An Introduction to Ergodic Theory

Normal Numbers: We Can't See Them, But They're Everywhere!

Joseph Horan

Department of Mathematics and Statistics
University of Victoria
Victoria, BC

December 5, 2013

An introduction

We need four things:

An introduction

We need four things:

- A set X. Here, $X = [0, 1]$.

An introduction

We need four things:

- A set X. Here, $X = [0, 1]$.
- A measure on X, ie. a set function on $\mathcal{P}(X)$ that is a notion of size. Here, we use the Lebesgue measure λ, which is length on the real line: $\lambda(a, b) = b - a$.

An introduction

We need four things:

- A set X. Here, $X = [0, 1]$.
- A measure on X, ie. a set function on $\mathcal{P}(X)$ that is a notion of size. Here, we use the Lebesgue measure λ, which is length on the real line: $\lambda(a, b) = b - a$.
- Sets inside of X which one can measure, ie. to which one can apply λ. A σ-algebra $\mathfrak{B}$, to be technical.

An introduction

We need four things:

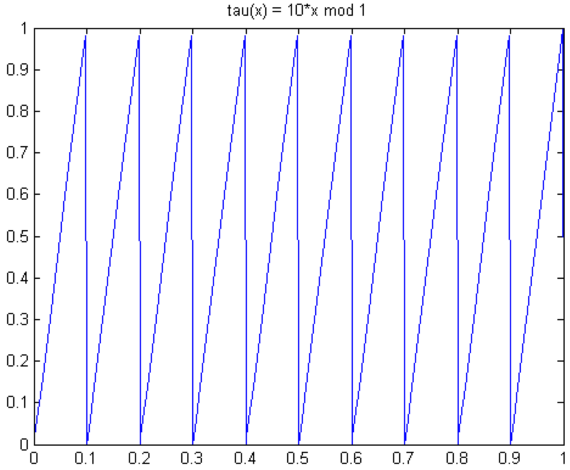
- A set X. Here, $X = [0, 1]$.
- A measure on X, ie. a set function on $\mathcal{P}(X)$ that is a notion of size. Here, we use the Lebesgue measure λ, which is length on the real line: $\lambda(a, b) = b - a$.
- Sets inside of X which one can measure, ie. to which one can apply λ. A σ-algebra $\mathfrak{B}$, to be technical.
- A map $\tau : X \to X$, which "preserves" the measure of a set under pullbacks: $\lambda(\tau^{-1}(A)) = \lambda(A)$. Here, we use $\tau(x) = 10x \mod 1$.

An introduction

We need four things:

- A set X. Here, $X = [0,1]$.
- A measure on X, ie. a set function on $\mathcal{P}(X)$ that is a notion of size. Here, we use the Lebesgue measure λ, which is length on the real line: $\lambda(a,b) = b - a$.
- Sets inside of X which one can measure, ie. to which one can apply λ. A σ-algebra $\mathfrak{B}$, to be technical.
- A map $\tau : X \to X$, which "preserves" the measure of a set under pullbacks: $\lambda(\tau^{-1}(A)) = \lambda(A)$. Here, we use $\tau(x) = 10x \mod 1$.

Together, this is called a *dynamical system*: $(X, \mathfrak{B}, \lambda, \tau)$. One can think of it like a state space, which evolves over time by way of iterating τ.


tau(x) = 10*x mod 1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

More definitions

We make three more definitions:

More definitions

We make three more definitions:

- *Almost everywhere* means *everywhere except a set of measure zero*.

More definitions

We make three more definitions:

- *Almost everywhere* means *everywhere except a set of measure zero*.
- The *orbit* of x under τ is $\{\tau^n(x)\}_{n=0}^{\infty}$. This is the 'future' of x, under τ, in the state space interpretation.

More definitions

We make three more definitions:

- *Almost everywhere* means *everywhere except a set of measure zero*.
- The *orbit* of x under τ is $\{\tau^n(x)\}_{n=0}^{\infty}$. This is the 'future' of x, under τ, in the state space interpretation.
- (λ, τ) is called *ergodic* if when $\tau^{-1}(A) = A$, then one of $\lambda(A), \lambda(X \setminus A)$ is zero.

More definitions

We make three more definitions:

- *Almost everywhere* means *everywhere except a set of measure zero*.
- The *orbit* of x under τ is $\{\tau^n(x)\}_{n=0}^{\infty}$. This is the 'future' of x, under τ, in the state space interpretation.
- (λ, τ) is called *ergodic* if when $\tau^{-1}(A) = A$, then one of $\lambda(A), \lambda(X \setminus A)$ is zero.

If we don't have an ergodic pair (λ, τ), then if $\tau^{-1}(A) = A$ with non-zero measure, we could study τ just on A instead, and so decompose our space. Here, our (λ, τ) are ergodic.

A Specific Case of Birkhoff's Ergodic Theorem

Theorem (Birkhoff)

Let everything be as above, and $f \in L^1([0,1])$. Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\tau^i(x)) = \int_{[0,1]} f \; d\lambda,$$

where the left-hand side converges almost everywhere with respect to λ.

Briefly, *time average* = *space average*.

The set of $x \in [0,1]$ for which this is true depends on τ and on f.

f is an observable on the state space, so it samples points.

Normal Numbers

What do we mean by "normal"? Essentially, a real number x is normal if for any base b, the frequency of finite strings of a fixed length in the base b representation of x is uniform, ie. whenever a word w has length n, the frequency with which it shows up is $\frac{1}{b^n}$, independent of which word it is.

Normal Numbers

What do we mean by "normal"? Essentially, a real number x is normal if for any base b, the frequency of finite strings of a fixed length in the base b representation of x is uniform, ie. whenever a word w has length n, the frequency with which it shows up is $\frac{1}{b^n}$, independent of which word it is.

For example, if $b = 2$, $x = 0.10101010\ldots$ is not normal, because 11 doesn't occur, when it should occur with frequency $\frac{1}{2^2} = \frac{1}{4}$.

Normal Numbers

What do we mean by "normal"? Essentially, a real number x is normal if for any base b, the frequency of finite strings of a fixed length in the base b representation of x is uniform, ie. whenever a word w has length n, the frequency with which it shows up is $\frac{1}{b^n}$, independent of which word it is.

For example, if $b = 2$, $x = 0.10101010\ldots$ is not normal, because 11 doesn't occur, when it should occur with frequency $\frac{1}{2^2} = \frac{1}{4}$.

As well, if x is rational, its base b representation will always be either terminating or repeating, so it cannot be normal.

Normal Numbers

What do we mean by "normal"? Essentially, a real number x is normal if for any base b, the frequency of finite strings of a fixed length in the base b representation of x is uniform, ie. whenever a word w has length n, the frequency with which it shows up is $\frac{1}{b^n}$, independent of which word it is.

For example, if $b = 2$, $x = 0.10101010\ldots$ is not normal, because 11 doesn't occur, when it should occur with frequency $\frac{1}{2^2} = \frac{1}{4}$.

As well, if x is rational, its base b representation will always be either terminating or repeating, so it cannot be normal.

Conjecture

There exists at least one normal number.

Borel Normal Number Theorem

Theorem (Borel, 1909)

Almost every real number, with respect to the Lebesgue measure, is normal.

Proof of the BNNT

We illustrate the general argument by looking at the specific case of $b = 10$, and looking only at the density of single digits.

Proof of the BNNT

We illustrate the general argument by looking at the specific case of $b = 10$, and looking only at the density of single digits.

The number of times that k appears in the first n digits of the expansion is

$$\sum_{i=0}^{n-1} \chi_{I_k}(\tau^i(x)),$$

where we have:

Proof of the BNNT

We illustrate the general argument by looking at the specific case of $b = 10$, and looking only at the density of single digits.

The number of times that k appears in the first n digits of the expansion is

$$\sum_{i=0}^{n-1} \chi_{I_k}(\tau^i(x)),$$

where we have:

- $I_k = [\frac{k}{10}, \frac{k+1}{10}]$, for integers $0 \leq k \leq 9$.

Proof of the BNNT

We illustrate the general argument by looking at the specific case of $b = 10$, and looking only at the density of single digits.

The number of times that k appears in the first n digits of the expansion is

$$\sum_{i=0}^{n-1} \chi_{I_k}(\tau^i(x)),$$

where we have:

- $I_k = [\frac{k}{10}, \frac{k+1}{10}]$, for integers $0 \leq k \leq 9$.
- $\tau(0.x_1x_2\ldots) = x_1.x_2x_3\ldots \mod 1 = 0.x_2x_3\ldots$

Proof of the BNNT

We illustrate the general argument by looking at the specific case of $b = 10$, and looking only at the density of single digits.

The number of times that k appears in the first n digits of the expansion is

$$\sum_{i=0}^{n-1} \chi_{I_k}(\tau^i(x)),$$

where we have:

- $I_k = [\frac{k}{10}, \frac{k+1}{10}]$, for integers $0 \leq k \leq 9$.
- $\tau(0.x_1x_2\ldots) = x_1.x_2x_3\ldots \mod 1 = 0.x_2x_3\ldots$
- $\chi_{I_k}(x)$ checks if the first digit of x is k. That is:

$$\chi_{I_k}(x) = \begin{cases} 1, & x = 0.kx_2x_3\ldots \\ 0, & \text{otherwise} \end{cases}$$

Proof continued

$\chi_{I_k} \in L^1([0,1])$, so we can apply the Birkhoff Ergodic Theorem:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_{I_k}(\tau^i(x)) = \int_{[0,1]} \chi_{I_k} \, d\lambda$$
$$= \lambda(I_k) = \frac{1}{10}, \; \lambda - a.e.$$

Proof continued

$\chi_{I_k} \in L^1([0,1])$, so we can apply the Birkhoff Ergodic Theorem:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_{I_k}(\tau^i(x)) = \int_{[0,1]} \chi_{I_k} \, d\lambda$$
$$= \lambda(I_k) = \frac{1}{10}, \; \lambda - a.e.$$

This means that the density of the digit k in the decimal expansion of x is $\frac{1}{10}$ for all $0 \leq k \leq 9$, and so the distribution of the digits is uniform, which is what we wanted to show.

Proof continued

$\chi_{I_k} \in L^1([0,1])$, so we can apply the Birkhoff Ergodic Theorem:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \chi_{I_k}(\tau^i(x)) = \int_{[0,1]} \chi_{I_k}\ d\lambda$$
$$= \lambda(I_k) = \frac{1}{10},\ \ \lambda - a.e.$$

This means that the density of the digit k in the decimal expansion of x is $\frac{1}{10}$ for all $0 \leq k \leq 9$, and so the distribution of the digits is uniform, which is what we wanted to show.

In general, we do this for a more general word instead of a single digit, and then we use countable sub-additivity of the measure to conclude the proof. □

Conclusion

We introduced ergodic theory, and applied it to a neat problem that seemed far removed from the abstract theory. Turns out that ergodic theory has other such surprising applications!

Ask or see the extended abstract for references.