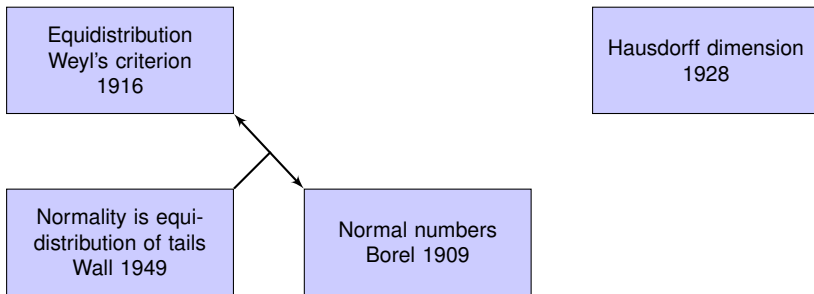
Equidistribution
Weyl's criterion
1916

Normal numbers
Borel 1909

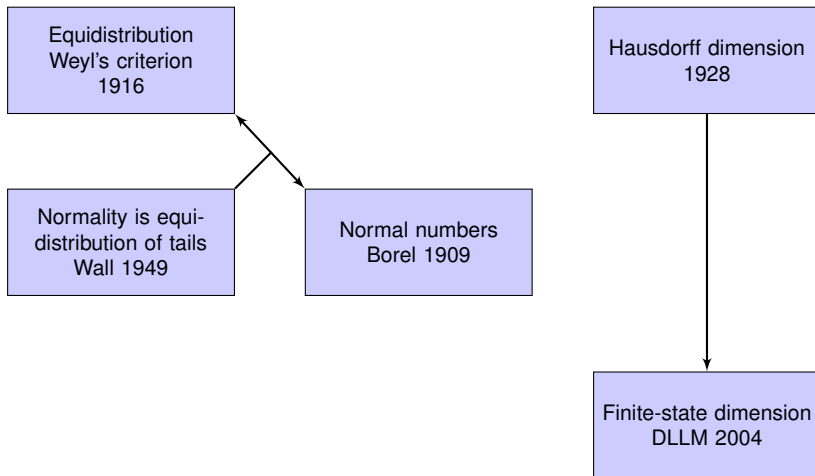
Very Big Picture



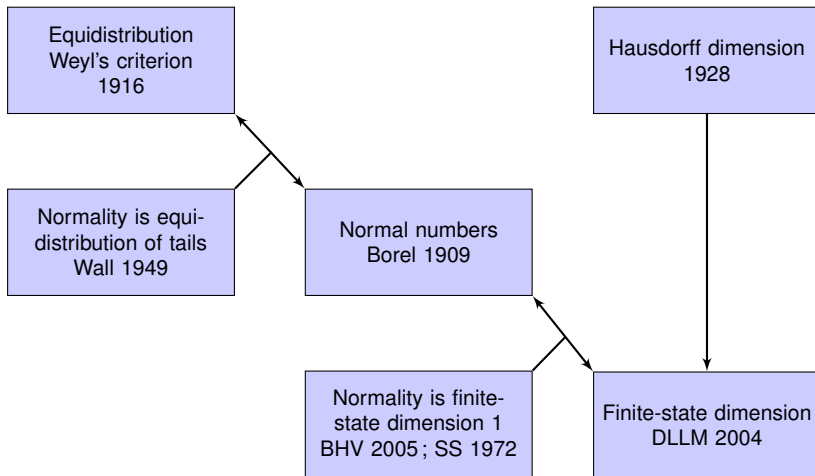
Very Big Picture



Very Big Picture



Very Big Picture



Equidistribution

A sequence $\alpha = (\alpha_n \mid n \in \mathbb{N})$ of real numbers is **uniformly distributed** (or **equidistributed**) **mod 1** if, for all $0 \leq a < b \leq 1$,

$$\lim_{m \rightarrow \infty} \frac{|\{n < m \mid \{\alpha_n\} \in [a, b]\}|}{m} = b - a$$

Equidistribution

A sequence $\alpha = (\alpha_n \mid n \in \mathbb{N})$ of real numbers is **uniformly distributed** (or **equidistributed**) **mod 1** if, for all $0 \leq a < b \leq 1$,

$$\lim_{m \rightarrow \infty} \frac{|\{n < m \mid \{\alpha_n\} \in [a, b]\}|}{m} = b - a$$

Weyl's criterion (1916)

This holds if and only if, for all $h \in \mathbb{Z}^+$,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=0}^{m-1} e^{2\pi i h \alpha_n} = 0$$

Normality

$\Sigma = \{0, 1, \dots, b-1\}$ ($2 \leq b \in \mathbb{N}$)

For $S \in \Sigma^\omega$, $w \in \Sigma^+$, and $n \in \mathbb{Z}^+$,

$$\begin{aligned} \text{freq}_n(w, S) &= \frac{|\{i < n \mid S[i..i+|w|-1] = w\}|}{n} \\ &= n\text{-th frequency of } w \text{ in } S \end{aligned}$$

Normality

$\Sigma = \{0, 1, \dots, b-1\}$ ($2 \leq b \in \mathbb{N}$)
 For $S \in \Sigma^\omega$, $w \in \Sigma^+$, and $n \in \mathbb{Z}^+$,

$$\text{freq}_n(w, S) = \frac{|\{i < n \mid S[i..i+|w|-1] = w\}|}{n}$$

= n -th frequency of w in S

Borel (1909)

A sequence $S \in \Sigma^\omega$ is **normal** if

$$(\forall w \in \Sigma^+) \lim_{n \rightarrow \infty} \text{freq}_n(w, S) = b^{-|w|}.$$

Normality

$\Sigma = \{0, 1, \dots, b-1\}$ ($2 \leq b \in \mathbb{N}$)
 For $S \in \Sigma^\omega$, $w \in \Sigma^+$, and $n \in \mathbb{Z}^+$,

$$\text{freq}_n(w, S) = \frac{|\{i < n \mid S[i..i+|w|-1] = w\}|}{n}$$

= n -th frequency of w in S

Borel (1909)

A sequence $S \in \Sigma^\omega$ is **normal** if

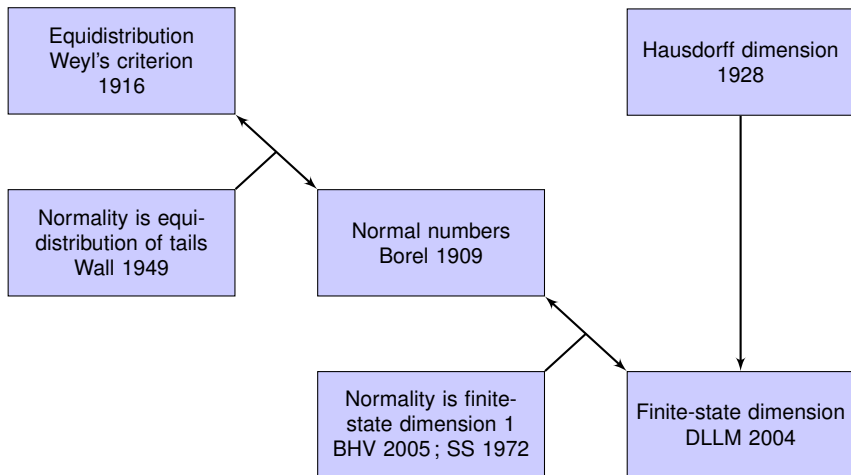
$$(\forall w \in \Sigma^+) \lim_{n \rightarrow \infty} \text{freq}_n(w, S) = b^{-|w|}.$$

A real number α is **normal in base** b if the base- b expansion of $\{\alpha\}$ is a normal sequence.

Wall (1949)

A real number α is normal in base b if and only if the sequence $(b^n \alpha \mid n \in \mathbb{N})$ is uniformly distributed mod 1.

Very Big Picture



Hausdorff dimension (1928)

Assigns a dimension $\dim_H(E)$ to every subset E of a given metric space.

Hausdorff dimension (1928)

Assigns a dimension $\dim_H(E)$ to every subset E of a given metric space.

Here we focus on the metric spaces Σ^ω equipped with the metric

$$d(S, T) = b^{-\min\{n \mid S[n] \neq T[n]\}}.$$

Today we use a very nonclassical characterization of classical Hausdorff dimension.

Today we use a very nonclassical characterization of classical Hausdorff dimension.

Fix a universal oracle Turing machine \mathcal{U} . The **Kolmogorov complexity** of a string $w \in \Sigma^*$ **relative to** an oracle $A \subseteq \mathbb{N}$ is

$$K^A(w) = \min \{ |\pi| \mid \pi \in \{0, 1\}^* \text{ and } \mathcal{U}^A(\pi) = w \}.$$

Today we use a very nonclassical characterization of classical Hausdorff dimension.

Fix a universal oracle Turing machine \mathcal{U} . The **Kolmogorov complexity** of a string $w \in \Sigma^*$ **relative to** an oracle $A \subseteq \mathbb{N}$ is

$$K^A(w) = \min \{ |\pi| \mid \pi \in \{0, 1\}^* \text{ and } \mathcal{U}^A(\pi) = w \}.$$

The **Kolmogorov complexity** of a string $w \in \Sigma^*$ is

$$K(w) = K^\emptyset(w).$$

The **lower** and **upper algorithmic dimensions** of a sequence $S \in \Sigma^\omega$ are

$$\dim(S) = \liminf_{w \rightarrow S} \frac{K(w)}{|w| \log |\Sigma|} \quad (\text{Lutz 2003, Mayordomo 2002})$$

and

$$\text{Dim}(S) = \limsup_{w \rightarrow S} \frac{K(w)}{|w| \log |\Sigma|}, \quad (\text{Athreya, Hitchcock, Lutz, Mayordomo 2007})$$

respectively.

The **lower** and **upper algorithmic dimensions** of a sequence $S \in \Sigma^\omega$ are

$$\dim(S) = \liminf_{w \rightarrow S} \frac{K(w)}{|w| \log |\Sigma|} \quad (\text{Lutz 2003, Mayordomo 2002})$$

and

$$\text{Dim}(S) = \limsup_{w \rightarrow S} \frac{K(w)}{|w| \log |\Sigma|}, \quad (\text{Athreya, Hitchcock, Lutz, Mayordomo 2007})$$

respectively.

In general, $0 \leq \dim(S) \leq \text{Dim}(S) \leq 1$.

The **lower** and **upper algorithmic dimensions** of a sequence $S \in \Sigma^\omega$ are

$$\dim(S) = \liminf_{w \rightarrow S} \frac{K(w)}{|w| \log |\Sigma|} \quad (\text{Lutz 2003, Mayordomo 2002})$$

and

$$\text{Dim}(S) = \limsup_{w \rightarrow S} \frac{K(w)}{|w| \log |\Sigma|}, \quad (\text{Athreya, Hitchcock, Lutz, Mayordomo 2007})$$

respectively.

In general, $0 \leq \dim(S) \leq \text{Dim}(S) \leq 1$.

Define the relativized algorithmic dimensions $\dim^A(S)$ and $\text{Dim}^A(S)$ analogously.

Today we use the following theorem as the definition of Hausdorff dimension.

Point-to-Set Principle (N. Lutz and J. Lutz 2018)

For every $E \subseteq \Sigma^\omega$

$$\dim_H(E) = \min_{A \in \mathbb{N}} \sup_{S \in E} \dim^A(S).$$

Today we use the following theorem as the definition of Hausdorff dimension.

Point-to-Set Principle (N. Lutz and J. Lutz 2018)

For every $E \subseteq \Sigma^\omega$

$$\dim_H(E) = \min_{A \in \mathbb{N}} \sup_{S \in E} \dim^A(S).$$

Also holds with \mathbb{R}^n in place of Σ^ω .

Today we use the following theorem as the definition of Hausdorff dimension.

Point-to-Set Principle (N. Lutz and J. Lutz 2018)

For every $E \subseteq \Sigma^\omega$

$$\dim_H(E) = \min_{A \in \mathbb{N}} \sup_{S \in E} \dim^A(S).$$

Also holds with \mathbb{R}^n in place of Σ^ω .

Allows one to prove a lower bound on $\dim_H(E)$ by proving that a judiciously chosen point $x \in E$ has large $\dim^A(x)$.

Today we use the following theorem as the definition of Hausdorff dimension.

Point-to-Set Principle (N. Lutz and J. Lutz 2018)

For every $E \subseteq \Sigma^\omega$

$$\dim_H(E) = \min_{A \in \mathbb{N}} \sup_{S \in E} \dim^A(S).$$

Also holds with \mathbb{R}^n in place of Σ^ω .

Allows one to prove a lower bound on $\dim_H(E)$ by proving that a judiciously chosen point $x \in E$ has large $\dim^A(x)$.

N. Lutz and Stull (2017) and N. Lutz (2017) have used this method to solve open problems in classical geometric measure theory.

But algorithmic dimensions are not effective enough to study normality.

But algorithmic dimensions are not effective enough to study normality.

Dai, Lathrop, Lutz, and Mayordomo (2004) developed the **lower** and **upper finite-state dimensions** $\dim_{\text{FS}}(S)$ and $\text{Dim}_{\text{FS}}(S)$ of sequences $S \in \Sigma^\omega$.

But algorithmic dimensions are not effective enough to study normality.

Dai, Lathrop, Lutz, and Mayordomo (2004) developed the **lower** and **upper finite-state dimensions** $\dim_{\text{FS}}(S)$ and $\text{Dim}_{\text{FS}}(S)$ of sequences $S \in \Sigma^\omega$.

Several equivalent characterizations are now known. Today we use the compression characterization.

Shannon (1948)

A **finite-state compressor (FSC)** is a 5-tuple

$$C = (Q, \Sigma, \delta, s, v),$$

where

- Q is a finite set of **states**;
- $\Sigma = \{0, 1, \dots, b-1\}$ is the **input alphabet**;
- $\delta : Q \times \Sigma \rightarrow Q$ is the **transition function**;
- $s \in Q$ is the **start state**; and
- $v : Q \times \Sigma \rightarrow \{0, 1\}^*$ is the **output function**.

Write $C(w)$ for the **cumulative output** of C on an input $w \in \Sigma^*$.

An **information-lossless finite-state compressor (ILFSC)** is an FSC $C = (Q, \Sigma, \delta, s, \nu)$ for which the function

$$\begin{aligned}\Sigma^* &\rightarrow \{0, 1\}^* \times Q \\ w &\mapsto (C(w), \delta(s, w))\end{aligned}$$

is one-to-one.

An **information-lossless finite-state compressor (ILFSC)** is an FSC $C = (Q, \Sigma, \delta, s, \nu)$ for which the function

$$\begin{aligned} \Sigma^* &\rightarrow \{0, 1\}^* \times Q \\ w &\mapsto (C(w), \delta(s, w)) \end{aligned}$$

is one-to-one.

Today's definitions :

$$\dim_{\text{FS}}(S) = \inf_C \liminf_{w \rightarrow S} \frac{C(w)}{|w| \log |\Sigma|}. \quad (\text{Dai, Lathrop, Lutz, Mayordomo 2004})$$

$$\text{Dim}_{\text{FS}}(S) = \inf_C \limsup_{w \rightarrow S} \frac{C(w)}{|w| \log |\Sigma|}. \quad (\text{Athreya, Hitchcock, Lutz, Mayordomo 2007})$$

(C ranges over all ILFSCs.)

An **information-lossless finite-state compressor (ILFSC)** is an FSC $C = (Q, \Sigma, \delta, s, \nu)$ for which the function

$$\begin{aligned} \Sigma^* &\rightarrow \{0, 1\}^* \times Q \\ w &\mapsto (C(w), \delta(s, w)) \end{aligned}$$

is one-to-one.

Today's definitions :

$$\dim_{\text{FS}}(S) = \inf_C \liminf_{w \rightarrow S} \frac{C(w)}{|w| \log |\Sigma|}. \quad (\text{Dai, Lathrop, Lutz, Mayordomo 2004})$$

$$\text{Dim}_{\text{FS}}(S) = \inf_C \limsup_{w \rightarrow S} \frac{C(w)}{|w| \log |\Sigma|}. \quad (\text{Athreya, Hitchcock, Lutz, Mayordomo 2007})$$

(C ranges over all ILFSCs.)

In general, $0 \leq \dim_{\text{FS}}(S) \leq \text{Dim}_{\text{FS}}(S) \leq 1$.

Schnorr & Stimm (1972); Bourke, Hitchcock, & Vinodchandren (2005) :

A sequence $S \in \Sigma^\omega$ is normal if and only if $\dim_{\text{FS}}(S) = 1$.

Schnorr & Stimm (1972); Bourke, Hitchcock, & Vinodchandren (2005) :

A sequence $S \in \Sigma^\omega$ is normal if and only if $\dim_{\text{FS}}(S) = 1$.

Schnorr and Stimm's method makes it clear that normality **is** finite-state randomness.

Schnorr & Stimm (1972); Bourke, Hitchcock, & Vinodchandren (2005) :

A sequence $S \in \Sigma^\omega$ is normal if and only if $\dim_{\text{FS}}(S) = 1$.

Schnorr and Stimm's method makes it clear that normality **is** finite-state randomness.

The finite-state world is the only one that we know of where dimension 1 implies randomness.

Schnorr & Stimm (1972); Bourke, Hitchcock, & Vinodchandren (2005) :

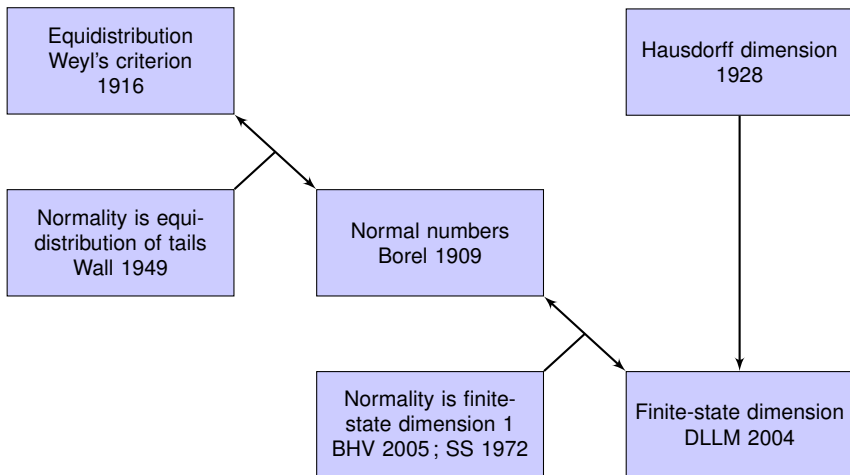
A sequence $S \in \Sigma^\omega$ is normal if and only if $\dim_{\text{FS}}(S) = 1$.

Schnorr and Stimm's method makes it clear that normality **is** finite-state randomness.

The finite-state world is the only one that we know of where dimension 1 implies randomness.

Elvira will tell you about very recent work extending Schnorr and Stimm's remarkable theorem.

Very Big Picture



We now know that a sequence $S \in \Sigma^\omega$ is normal if and only if $\dim_{\text{FS}}(S) = 1$.

We now know that a sequence $S \in \Sigma^\omega$ is normal if and only if $\dim_{FS}(S) = 1$.

QUESTION : Which theorems about normality can be better understood as the dimension-1 special case of quantitative theorems about finite-state dimension ?

EXAMPLE 1 : Copeland-Erdős sequences

$$2 \leq b \in \mathbb{N}, \Sigma = \{0, 1, \dots, b-1\}.$$

For $n \in \mathbb{Z}^+$, $\sigma_b(n)$ = the base- b expansion of n .

EXAMPLE 1 : Copeland-Erdős sequences

$$2 \leq b \in \mathbb{N}, \Sigma = \{0, 1, \dots, b-1\}.$$

For $n \in \mathbb{Z}^+$, $\sigma_b(n)$ = the base- b expansion of n .

The **base- b Copeland-Erdős sequence** of an infinite set

$$A = \{a_1 < a_2 < \dots\} \subseteq \mathbb{Z}^+$$

is the sequence $CE_b(A) = \sigma_b(a_1)\sigma_b(a_2)\dots \in \Sigma^\omega$.

Champernowne (1933)

The *decimal Champernowne sequence*

$$CE_{10}(\mathbb{Z}^+) = 123456789101112 \dots$$

is normal.

Champernowne (1933)

The *decimal Champernowne sequence*

$$CE_{10}(\mathbb{Z}^+) = 123456789101112 \dots$$

is normal.

Champernowne conjectured that

$$CE_{10}(PRIMES) = 2357111317 \dots$$

is also normal.

Copeland-Erdős (1946)

*For all b ,
 $CE_b(\text{PRIMES})$ is normal.*

Copeland-Erdős (1946)

For all b ,
 $CE_b(\text{PRIMES})$ is normal.

Outline of proof

- 1 For all sufficiently dense $A \subseteq \mathbb{Z}^+$,
 $CE_b(A)$ is normal.
- 2 PRIMES is sufficiently dense. \square

Zeta dimensions (invented **many** times over the past 200 years)

For $A \subseteq \mathbb{Z}^+$, define the **A-zeta function**

$$\zeta_A : [0, \infty) \rightarrow [0, \infty]$$

$$\zeta_A(s) = \sum_{n \in A} n^{-s}.$$

Zeta dimensions (invented **many** times over the past 200 years)

For $A \subseteq \mathbb{Z}^+$, define the **A-zeta function**

$$\zeta_A : [0, \infty) \rightarrow [0, \infty]$$

$$\zeta_A(s) = \sum_{n \in A} n^{-s}.$$

The **zeta-dimension** of A is

$$\text{Dim}_\zeta(A) = \inf\{s \mid \zeta_A(s) < \infty\}.$$

Zeta dimensions (invented **many** times over the past 200 years)

For $A \subseteq \mathbb{Z}^+$, define the **A-zeta function**

$$\zeta_A : [0, \infty) \rightarrow [0, \infty]$$

$$\zeta_A(s) = \sum_{n \in A} n^{-s}.$$

The **zeta-dimension** of A is

$$\text{Dim}_\zeta(A) = \inf\{s \mid \zeta_A(s) < \infty\}.$$

Cahen (1894)

$$\text{Dim}_\zeta(A) = \limsup_{n \rightarrow \infty} \frac{\log |A \cap \{1, \dots, n\}|}{\log n}$$

Zeta dimensions (invented **many** times over the past 200 years)

For $A \subseteq \mathbb{Z}^+$, define the **A-zeta function**

$$\zeta_A : [0, \infty) \rightarrow [0, \infty]$$

$$\zeta_A(s) = \sum_{n \in A} n^{-s}.$$

The **zeta-dimension** of A is

$$\text{Dim}_\zeta(A) = \inf\{s \mid \zeta_A(s) < \infty\}.$$

Cahen (1894)

$$\text{Dim}_\zeta(A) = \limsup_{n \rightarrow \infty} \frac{\log|A \cap \{1, \dots, n\}|}{\log n}$$

The **lower zeta-dimension** of A is

$$\text{dim}_\zeta(A) = \liminf_{n \rightarrow \infty} \frac{\log|A \cap \{1, \dots, n\}|}{\log n}$$

Zeta dimensions (invented **many** times over the past 200 years)

For $A \subseteq \mathbb{Z}^+$, define the **A-zeta function**

$$\zeta_A : [0, \infty) \rightarrow [0, \infty]$$

$$\zeta_A(s) = \sum_{n \in A} n^{-s}.$$

The **zeta-dimension** of A is

$$\text{Dim}_\zeta(A) = \inf\{s \mid \zeta_A(s) < \infty\}.$$

Cahen (1894)

$$\text{Dim}_\zeta(A) = \limsup_{n \rightarrow \infty} \frac{\log|A \cap \{1, \dots, n\}|}{\log n}$$

The **lower zeta-dimension** of A is

$$\dim_\zeta(A) = \liminf_{n \rightarrow \infty} \frac{\log|A \cap \{1, \dots, n\}|}{\log n}$$

Clearly, $0 \leq \dim_\zeta(A) \leq \text{Dim}_\zeta(A) \leq 1$.

Gu, Lutz, and Moser (2007)

For every infinite $A \subseteq \mathbb{Z}^+$,

$$\dim_{FS}(CE_b(A)) \geq \dim_{\zeta}(A)$$

and

$$\text{Dim}_{FS}(CE_b(A)) \geq \text{Dim}_{\zeta}(A)$$

Gu, Lutz, and Moser (2007)

For every infinite $A \subseteq \mathbb{Z}^+$,

$$\dim_{FS}(CE_b(A)) \geq \dim_{\zeta}(A)$$

and

$$\text{Dim}_{FS}(CE_b(A)) \geq \text{Dim}_{\zeta}(A)$$

These inequalities are tight : For any real numbers

$$\begin{array}{ccccccc}
 & & \gamma & \leq & \delta & \leq & 1 \\
 & & \vee & & \vee & & \\
 0 & \leq & \alpha & \leq & \beta & &
 \end{array}$$

there exists an infinite $A \subseteq \mathbb{Z}^+$ with $\dim_{\zeta}(A) = \alpha$, $\text{Dim}_{\zeta}(A) = \beta$, $\dim_{FS}(CE_b(A)) = \gamma$, and $\text{Dim}_{FS}(CE_b(A)) = \delta$.

Gu, Lutz, and Moser (2007)

For every infinite $A \subseteq \mathbb{Z}^+$,

$$\dim_{FS}(\text{CE}_b(A)) \geq \dim_{\zeta}(A)$$

and

$$\text{Dim}_{FS}(\text{CE}_b(A)) \geq \text{Dim}_{\zeta}(A)$$

These inequalities are tight : For any real numbers

$$0 \leq \underset{\forall I}{\alpha} \leq \underset{\forall I}{\beta} \leq \underset{\forall I}{\gamma} \leq \underset{\forall I}{\delta} \leq 1$$

there exists an infinite $A \subseteq \mathbb{Z}^+$ with $\dim_{\zeta}(A) = \alpha$, $\text{Dim}_{\zeta}(A) = \beta$, $\dim_{FS}(\text{CE}_b(A)) = \gamma$, and $\text{Dim}_{FS}(\text{CE}_b(A)) = \delta$.

Remarks :

- The Copeland-Erdős density criterion is equivalent to $\dim_{\zeta}(A) = 1$.
- The dimension theorem required a different method.

EXAMPLE 2 : Finite-state dimension and real arithmetic.

Wall (1949)

*If q is a non-zero rational, then,
for every real number α ,*

α is normal in base b

$\implies q + \alpha$ and $q\alpha$ are normal in base b .

EXAMPLE 2 : Finite-state dimension and real arithmetic.**Wall (1949)**

*If q is a non-zero rational, then,
for every real number α ,
 α is normal in base b
 $\implies q + \alpha$ and $q\alpha$ are normal in base b .*

Doty, Lutz, and Nandakumar (2007)

If q is a non-zero rational, then, for every real number α , the base- b expansions of α , $q + \alpha$, and $q\alpha$ all have the same finite-state dimension and the same strong finite-state dimension.

EXAMPLE 2 : Finite-state dimension and real arithmetic.**Wall (1949)**

*If q is a non-zero rational, then,
for every real number α ,
 α is normal in base b
 $\implies q + \alpha$ and $q\alpha$ are normal in base b .*

Doty, Lutz, and Nandakumar (2007)

If q is a non-zero rational, then, for every real number α , the base- b expansions of α , $q + \alpha$, and $q\alpha$ all have the same finite-state dimension and the same strong finite-state dimension.

The dimension result required a **very** different method.

I'm asking again, and this time I'm asking you !

QUESTION : Which theorems about normality can be better understood as the dimension-1 special case of quantitative theorems about finite-state dimension ?

Further Directions

- (1) It is routine to generalize normality to α -normality, where α is a probability measure on Σ . One can also define finite-state dimensions with respect to such α .

Further Directions

- (1) It is routine to generalize normality to α -normality, where α is a probability measure on Σ . One can also define finite-state dimensions with respect to such α .

Lutz (2011)

If α and β are positive probability measures on Σ and $R \in \Sigma^\omega$ is α -normal, then

$$\dim_{FS}^\beta(R) = \text{Dim}_{FS}^\beta(R) = \frac{\mathcal{H}(\alpha)}{\mathcal{H}(\alpha) + \mathcal{D}(\alpha\|\beta)},$$

where

$$\mathcal{H} = E_\alpha \log \frac{1}{\alpha}, \mathcal{D}(\alpha\|\beta) = E_\alpha \log \frac{\alpha}{\beta}$$

are the Shannon entropy and the Kullback-Leibler divergence.

Further Directions

- (1) It is routine to generalize normality to α -normality, where α is a probability measure on Σ . One can also define finite-state dimensions with respect to such α .

Lutz (2011)

If α and β are positive probability measures on Σ and $R \in \Sigma^\omega$ is α -normal, then

$$\dim_{FS}^\beta(R) = \text{Dim}_{FS}^\beta(R) = \frac{\mathcal{H}(\alpha)}{\mathcal{H}(\alpha) + \mathcal{D}(\alpha\|\beta)},$$

where

$$\mathcal{H} = E_\alpha \log \frac{1}{\alpha}, \mathcal{D}(\alpha\|\beta) = E_\alpha \log \frac{\alpha}{\beta}$$

are the Shannon entropy and the Kullback-Leibler divergence.

Can this be usefully generalized?

- (2) In many contexts, it is natural to ask “how much randomness” a probability theorem requires. Classical examples include the law of large numbers and ergodic theorems. Current examples include stochastic chemical reaction networks and Banach-Tarski type constructions. The answers range from normality to algorithmic randomness and beyond. More general notions of normality, along lines that we’ve seen this week, may be needed to address such questions.

Thank You !