Computability, randomness and the ergodic decomposition

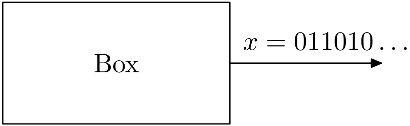
Mathieu Hoyrup
(Матьё Уарю from June 10 to June 17)

INRIA Nancy - France

June 14, 2010

Probabilistic process

We consider a probabilistic process that produces bits. It is fully described by a stationary probability measure P over $\{0,1\}^{\mathbb{N}}$.



Each $w \in \{0,1\}^*$ has a probability $P(w)$ of appearing at time 0.
P is stationary: w appears at time n with the same probability as at time 0, for every n.

Limit frequencies

Theorem (Birkhoff, 1931)

For P-almost every $x \in \{0,1\}^{\mathbb{N}}$, for each $w \in \{0,1\}^*$ the following limit exists:

$$P_x(w) := \lim_{n\to\infty} \frac{\#_{occ}(w, x_0 x_1 \ldots x_{n-1})}{n}.$$

Definition

A sequence x is ***generic*** if $P_x(w)$ exists for every $w \in \{0,1\}^*$.

Property

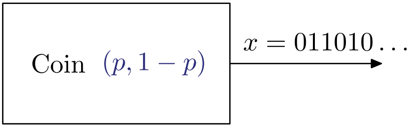
For every generic x, P_x is a stationary probability measure.

Question

Can we say more about P_x?

Example 1

Coin flipping



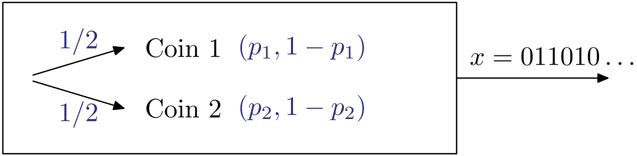
$$B_p(w) = p^{|w|_1}(1-p)^{|w|_0}$$

Strong law of large numbers

B_p-almost surely, the limit frequency $P_x(w)$ of occurrences of w is $B_p(w)$. Hence $P_x = B_p$ for B_p-almost every x.

Example 2

Coins flipping



- First step: choose coin 1 or 2 at random ($(1/2, 1/2)$, say), *once for all*.
- Following steps: flip the chosen coin.

$$P = \frac{1}{2}(B_{p_1} + B_{p_2}).$$

With probability $1/2$, the induced measure will be B_{p_1}. With probability $1/2$, it will be B_{p_2}.

Ergodicity

Definition

A stationary measure P has a ***decomposition*** if $P = \alpha P_1 + (1-\alpha)P_2$ where:

- $0 < \alpha < 1$,
- P_1 and P_2 are stationary,
- $P_1 \neq P_2$.

A stationary measure is ***ergodic*** if it has no decomposition.

The 2 examples

1. The Bernoulli measure B_p is ergodic for every p.
2. Of course, $\frac{1}{2}(B_{p_1} + B_{p_2})$ is not ergodic if $p_1 \neq p_2$.

Ergodic decomposition

Question

Can we say more about P_x?

The ergodic case

Theorem (Birkhoff, 1931)

Let P be an ergodic stationary measure. For P-almost every sequence x, $P_x = P$.

The non-ergodic case

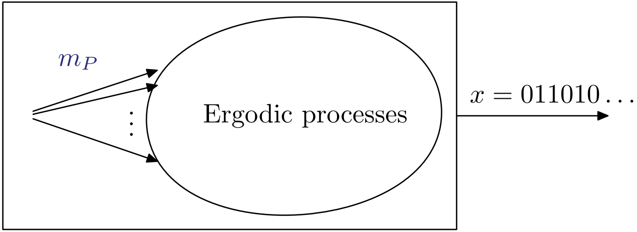
Theorem (Ergodic decomposition)

Let P be a stationary measure. For P-almost every sequence x, P_x is ergodic.

Ergodic decomposition

Every stationary process can be decomposed into:

- First step: pick an ergodic process at random.
- Following steps: run the chosen process.

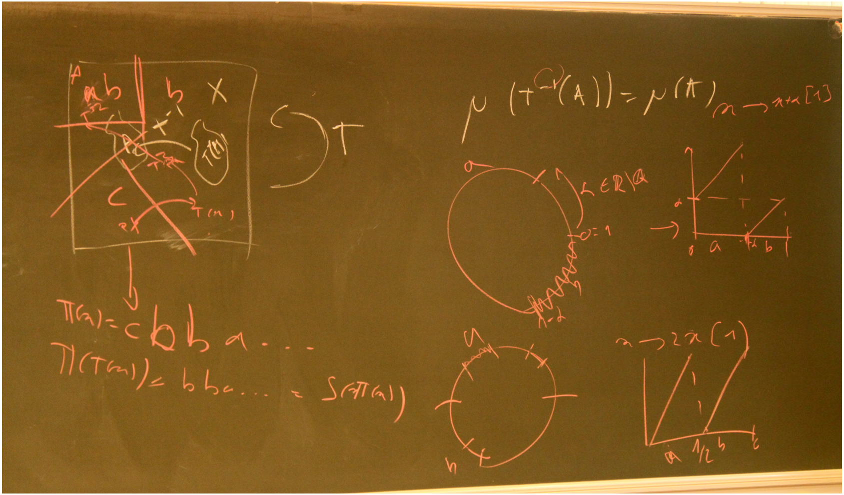


Every stationary measure $P \in \mathcal{P}(X)$ is a barycenter of the ergodic measures: there is a probability measure $m_P \in \mathcal{P}(\mathcal{P}(X))$ supported on the ergodic measures such that

$$P(w) = \int Q(w)\,\mathrm{d}m_P(Q)$$

for every $w \in \{0,1\}^*$.

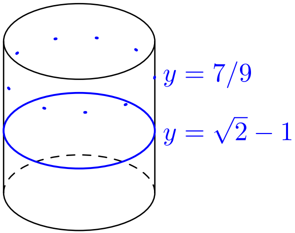
Dynamical systems



(thanks to Thierry)

Dynamical systems

- $X = S \times [0,1]$ where $S = [0,1] \mod 1$ is the unit circle.
- $T(x,y) = (x + y \mod 1, y)$.



1. Ergodic decomposition

2. Randomness and Computability
    a. Effective decomposition
    b. The ergodic case

Randomness

Let $P = B_{1/2}$.

$000000000000000000000\dots$	non-random
$101101101101101101101\dots$	non-random
$010111010001001010110 0\dots$	possibly random

Martin-Löf, 1966

To each probability measure P is associated the set $\mathcal{R}_P$ of P-***random sequences***, defined as the intersection of all the "constructive" sets of measure one.

$$P(R_P) = 1.$$

The theory can be extended to many separable metric spaces: $\mathbb{R}^n$, $C([0,1])$, $\mathcal{K}(\mathbb{R}^n)$, $\mathcal{P}(\{0,1\}^{\mathbb{N}})$, ...

Randomness vs ergodic theory

General direction

Understand the properties of the sequences that are random with respect to invariant measures.

Kučera (1985), V'yugin (1997, 1998), Nakamura (2005), Gács, Galatolo, H., Rojas (2008, 2009), Bienvenu, Day, Mezhirov, Shen (2010)

Randomness

V'yugin (1997): Bikhoff's ergodic theorem holds for random sequences.

Theorem (V'yugin, 1997)

Let P be a stationary probability measure:

- *Every P-random sequence x is generic.*
- *When P is ergodic, $P_x = P$ for every P-random x.*

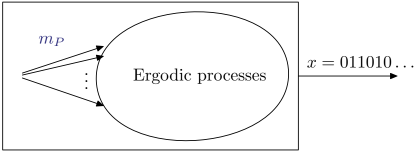
Ergodic decomposition for random sequences?

Let P be a stationary probability measure. If x be P-random, is P_x ergodic?

a. Effective decomposition

Reminder: if $P \in \mathcal{P}(X)$ is stationary, then there exists $m_P \in \mathcal{P}(\mathcal{P}(X))$ such that for every $w \in \{0,1\}^*$,

$$P(w) = \int Q(w)\,\mathrm{d}m_P(Q).$$



Definition

Let P be a computable stationary measure. The ergodic decomposition of P is ***effective*** if the measure m_P is computable.

a. Effective decomposition

Theorem

Let P be an effectively decomposable stationary measure. The following statements are equivalent:

- *x is P-random,*
- *there is an m_P-random measure P' such that x is P'-random.*

Lemma

Every m_P-random measure is ergodic.

Corollary

If x is P-random, then

- *P_x is m_P-random,*
- *P_x is ergodic,*
- *x is P_x-random.*

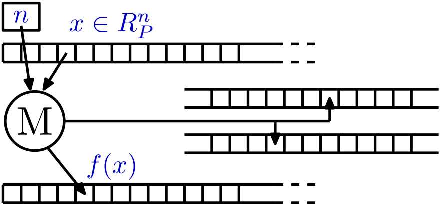
a. Effective decomposition

Reminder: decomposition of the set of random points

$$\mathcal{R}_P = \bigcup_{n\in\mathbb{N}} \mathcal{R}_P^n \quad \text{with } \mathcal{R}_P^n \subseteq \mathcal{R}_P^{n+1} \text{ and } P(\mathcal{R}_P^n) > 1 - 2^{-n}.$$

Definition

A function $f : X \to \mathbb{R}$ is P-***layerwise computable*** if there is a machine that computes (successive approximations of) $f(x)$ from x and n such that $x \in \mathcal{R}_P^n$.



a. Effective decomposition

Theorem

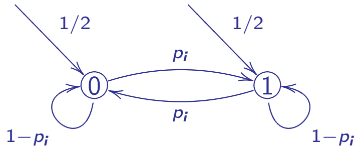
Let P be a computable stationary measure. The following statements are equivalent:

- *P is effectively decomposable (i.e., m_P is computable),*
- *the function $x \mapsto P_x$ is P-layerwise computable.*

a. Effective decomposition

A counter-example due to V'yugin (1997)

- First step: pick $i \in \{1, 2, 3, \dots\}$ with probability 2^{-i}.
  Let $p_i = 2^{-t_i}$ where t_i is the halting time of Turing machine M_i ($p_i = 0$ when M_i does not halt).
- Following steps: run the following Markov chain



The mixture $P = \sum_i 2^{-i} P_i$ is computable, but m_P is ***not*** computable.

Open question

- What about ***finitely*** decomposable invariant measures?
- Let $P = \frac{1}{2}(P_1 + P_2)$ with P_1, P_2 ergodic and $P_1 \neq P_2$. If P is computable, are P_1, P_2 computable?

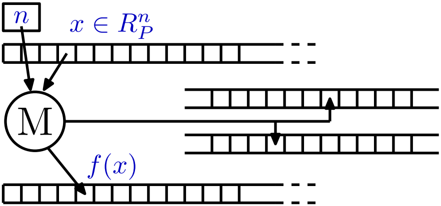
b. The ergodic case

Open question

Let P be an ergodic stationary measure, which is not computable. Is the constant function $f(x) = P$ layerwise computable?
Weaker question: given a P-random sequence x, is P computable relative to x?

Theorem

If P belongs to an effective closed set of ergodic measures, then the constant function $x \mapsto P$ is P-layerwise computable.



b. The ergodic case

Effective convergence

Theorem

The following are equivalent:

1. P *belongs to some effective closed class of ergodic measures,*
2. *there is a computable function* $n(i, w, \epsilon)$ *such that for every* $x \in \mathcal{R}_P^i$ *and every* $n \geq n(i, w, \epsilon)$,

$$\left| \frac{\#_{occ}(w, x_0 x_1 \ldots x_{n-1})}{n} - P(w) \right| < \epsilon.$$

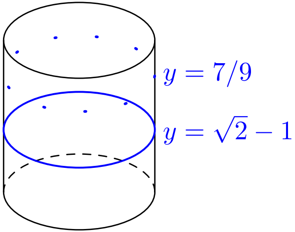
(the convergence of frequencies is P*-layerwise effective)*

Observation

Using Baire's theorem, there exist ergodic measures that do not satisfy this property.
Question: if w is fixed, is the convergence always effective?

Example

$X = S \times [0,1]$ where $S = [0,1] \mod 1$ is the unit circle.
$T(x,y) = (x + y \mod 1, y)$.



Specific answers

- Every point is generic and induces an ergodic measure $P_{(x,y)}$.
- For every (x,y), the induced measure $P_{(x,y)}$ is computable relative to (x,y).

Open questions

- If P is a computable stationary measure that has a finite decomposition (e.g. $P = \frac{1}{2}(P_1 + P_2)$), are P_1, P_2 computable?
- If P is ergodic and x is P-random, is P computable from x?
- Given a stationary measure P and a P-random sequence x,
  - is P_x always ergodic?
  - is P_x always m_P-random?
  - is x always P_x-random?
  - is P_x computable relative to x?

Open questions

- If P is a computable stationary measure that has a finite decomposition (e.g. $P = \frac{1}{2}(P_1 + P_2)$), are P_1, P_2 computable?
- If P is ergodic and x is P-random, is P computable from x?
- Given a stationary measure P and a P-random sequence x,
    - is P_x always ergodic?
    - is P_x always m_P-random?
    - is x always P_x-random?
    - is P_x computable relative to x?

Thank you!