Random Bits and Pieces

An Introduction to Symbolic Dynamics

Davar Khoshnevisan

Department of Mathematics
University of Utah
http://www.math.utah.edu/~davar

Undergraduate Colloquium
University of Utah
November 15, 2006

Introduction

This is a talk about sequences of zeros and ones:

Introduction

This is a talk about sequences of zeros and ones:

- Random Number Generators;

Introduction

This is a talk about sequences of zeros and ones:

- Random Number Generators;
- Random search on a binary tree [philogenetic];

Introduction

This is a talk about sequences of zeros and ones:

- Random Number Generators;
- Random search on a binary tree [philogenetic];
- Binary encoding of numbers.

Introduction

This is a talk about sequences of zeros and ones:

- Random Number Generators;
- Random search on a binary tree [philogenetic];
- Binary encoding of numbers.

Introduction

This is a talk about sequences of zeros and ones:

- Random Number Generators;
- Random search on a binary tree [philogenetic];
- Binary encoding of numbers.

We work by examples, and in random order.

Binary Encoding of Numbers

- Let x be a number between zero and one.

Binary Encoding of Numbers

- Let x be a number between zero and one.
- We can write

$$x = \frac{x_1}{2} + \frac{x_2}{4} + \frac{x_3}{8} + \cdots$$
$$= \sum_{j=1}^{\infty} \frac{x_j}{2^j},$$

where $x_1, x_2, \ldots$ are either zero or one.

Binary Encoding of Numbers

- Let x be a number between zero and one.
- We can write

$$\begin{aligned} x &= \frac{x_1}{2} + \frac{x_2}{4} + \frac{x_3}{8} + \cdots \\ &= \sum_{j=1}^{\infty} \frac{x_j}{2^j}, \end{aligned}$$

  where $x_1, x_2, \ldots$ are either zero or one.
- If there are two ways of doing this [dyadic rationals] then opt for the non-terminating expansion.

Examples

- We might write $x = [x_1, x_2, \ldots]$ instead of $x = \sum_{j=1}^{\infty} 2^{-j} x_j$.

Examples

- We might write $x = [x_1, x_2, \ldots]$ instead of $x = \sum_{j=1}^{\infty} 2^{-j} x_j$.
- $0 = [0, 0, \ldots]$

Examples

- We might write $x = [x_1, x_2, \ldots]$ instead of $x = \sum_{j=1}^{\infty} 2^{-j} x_j$.
- $0 = [0, 0, \ldots]$
- $1 = [1, 1, \ldots]$ because $\sum_{j=1}^{\infty} 2^{-j} = 1$

Examples

- We might write $x = [x_1\,, x_2\,, \ldots]$ instead of $x = \sum_{j=1}^{\infty} 2^{-j} x_j$.
- $0 = [0\,, 0\,, \ldots]$
- $1 = [1\,, 1\,, \ldots]$ because $\sum_{j=1}^{\infty} 2^{-j} = 1$
- 0.5 can be written in two different ways.

Examples

- We might write $x = [x_1\,, x_2\,, \ldots]$ instead of $x = \sum_{j=1}^{\infty} 2^{-j} x_j$.
- $0 = [0\,, 0\,, \ldots]$
- $1 = [1\,, 1\,, \ldots]$ because $\sum_{j=1}^{\infty} 2^{-j} = 1$
- 0.5 can be written in two different ways.
  - Here is one:

$$0.5 = \frac{1}{2} + \frac{0}{4} + \frac{0}{8} + \cdots$$

Examples

- We might write $x = [x_1\,, x_2\,, \ldots]$ instead of $x = \sum_{j=1}^{\infty} 2^{-j} x_j$.
- $0 = [0\,, 0\,, \ldots]$
- $1 = [1\,, 1\,, \ldots]$ because $\sum_{j=1}^{\infty} 2^{-j} = 1$
- 0.5 can be written in two different ways.
  - Here is one:
    $$0.5 = \frac{1}{2} + \frac{0}{4} + \frac{0}{8} + \cdots$$
  - Here is another:
    $$0.5 = \frac{0}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$$
    This works because $\sum_{j=2}^{\infty} 2^{-j} = 1/2$.

Examples

- We might write $x = [x_1, x_2, \ldots]$ instead of $x = \sum_{j=1}^{\infty} 2^{-j} x_j$.
- $0 = [0, 0, \ldots]$
- $1 = [1, 1, \ldots]$ because $\sum_{j=1}^{\infty} 2^{-j} = 1$
- 0.5 can be written in two different ways.
  - Here is one:

$$0.5 = \frac{1}{2} + \frac{0}{4} + \frac{0}{8} + \cdots$$

  - Here is another:

$$0.5 = \frac{0}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$$

    This works because $\sum_{j=2}^{\infty} 2^{-j} = 1/2$.
  - Infinite-option convention yields:

$$0.5 = [0, 1, 1, \ldots].$$

An Algorithm for Finding the Digits

- Let x be a fixed number between zero and one

An Algorithm for Finding the Digits

- Let x be a fixed number between zero and one
- Ask twenty-twenty style:

An Algorithm for Finding the Digits

- Let x be a fixed number between zero and one
- Ask twenty-twenty style:
  - Is $x \leq 0.5$? If yes then $x_1 = 0$; else, $x_1 = 1$

An Algorithm for Finding the Digits

- Let x be a fixed number between zero and one
- Ask twenty-twenty style:
  - Is $x \leq 0.5$? If yes then $x_1 = 0$; else, $x_1 = 1$
  - Is $y_1 = 2(x - \frac{1}{2}x_1) \leq 0.5$? If yes then $x_2 = 0$; else, $x_2 = 1$

An Algorithm for Finding the Digits

- Let x be a fixed number between zero and one
- Ask twenty-twenty style:
  - Is $x \leq 0.5$? If yes then $x_1 = 0$; else, $x_1 = 1$
  - Is $y_1 = 2(x - \frac{1}{2}x_1) \leq 0.5$? If yes then $x_2 = 0$; else, $x_2 = 1$
  - Is $y_2 = 2(y_1 - \frac{1}{2}x_2) \leq 0.5$? If yes then $x_3 = 0$; else, $x_3 = 1$

An Algorithm for Finding the Digits

- Let x be a fixed number between zero and one
- Ask twenty-twenty style:
  - Is $x \leq 0.5$? If yes then $x_1 = 0$; else, $x_1 = 1$
  - Is $y_1 = 2(x - \frac{1}{2}x_1) \leq 0.5$? If yes then $x_2 = 0$; else, $x_2 = 1$
  - Is $y_2 = 2(y_1 - \frac{1}{2}x_2) \leq 0.5$? If yes then $x_3 = 0$; else, $x_3 = 1$
  - $\vdots$

An Algorithm for Finding the Digits

- Let x be a fixed number between zero and one
- Ask twenty-twenty style:
  - Is $x \leq 0.5$? If yes then $x_1 = 0$; else, $x_1 = 1$
  - Is $y_1 = 2(x - \frac{1}{2}x_1) \leq 0.5$? If yes then $x_2 = 0$; else, $x_2 = 1$
  - Is $y_2 = 2(y_1 - \frac{1}{2}x_2) \leq 0.5$? If yes then $x_3 = 0$; else, $x_3 = 1$
  - $\vdots$
- Why does this work? Hint:

$$y_1 = \sum_{j=1}^{\infty} \frac{x_{j+1}}{2^j}$$

Symbolic Dynamics [An Alternative]

- Split $[0\,,1]$ into two subintervals $[0\,,0.5]$ and $(0.5\,,1]$;

Symbolic Dynamics [An Alternative]

- Split $[0\,,1]$ into two subintervals $[0\,,0.5]$ and $(0.5\,,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;

Symbolic Dynamics [An Alternative]

- Split $[0\,,1]$ into two subintervals $[0\,,0.5]$ and $(0.5\,,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;
- Call whichever [dyadic] interval x fell in last I_1;

Symbolic Dynamics [An Alternative]

- Split $[0\,,1]$ into two subintervals $[0\,,0.5]$ and $(0.5\,,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;
- Call whichever [dyadic] interval x fell in last I_1;
- Split I_1 into two subintervals each half the length of I_1;

Symbolic Dynamics [An Alternative]

- Split $[0\,,1]$ into two subintervals $[0\,,0.5]$ and $(0.5\,,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;
- Call whichever [dyadic] interval x fell in last I_1;
- Split I_1 into two subintervals each half the length of I_1;
- If x falls in the left one $x_2 = 0$; if x falls in the right one $x_2 = 1$;

Symbolic Dynamics [An Alternative]

- Split $[0,1]$ into two subintervals $[0,0.5]$ and $(0.5,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;
- Call whichever [dyadic] interval x fell in last I_1;
- Split I_1 into two subintervals each half the length of I_1;
- If x falls in the left one $x_2 = 0$; if x falls in the right one $x_2 = 1$;
- Call whichever [dyadic] interval x fell in last I_2;

Symbolic Dynamics [An Alternative]

- Split $[0\,,1]$ into two subintervals $[0\,,0.5]$ and $(0.5\,,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;
- Call whichever [dyadic] interval x fell in last I_1;
- Split I_1 into two subintervals each half the length of I_1;
- If x falls in the left one $x_2 = 0$; if x falls in the right one $x_2 = 1$;
- Call whichever [dyadic] interval x fell in last I_2;
- $\vdots$

Symbolic Dynamics [An Alternative]

- Split $[0,1]$ into two subintervals $[0,0.5]$ and $(0.5,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;
- Call whichever [dyadic] interval x fell in last I_1;
- Split I_1 into two subintervals each half the length of I_1;
- If x falls in the left one $x_2 = 0$; if x falls in the right one $x_2 = 1$;
- Call whichever [dyadic] interval x fell in last I_2;
- $\vdots$
- Try it for $x = 0.5$ $\qquad$ $0.5 = [0,1,1,\ldots]$

Symbolic Dynamics [An Alternative]

- Split $[0\,,1]$ into two subintervals $[0\,,0.5]$ and $(0.5\,,1]$;
- If x falls in the left interval then $x_1 = 0$; if x falls in the right one then $x_1 = 1$;
- Call whichever [dyadic] interval x fell in last I_1;
- Split I_1 into two subintervals each half the length of I_1;
- If x falls in the left one $x_2 = 0$; if x falls in the right one $x_2 = 1$;
- Call whichever [dyadic] interval x fell in last I_2;
- $\vdots$
- Try it for $x = 0.5$ $\qquad 0.5 = [0\,,1\,,1\,,\ldots]$
- What if you split into $[0\,,0.5)$ and $[0.5\,,1]$ etc.?

Dyadic Intervals

- These are the intervals we obtained by subdividing.

Dyadic Intervals

- These are the intervals we obtained by subdividing.
- A dyadic interval is a subintervals of $[0\,,1]$ that has length 2^{-n} for some integer $n \geq 0$.

Dyadic Intervals

- These are the intervals we obtained by subdividing.
- A dyadic interval is a subintervals of $[0\,,1]$ that has length 2^{-n} for some integer $n \geq 0$.
- A dyadic interval of length $n \geq 1$ can be written as

$$\left[\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 0, 2, \ldots \text{ is even,}$$

Dyadic Intervals

- These are the intervals we obtained by subdividing.
- A dyadic interval is a subintervals of $[0\,,1]$ that has length 2^{-n} for some integer $n \geq 0$.
- A dyadic interval of length $n \geq 1$ can be written as

$$\left[\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 0, 2, \ldots \text{ is even,}$$

Dyadic Intervals

- These are the intervals we obtained by subdividing.
- A dyadic interval is a subintervals of $[0\,,1]$ that has length 2^{-n} for some integer $n \geq 0$.
- A dyadic interval of length $n \geq 1$ can be written as

$$\left[\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 0, 2, \ldots \text{ is even,}$$
$$\left(\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 1, 3, \ldots \text{ is odd}$$

Dyadic Intervals

- These are the intervals we obtained by subdividing.
- A dyadic interval is a subintervals of $[0\,,1]$ that has length 2^{-n} for some integer $n \geq 0$.
- A dyadic interval of length $n \geq 1$ can be written as

$$\left[\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 0, 2, \ldots \text{ is even,}$$
$$\left(\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 1, 3, \ldots \text{ is odd}$$

- Let $\mathscr{D}_n$ denote all dyadic intervals of length 2^{-n}.

Dyadic Intervals

- These are the intervals we obtained by subdividing.
- A dyadic interval is a subintervals of $[0\,,1]$ that has length 2^{-n} for some integer $n \geq 0$.
- A dyadic interval of length $n \geq 1$ can be written as

$$\left[\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 0, 2, \ldots \text{ is even,}$$
$$\left(\frac{j}{2^n}\,,\frac{j+1}{2^n}\right] \quad \text{if } j = 1, 3, \ldots \text{ is odd}$$

- Let $\mathscr{D}_n$ denote all dyadic intervals of length 2^{-n}.
- $\#\mathscr{D}_n = 2^n$ (check!)

Uniform Sampling

- Let $X_1, X_2, \ldots$ be independent random variables

Uniform Sampling

- Let $X_1, X_2, \ldots$ be independent random variables
- $\Pr\{X_j = 0\} = \Pr\{X_j = 1\} = \frac{1}{2}$ for all $j \geq 1$

Uniform Sampling

- Let $X_1, X_2, \ldots$ be independent random variables
- $\Pr\{X_j = 0\} = \Pr\{X_j = 1\} = \frac{1}{2}$ for all $j \geq 1$
- For all sequences $a_1, \ldots, a_n$ of zeros and ones,

$$\Pr\{X_1 = a_1, \ldots, X_n = a_n\} = \prod_{j=1}^{n} \Pr\{X_j = a_j\} = \frac{1}{2^n}. \tag{1}$$

Uniform Sampling

- Let $X_1, X_2, \ldots$ be independent random variables
- $\Pr\{X_j = 0\} = \Pr\{X_j = 1\} = \frac{1}{2}$ for all $j \geq 1$
- For all sequences $a_1, \ldots, a_n$ of zeros and ones,

$$\Pr\{X_1 = a_1, \ldots, X_n = a_n\} = \prod_{j=1}^{n} \Pr\{X_j = a_j\} = \frac{1}{2^n}. \tag{1}$$

- Let X be a random variable who [random] binary digits are $X_1, X_2, \ldots$. I.e.,

$$X = \sum_{j=1}^{\infty} \frac{X_j}{2^j}.$$

Uniform Sampling

- Let $X_1, X_2, \ldots$ be independent random variables
- $\Pr\{X_j = 0\} = \Pr\{X_j = 1\} = \frac{1}{2}$ for all $j \geq 1$
- For all sequences $a_1, \ldots, a_n$ of zeros and ones,

$$\Pr\{X_1 = a_1, \ldots, X_n = a_n\} = \prod_{j=1}^{n} \Pr\{X_j = a_j\} = \frac{1}{2^n}. \qquad (1)$$

- Let X be a random variable who [random] binary digits are $X_1, X_2, \ldots$. I.e.,

$$X = \sum_{j=1}^{\infty} \frac{X_j}{2^j}.$$

- By (1), $\Pr\{X \in I\} = 2^{-n}$ for all $I \in \mathscr{D}_n$.

Zero-One Construction of Length [Lebesgue Measure]

- We just argued that $\Pr\{X \in I\} = \text{length}(I)$ for all dyadic intervals I.

Zero-One Construction of Length [Lebesgue Measure]

- We just argued that $\Pr\{X \in I\} = \text{length}(I)$ for all dyadic intervals I.
- General measure theory tells us that for all sets $I \subseteq [0\,,1]$,

$$\Pr\{X \in I\} = \text{length}(I),$$

provided that we can attribute "length" to I.

Zero-One Construction of Length [Lebesgue Measure]

- We just argued that $\Pr\{X \in I\} = \text{length}(I)$ for all dyadic intervals I.
- General measure theory tells us that for all sets $I \subseteq [0\,,1]$,

$$\Pr\{X \in I\} = \text{length}(I),$$

  provided that we can attribute "length" to I.
- X is "distributed uniformly on $[0\,,1]$"

Borel's Strong Law of Large Numbers

- Recall $X_1, X_2, \ldots$ are independent, and

$$X_j = \begin{cases} 1, & \text{with probab. } \frac{1}{2} \\ 0, & \text{with probab. } \frac{1}{2}. \end{cases}$$

Borel's Strong Law of Large Numbers

- Recall $X_1, X_2, \ldots$ are independent, and

$$X_j = \begin{cases} 1, & \text{with probab. } \frac{1}{2} \\ 0, & \text{with probab. } \frac{1}{2}. \end{cases}$$

- (Expectations)

$$EX_j = \left(1 \times \frac{1}{2}\right) + \left(0 \times \frac{1}{2}\right) = \frac{1}{2} \qquad \text{for all } j$$

Borel's Strong Law of Large Numbers

- Recall $X_1, X_2, \ldots$ are independent, and

$$X_j = \begin{cases} 1, & \text{with probab. } \frac{1}{2} \\ 0, & \text{with probab. } \frac{1}{2}. \end{cases}$$

- (Expectations)

$$EX_j = \left(1 \times \frac{1}{2}\right) + \left(0 \times \frac{1}{2}\right) = \frac{1}{2} \qquad \text{for all } j$$

- (Borel's Theorem, 1909) With probability one:

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \lim_{n\to\infty} \frac{EX_1 + \cdots + EX_n}{n} = \frac{1}{2}.$$

Normal Numbers

- Borel's theorem: With probab. one, $\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \frac{1}{2}$.

Normal Numbers

- Borel's theorem: With probab. one, $\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \frac{1}{2}$.
- $\frac{X_1+\cdots+X_n}{n}$ is also the fraction of 1's in the first n digits of X

Normal Numbers

- Borel's theorem: With probab. one, $\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \frac{1}{2}$.
- $\frac{X_1+\cdots+X_n}{n}$ is also the fraction of 1's in the first n digits of X
- Since $\Pr\{X \in I\} = \text{length}(I)$,

$$\text{Length}\left\{x : \text{ asymp. fraction of ones} = \frac{1}{2}\right\} = 1.$$

Normal Numbers

- Borel's theorem: With probab. one, $\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \frac{1}{2}$.
- $\frac{X_1+\cdots+X_n}{n}$ is also the fraction of 1's in the first n digits of X
- Since $\Pr\{X \in I\} = \text{length}(I)$,

$$\text{Length}\left\{x : \text{ asymp. fraction of ones} = \frac{1}{2}\right\} = 1.$$

- A number $x \in [0\,,1]$ is *normal* if $\lim_{n\to\infty} \frac{x_1+\cdots+x_n}{n} = \frac{1}{2}$.

Normal Numbers

- Borel's theorem: With probab. one, $\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \frac{1}{2}$.
- $\frac{X_1+\cdots+X_n}{n}$ is also the fraction of 1's in the first n digits of X
- Since $\Pr\{X \in I\} = \text{length}(I)$,

$$\text{Length}\left\{x : \text{ asymp. fraction of ones} = \frac{1}{2}\right\} = 1.$$

- A number $x \in [0\,,1]$ is *normal* if $\lim_{n\to\infty} \frac{x_1+\cdots+x_n}{n} = \frac{1}{2}$.
- Borel's theorem: Nonnormal numbers are of length zero.

Normal Numbers

Normal numbers make sense also in base-ten arith. (or any other base ≥ 2 for that matter):

- $x = \sum_{j=1}^{\infty} 10^{-j} x_j$, where $x_j \in \{0, \ldots, 9\}$.

Normal Numbers

Normal numbers make sense also in base-ten arith. (or any other base ≥ 2 for that matter):

- $x = \sum_{j=1}^{\infty} 10^{-j} x_j$, where $x_j \in \{0, \ldots, 9\}$.
- x is normal in base ten if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} I\{x_j = 0\} = \cdots = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} I\{x_j = 9\} = \frac{1}{10}.$$

Normal Numbers

Normal numbers make sense also in base-ten arith. (or any other base ≥ 2 for that matter):

- $x = \sum_{j=1}^{\infty} 10^{-j} x_j$, where $x_j \in \{0, \ldots, 9\}$.
- x is normal in base ten if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} I\{x_j = 0\} = \cdots = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} I\{x_j = 9\} = \frac{1}{10}.$$

- Borel's theorem: Almost every number is normal in base ten. In fact, almost every number is normal in all bases!

Amusing Facts

- There are no numbers that are known to be normal in all bases.

Amusing Facts

- There are no numbers that are known to be normal in all bases.
- (Champernowne, 1933) $0.1234567891011121314\ldots$ is normal in base ten

Amusing Facts

- There are no numbers that are known to be normal in all bases.
- (Champernowne, 1933) 0.1234567891011121314... is normal in base ten
- Champernnown's number is also transcendental (Mahler)

Amusing Facts

- There are no numbers that are known to be normal in all bases.
- (Champernowne, 1933) $0.1234567891011121314\ldots$ is normal in base ten
- Champernnown's number is also transcendental (Mahler)
- (Copeland and Erdős, 1946) $0.23571113\ldots$ is normal in base ten [conjectured by Champernowne, 1933]

Amusing Facts

- There are no numbers that are known to be normal in all bases.
- (Champernowne, 1933) $0.1234567891011121314\ldots$ is normal in base ten
- Champernnown's number is also transcendental (Mahler)
- (Copeland and Erdős, 1946) $0.23571113\ldots$ is normal in base ten [conjectured by Champernowne, 1933]
- and a few others

Amusing Facts

- There are no numbers that are known to be normal in all bases.
- (Champernowne, 1933) $0.1234567891011121314\ldots$ is normal in base ten
- Champernnown's number is also transcendental (Mahler)
- (Copeland and Erdős, 1946) $0.23571113\ldots$ is normal in base ten [conjectured by Champernowne, 1933]
- and a few others
- Is $\pi/10$ normal? How about $\sqrt{2}/10$?

Random-Number Generators

- Your computer generates X uniformly between 0 and 1.

Random-Number Generators

- Your computer generates X uniformly between 0 and 1.
- Is it the case that X has the correct distribution?

Random-Number Generators

- Your computer generates X uniformly between 0 and 1.
- Is it the case that X has the correct distribution?
- The binary digits $X_1, X_2, \ldots$ have lots of structure; so they need to pass various statistical tests (lots known)

Random-Number Generators

- Your computer generates X uniformly between 0 and 1.
- Is it the case that X has the correct distribution?
- The binary digits $X_1, X_2, \ldots$ have lots of structure; so they need to pass various statistical tests (lots known)
- All RNG's will fail the true test of randomness: X_j's have to be normal in all bases.

Ternary Expansions

- Let $x = [0\,,1]$, and write uniquely,

$$x = \sum_{j=1}^{\infty} \frac{x_j}{3^j}, \qquad \text{where } x_j \in \{0\,,1\,,2\}.$$

Ternary Expansions

- Let $x = [0\,,1]$, and write uniquely,

$$x = \sum_{j=1}^{\infty} \frac{x_j}{3^j}, \qquad \text{where } x_j \in \{0\,,1\,,2\}.$$

- The ternary Cantor set $\mathscr{C}$:

$$\mathscr{C} = \text{closure of } \left\{x \in [0\,,1] : \; x_j \in \{0\,,2\}\right\}$$

Ternary Expansions

- Let $x = [0\,,1]$, and write uniquely,

$$x = \sum_{j=1}^{\infty} \frac{x_j}{3^j}, \qquad \text{where } x_j \in \{0\,,1\,,2\}.$$

- The ternary Cantor set $\mathscr{C}$:

$$\mathscr{C} = \text{closure of } \{x \in [0\,,1] : \ x_j \in \{0\,,2\}\}$$

- $x = 1/3$ is in the Cantor set; in fact, $x = [0\,,2\,,2\,,\ldots]$

Ternary Expansions

- Let $x = [0\,,1]$, and write uniquely,

$$x = \sum_{j=1}^{\infty} \frac{x_j}{3^j}, \qquad \text{where } x_j \in \{0\,,1\,,2\}.$$

- The ternary Cantor set $\mathscr{C}$:

$$\mathscr{C} = \text{closure of } \{x \in [0\,,1] : \; x_j \in \{0\,,2\}\}$$

- $x = 1/3$ is in the Cantor set; in fact, $x = [0\,,2\,,2\,,\ldots]$
- If $\frac{1}{3} < x < \frac{2}{3}$ then $x \notin \mathscr{C}$ etc.

Ternary Expansions

- Let $x = [0\,,1]$, and write uniquely,

$$x = \sum_{j=1}^{\infty} \frac{x_j}{3^j}, \qquad \text{where } x_j \in \{0\,,1\,,2\}.$$

- The ternary Cantor set $\mathscr{C}$:

$$\mathscr{C} = \text{closure of } \{x \in [0\,,1] : \ x_j \in \{0\,,2\}\}$$

- $x = 1/3$ is in the Cantor set; in fact, $x = [0\,,2\,,2\,,\ldots]$
- If $\frac{1}{3} < x < \frac{2}{3}$ then $x \notin \mathscr{C}$ etc.
- $\mathscr{C}$ = The middle-thirds Cantor set

Devil's Staircase

- Let $X_1, X_2, \ldots$ be independent,

$$\Pr\{X_1 = 0\} = \Pr\{X_1 = 2\} = \frac{1}{2}.$$

Devil's Staircase

- Let $X_1, X_2, \ldots$ be independent,

$$\Pr\{X_1 = 0\} = \Pr\{X_1 = 2\} = \frac{1}{2}.$$

- Let X be "uniformly distributed" on $\mathscr{C}$; i.e.,

$$X = \sum_{j=1}^{\infty} \frac{X_j}{3^j} \qquad \Longrightarrow \qquad \Pr\{X \in \mathscr{C}\} = 1.$$

Devil's Staircase

- Let $X_1, X_2, \ldots$ be independent,

$$\Pr\{X_1 = 0\} = \Pr\{X_1 = 2\} = \frac{1}{2}.$$

- Let X be "uniformly distributed" on $\mathscr{C}$; i.e.,

$$X = \sum_{j=1}^{\infty} \frac{X_j}{3^j} \qquad \Longrightarrow \qquad \Pr\{X \in \mathscr{C}\} = 1.$$

- Distribution function of X,

$$F(x) := \Pr\{X \leq x\} \qquad \text{"devil's staircase"}$$

Devil's Staircase

- Let $X_1, X_2, \ldots$ be independent,

$$\Pr\{X_1 = 0\} = \Pr\{X_1 = 2\} = \frac{1}{2}.$$

- Let X be "uniformly distributed" on $\mathscr{C}$; i.e.,

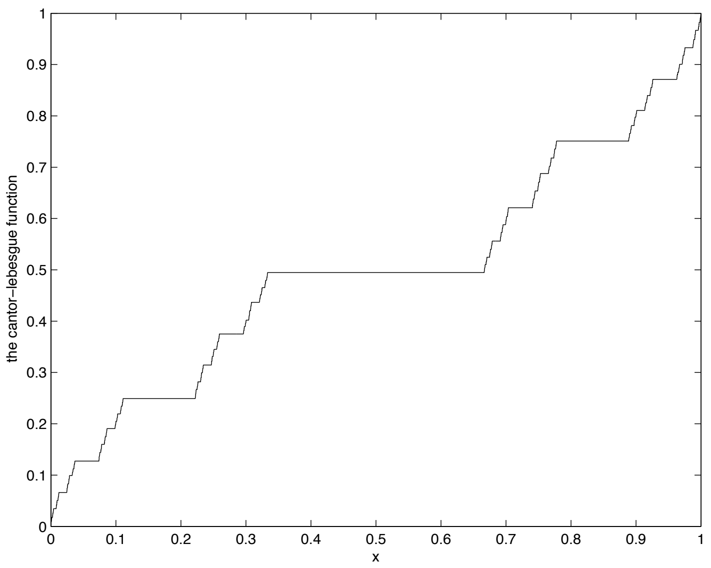
$$X = \sum_{j=1}^{\infty} \frac{X_j}{3^j} \qquad \Longrightarrow \qquad \Pr\{X \in \mathscr{C}\} = 1.$$

- Distribution function of X,

$$F(x) := \Pr\{X \leq x\} \qquad \text{"devil's staircase"}$$

- Aka Cantor–Lebesgue function

The Cantor–Lebesgue Function



The Cantor–Lebesgue Function

Theorem (Cantor)

The Cantor–Lebesgue Function

Theorem (Cantor)

- $C := \{x : F'(x)$ *exists and is* $= 0\}$ *has length one.*

The Cantor–Lebesgue Function

Theorem (Cantor)

- $C := \{x : F'(x)$ *exists and is* $= 0\}$ *has length one.*
- F *is nondecreasing and continuous*

The Cantor–Lebesgue Function

Theorem (Cantor)

- $C := \{x : F'(x)$ *exists and is* $= 0\}$ *has length one.*
- F *is nondecreasing and continuous*
- $F(0) = 0$

The Cantor–Lebesgue Function

Theorem (Cantor)

- $C := \{x : F'(x)$ *exists and is* $= 0\}$ *has length one.*
- F *is nondecreasing and continuous*
- $F(0) = 0$
- $F(1) = 1$

The Cantor–Lebesgue Function

Theorem (Cantor)

- $C := \{x : F'(x) \text{ exists and is} = 0\}$ *has length one.*
- F *is nondecreasing and continuous*
- $F(0) = 0$
- $F(1) = 1$

The Cantor–Lebesgue Function

Theorem (Cantor)

- *$C := \{x : F'(x)$ exists and is $= 0\}$ has length one.*
- *F is nondecreasing and continuous*
- $F(0) = 0$
- $F(1) = 1$

Fundamental theorem of calculus(?):

$$1 = F(1) - F(0) = \int_0^1 F'(x)\,dx = 0$$

The Cantor–Lebesgue Function

Theorem (Cantor)

- *$C := \{x : F'(x)$ exists and is $= 0\}$ has length one.*
- *F is nondecreasing and continuous*
- $F(0) = 0$
- $F(1) = 1$

Fundamental theorem of calculus(?):

$$1 = F(1) - F(0) = \int_0^1 F'(x)\,dx = 0$$

Mind those technical conditions of theorems!

Hausdorff Dimension

Let S be a set in $\mathbf{R}^n$. Roughly speaking, its *Hausdorff dimension*

$$\dim_H S := \max\left\{ s \in [0\,,n] :\ E\left(\frac{1}{\|X-Y\|^s}\right) < \infty \right\},$$

where X and Y are:

- independent;

Hausdorff Dimension

Let S be a set in $\mathbf{R}^n$. Roughly speaking, its *Hausdorff dimension*

$$\dim_{_H} S := \max\left\{ s \in [0\,,n] : \; E\left(\frac{1}{\|X-Y\|^s}\right) < \infty \right\},$$

where X and Y are:

- independent;
- both distributed "uniformly" on S

Hausdorff Dimension

Let S be a set in $\mathbf{R}^n$. Roughly speaking, its *Hausdorff dimension*

$$\dim_H S := \max\left\{ s \in [0\,, n] : \; E\left(\frac{1}{\|X - Y\|^s} \right) < \infty \right\},$$

where X and Y are:

- independent;
- both distributed "uniformly" on S

Hausdorff Dimension

Let S be a set in $\mathbf{R}^n$. Roughly speaking, its *Hausdorff dimension*

$$\dim_H S := \max\left\{ s \in [0\,, n] : \ E\left(\frac{1}{\|X - Y\|^s} \right) < \infty \right\},$$

where X and Y are:

- independent;
- both distributed "uniformly" on S

(Frostman, 1935)

Dimension of the Cantor Set

Theorem (Hausdorff, 1919)

$\dim_H \mathscr{C} = \log_3(2) = \ln 2 / \ln 3 \simeq 0.7615$

Dimension of the Cantor Set

Theorem (Hausdorff, 1919)

$\dim_H \mathscr{C} = \log_3(2) = \ln 2/\ln 3 \simeq 0.7615$

Dimension of the Cantor Set

Theorem (Hausdorff, 1919)

$\dim_H \mathscr{C} = \log_3(2) = \ln 2 / \ln 3 \simeq 0.7615$

Strategy: Let X and Y be uniformly distributed on $\mathscr{C}$, both independent. Then we wish to demonstrate that:

- if $s > \log_3(2)$ then $E(|X - Y|^{-s}) = \infty$;

Dimension of the Cantor Set

Theorem (Hausdorff, 1919)

$\dim_H \mathscr{C} = \log_3(2) = \ln 2 / \ln 3 \simeq 0.7615$

Strategy: Let X and Y be uniformly distributed on $\mathscr{C}$, both independent. Then we wish to demonstrate that:

- if $s > \log_3(2)$ then $E(|X - Y|^{-s}) = \infty$;
- if $s < \log_3(2)$ then $E(|X - Y|^{-s}) < \infty$.

Finally, a Proof

- Let us prove that if $s < \log_3(2)$ then $E(|X - Y|^{-s}) < \infty$. This proves that $\dim_H \mathscr{C} \geq \log_3(2)$, and is in fact the harder bound.

Finally, a Proof

- Let us prove that if $s < \log_3(2)$ then $E(|X - Y|^{-s}) < \infty$. This proves that $\dim_H \mathscr{C} \geq \log_3(2)$, and is in fact the harder bound.

Finally, a Proof

- Let us prove that if $s < \log_3(2)$ then $E(|X - Y|^{-s}) < \infty$. This proves that $\dim_{_H} \mathscr{C} \geq \log_3(2)$, and is in fact the harder bound.
- Let $N := \min\{j \geq 1 : X_j \neq Y_j\}$; then $\Pr\{N > k\} = 2^{-k}$ for all $k \geq 0$.

Finally, a Proof

- Let us prove that if $s < \log_3(2)$ then $E(|X - Y|^{-s}) < \infty$. This proves that $\dim_H \mathscr{C} \geq \log_3(2)$, and is in fact the harder bound.
- Let $N := \min\{j \geq 1 : X_j \neq Y_j\}$; then $\Pr\{N > k\} = 2^{-k}$ for all $k \geq 0$.
- Therefore, $\Pr\{N = k\} = \Pr\{N > k - 1\} - \Pr\{N > k\} = 2^{-k}$.

Finally, a Proof

- Let us prove that if $s < \log_3(2)$ then $E(|X - Y|^{-s}) < \infty$. This proves that $\dim_H \mathscr{C} \geq \log_3(2)$, and is in fact the harder bound.
- Let $N := \min\{j \geq 1 : X_j \neq Y_j\}$; then $\Pr\{N > k\} = 2^{-k}$ for all $k \geq 0$.
- Therefore, $\Pr\{N = k\} = \Pr\{N > k - 1\} - \Pr\{N > k\} = 2^{-k}$.
- We have

$$\frac{1}{|X - Y|^s} \leq \frac{1}{3^{Ns}}$$

Finally, a Proof

- Let us prove that if $s < \log_3(2)$ then $E(|X - Y|^{-s}) < \infty$. This proves that $\dim_H \mathscr{C} \geq \log_3(2)$, and is in fact the harder bound.
- Let $N := \min\{j \geq 1 : X_j \neq Y_j\}$; then $\Pr\{N > k\} = 2^{-k}$ for all $k \geq 0$.
- Therefore, $\Pr\{N = k\} = \Pr\{N > k - 1\} - \Pr\{N > k\} = 2^{-k}$.
- We have

$$\frac{1}{|X - Y|^s} \leq \frac{1}{3^{Ns}}$$

- If $s < \log_3(2)$ then

$$E\left(\frac{1}{|X - Y|^s}\right) \leq E\left(\frac{1}{3^{Ns}}\right) = \sum_{k=1}^{\infty} \frac{1}{3^{ks}} \times 2^{-k} < \infty.$$