Ergodic Theory and Applications

Qijun Lin, Michael Brady, Dylan Wallace

Wednesday, December 5th 2024

Binary representation of various numbers in [0,1]

$0 = 0.00000000000\cdots$
$1/3 = 0.101010101010\cdots$
$0.5 = 0.100000000\cdots = 0.011111111111\cdots$
$1 = 1.0000000000000\cdots = 0.11111111111\cdots$
What can we say about the frequency of certain types of binary numbers?

Borel's Theorem

A randomly chosen number in [0, 1] has the same frequency of 0's and 1's in its binary representation with probability 1.

Itinerary

1. Introduction to Measure
2. Measure Preserving Transformations
3. Ergodicity
4. Ergodic Theorem
5. Borel's Theorem

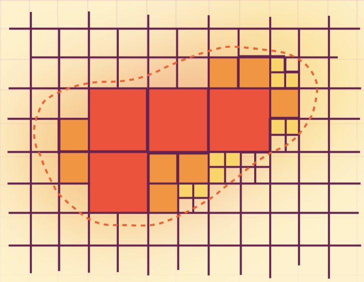
Measure = "Size" of a given set

Axioms: Let (X, β, m) be a measure space. Then:

1. $m(E) \geq 0, \forall E \in \beta$
2. $m(\varnothing) = 0$
3. If a collection of $\{E_i\}$ are disjoint then $m(\cup_i E_i) = \sum_i m(E_i)$

Measure in this context

- There are many different "measures" on a set, but in this context we care about the Lebesgue measure
- On an Euclidean space, we can approximate each subset by a union of "interval" sets (whose measure is their Euclidean size) that either contain or are contained by the subset
- When the measures of the two approximating sets converge, then we can let the measure of our subset be the same as the measure of the limit. This is the Lebesgue measure

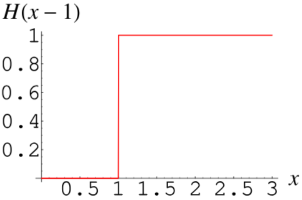


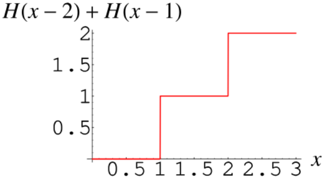
Measurable Functions

We often care about functions $f : X \to \mathbb{R}$ for measure spaces X. We give $\mathbb{R}$ the Lebesgue measure.

We have $f : X \to \mathbb{R}$ is measurable if for every measurable $D \subseteq \mathbb{R}$, we have $f^{-1}(D)$ measurable also. We say $f, g : X \to \mathbb{R}$ are equal *almost everywhere* if $m(\{x : f(x) \neq g(x)\}) = 0$.

For a "simple" function (i.e. a piecewise constant function), we can define $\int f dm = \sum_{i=1}^{n} a_i m(A_i)$ (where A_i are the sets on which f is piecewise).





For non-simple functions, we can construct a sequence of simple functions that converge to f, and we can define $\int f dm$ as the limit of the integrals of the simple functions.

Measure Preserving Transformations

Definition: Measure Preserving

A transformation T between measure spaces (X_1, β_1, m_1) and (X_2, β_2, m_2) is *measure preserving* if T is measurable and $m_1(T^{-1}(B_2)) = m_2(B_2)$ $\forall B_2 \in \beta_2$

- We often care about measure-preserving transformations from a space X to itself, and care about what transformations $T^n = T \circ \cdots_{\text{n times}} \cdots \circ T$ look like

Poincaré Recurrence Theorem

Theorem (Poincaré)

Suppose $T : X \to X$ is measure-preserving with $m(X) = 1$, and let $A \subset X$ be measurable with $m(A) > 0$. Then for *almost every* point $x \in A$, there exists $n \in \mathbb{N}$ for which $T^n(x) \in A$ also.

- Note that the $x \in A$ don't all have to return to A at the same time. Thus the Poincaré Recurrence Theorem is a sort of weak recurrence. Contrast this to when T is Ergodic

Ergodic Transformation

Definition: Ergodic

A measure-preserving map $T : (X, \beta, m) \to (X, \beta, m)$ is *ergodic* if for every $A \in \beta$ s.t. $T^{-1}(A) = A$ we have $m(A) = 0$ or $m(A) = 1$.

- If $m(A) = 0$, then A is "too small to matter". If $m(A) = 1$, then $X \setminus A$ is "too small to matter".
- So if T is ergodic and $T^{-1}(A) = A$, then A is "close to a trivial set" in a measure-theoretic sense

An Equivalent Definition

Definition: Ergodic (Alternative)

A map $T : X \to X$ is ergodic if

$$\forall A \subseteq X,\ m(A) > 0 \implies m\left(\cup_{i=1}^{\infty} T^{-i}(A)\right) = 1$$

- We can think of the set $\cup_{i=1}^{\infty} T^{-i}(A)$ as the set of points in X that eventually get mapped to A
- So T is ergodic if for *any* set $A \subseteq X$ with "size", almost every point in X eventually ends up in A. Contrast with Poincaré Recurrence Theorem

Proof of Equivalency

$\Longrightarrow$: Suppose T is ergodic, $m(A) > 0$, and let $C = \cup_{i=1}^{\infty} T^{-i}(A)$. Then $T^{-1}(C) \subset C$, and T is measure-preserving so $m(T^{-1}(C)) = m(C)$. Therefore $m(C) = 0$ or $m(C) = 1$; but $A \subseteq C$ and $m(A) > 0$ so $m(C) = 1$. ☺

$\Longleftarrow$: For $0 < m(A) < 1$, let $C = \cup_{i=1}^{\infty} T^{-i}(A)$; then $m(C) = 1$. Suppose $T^{-1}(A) = A$. Then by induction $T^{-n}(A) = A$ for any $n \in \mathbb{N}$, so $C = \cup_{i=1}^{\infty} A = A$. Thus $m(C) = m(A) \in (0, 1)$. By contradiction, we must have $m(A) = 0$ or $m(A) = 1$. ☺

Birkhoff Ergodic Theorem

Theorem (Birkhoff)

Let $T : X \to X$ be measure-preserving, and $f : X \to \mathbb{R}$ a measurable function. Then the functions $f_n(x) = \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$ converge to a measurable function f^*, with $f^*(T(x)) = f^*(x)$ almost everywhere. If $m(X) < \infty$, then we also have $\int f^* \, dm = \int f \, dm$.

Intuition of Ergodic Theorem

- We can think of $f_n(x) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ as averaging the value of f on points $\{x, T(x), T^2(x), \ldots, T^{n-1}(x)\}$
- T sufficiently spreads out points in $X \implies \{x, T(x), T^2(x), \ldots\}$ is a representative sample of points in $X \implies f^*(x)$ is average of $f(x)$, $x \in X$.

In fact, this "sufficiently spreading out" property that we want from T is satisfied when T is ergodic!

Intuition of Ergodic Theorem

- We can think of $f_n(x) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ as averaging the value of f on points $\{x, T(x), T^2(x), \ldots, T^{n-1}(x)\}$
- T sufficiently spreads out points in $X \implies \{x, T(x), T^2(x), \ldots\}$ is a representative sample of points in $X \implies f^*(x)$ is average of $f(x)$, $x \in X$.

In fact, this "sufficiently spreading out" property that we want from T is satisfied when T is ergodic!

$\{x, T(x), T^2(x), \ldots\}$ is a time process; $f_n(x)$ is summing across time. $\int f\, dm$ is a sum on X, so it is a sum across space. $\implies$ Birkhoff's Ergodic Theorem demonstrates a correspondence between time and space for ergodic maps

Borel's Theorem on Normal Numbers

Almost all numbers in $[0,1)$ are normal to base 2. In other words, for *almost all* $x \in [0,1)$ the frequency of 1's in the binary expansion of x is $\frac{1}{2}$.

Borel's Theorem on Normal Numbers

Almost all numbers in $[0, 1)$ are normal to base 2. In other words, for *almost all* $x \in [0, 1)$ the frequency of 1's in the binary expansion of x is $\frac{1}{2}$.

Strategy: Birkhoff's theorem lets us write the average of some time processes as an integral.
If we write $x \in [0, 1)$ as $x = 0.a_1 a_2 a_3 \ldots$, then $\{a_1, a_2, a_3, \ldots\}$ is like a time process.
$\implies$ We want some ergodic map $T : [0, 1) \to [0, 1)$ and function $f : [0, 1) \to \mathbb{R}$ such that $f(T^n(x)) = a_{n+1}$!

Proof of Borel's Theorem

Let $T : [0,1) \to [0,1)$ by $T(x) = 2x \mod 1$, and
$f : [0,1) \to \mathbb{R}$ by $f(x) = \begin{cases} 1 & x \in [\frac{1}{2}, 1) \\ 0 & x \in [0, \frac{1}{2}) \end{cases}$.

We will

- prove T is measure-preserving and ergodic;
- show that $f(T^n(x)) = a_{n+1}$;
- use the above two to prove Borel's theorem.

Proof T is Measure Preserving

Proof: T is Measure-Preserving

Let $O = [a, b)$ be any interval. We want $m(T^{-1}(O)) = m(O)$.

We know $T^{-1}(O) = [\frac{a}{2}, \frac{b}{2}) \cup [\frac{a+1}{2}, \frac{b+1}{2})$. Then $m(O) = b - a$ and $m(T^{-1}(O)) = m(\frac{a}{2}, \frac{b}{2}) + m(\frac{a+1}{2}, \frac{b+1}{2}) = \frac{b-a}{2} + \frac{b-a}{2} = b - a$.

Intervals $[a, b)$ generate measurable sets on $[0, 1) \implies T$ is measure-preserving ☺

Proof T is Ergodic

Proof: T is Ergodic

Let $A = [a, b) \subset [0, 1)$ s.t. $T^{-1}(A) = A$.

Then $[\frac{a}{2}, \frac{b}{2}) \cup [\frac{a+1}{2}, \frac{b+1}{2}) = [a, b)$, so we have $\frac{a}{2} = a$, $\frac{b+1}{2} = b$, and $\frac{b}{2} = \frac{a+1}{2}$.

Solving gives us $a = 0$ and $b = 1$. Then $m(A) = 1$, so every interval A with $T^{-1}(A) = A$ has measure 1.

Intervals generate the measurable sets on $[0, 1) \implies T$ is ergodic ☺

Proof Continued

Let $Y \subseteq [0,1)$ be the set of points with a unique binary expansion (e.g. $0.1000000\cdots = 0.01111111\ldots$ is not unique). Then the set of points with non-unique binary expansion are countable and have measure 0, so $m(Y) = 1$.

Proof Continued

Let $Y \subseteq [0,1)$ be the set of points with a unique binary expansion (e.g. $0.1000000 \cdots = 0.01111111\ldots$ is not unique). Then the set of points with non-unique binary expansion are countable and have measure 0, so $m(Y) = 1$.

Suppose $x \in Y$ and $x = 0.a_1a_2a_3\ldots$. Then $x = \frac{a_1}{2} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \ldots$, so $T(x) = (a_1 + \frac{a_2}{2} + \frac{a_3}{2^2} + \frac{a_4}{2^3} + \ldots) \mod 1 = \frac{a_2}{2} + \frac{a_3}{2^2} + \frac{a_4}{2^3} + \ldots$.

$\implies$ By induction, we see that $T^i(x) = \frac{a_{i+1}}{2} + \frac{a_{i+2}}{2^2} + \frac{a_{i+3}}{2^3} + \ldots$

Proof Continued

Let $f = \chi_{[\frac{1}{2},1)} = \begin{cases} 1 & x \in [\frac{1}{2},1) \\ 0 & x \in [0,\frac{1}{2}) \end{cases}$. Then we have:

$$f(T^i(x)) = f\left(\frac{a_{i+1}}{2} + \frac{a_{i+2}}{2^2} + \dots\right) = \begin{cases} 1 \textit{ iff } a_{i+1} = 1 \\ 0 \textit{ iff } a_{i+1} = 0 \end{cases}$$

because $\frac{a_{i+2}}{2^2} + \frac{a_{i+3}}{2^3} + \cdots < \frac{1}{2}$, so $T^i(x) \geq \frac{1}{2}$ only if $\frac{a_{i+1}}{2} = \frac{1}{2}$

Proof Continued

Let $f = \chi_{[\frac{1}{2},1)} = \begin{cases} 1 & x \in [\frac{1}{2}, 1) \\ 0 & x \in [0, \frac{1}{2}) \end{cases}$. Then we have:

$$f(T^i(x)) = f\left(\frac{a_{i+1}}{2} + \frac{a_{i+2}}{2^2} + \dots\right) = \begin{cases} 1 \; iff \; a_{i+1} = 1 \\ 0 \; iff \; a_{i+1} = 0 \end{cases}$$

because $\frac{a_{i+2}}{2^2} + \frac{a_{i+3}}{2^3} + \dots < \frac{1}{2}$, so $T^i(x) \geq \frac{1}{2}$ only if $\frac{a_{i+1}}{2} = \frac{1}{2}$

Thus $f(T^i(x)) = 1$ iff $a_{i+1} = 1$, so $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ is the frequency of 1's in $x = 0.a_1 a_2 \dots a_n$.

$\implies \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) = \int f\, dm$ by Birkhoff.

f is piecewise constant, and we can see that $\int f\, dm = \frac{1}{2}$. Thus for every $x \in Y$, its binary expansion has 1's with frequency $\frac{1}{2}$! ☺