# Outline of Lecture 1

**Martin-Löf tests and martingales**

- The Cantor space.
- Lebesgue measure on Cantor space.
- Martin-Löf tests.
- Basic properties of random sequences.
- Betting games and martingales.
- Equivalence of Martin-Löf tests and effective martingales.
- Alternative randomness concepts.
Cantor Space

We will study randomness for infinite binary sequences.

Cantor space: set of all such sequences, denoted by $2^\mathbb{N}$.

Ways to interpret sequences $X \in 2^\mathbb{N}$:
- sets of natural numbers, $S_X = \{ n \in \mathbb{N} : X(n) = 1 \}$,
- real numbers in $[0, 1]$, $\alpha_X = \sum_n X(n)2^{-n}$.

Metric

$$d(X, Y) = \begin{cases} 2^{-N(X, Y)} & \text{if } X \neq Y \\ 0 & \text{if } X = Y. \end{cases}$$

where $N(X, Y) = \min\{n : X(n) \neq Y(n)\}$. 
Cantor Space

Topological properties of $2^\mathbb{N}$

- compact
- perfect
- totally disconnected

$2^\mathbb{N}$ has a countable basis of clopen sets, the so-called cylinder sets

$$[\sigma] = \{X : X \upharpoonright_n = \sigma\},$$

where $\sigma$ is a finite binary sequence (string) and $X \upharpoonright_n$ denotes the first $n$ bits of $X$.

The open subsets of $2^\mathbb{N}$ are unions of cylinder sets. They can be represented by a set $W \subseteq 2^{<\mathbb{N}}$. We write $[W]$ to denote the open set induced by $W$. 
Lebesgue Measure on Cantor Space

Over $\mathbb{R}$: Lebesgue measure $\lambda$ unique Borel measure that is translation invariant and assigns every interval $(a, b)$ measure $|b - a|$.

Over $2^\mathbb{N}$:

- **Diameter** of a basic open cylinder $[[\sigma]]$ is $2^{-|\sigma|}$.

Hence we will set $\lambda[[\sigma]] = 2^{-|\sigma|}$.

Some basic results of measure theory ensure that $\lambda$ can be uniquely extended to all Borel sets.

- *We will return to this in more detail in Lecture 4.*
Lebesgue Measure on Cantor Space

Alternative view of Lebesgue measure:

- \( \chi \mapsto \alpha_\chi = \sum_n \chi(n)2^{-n} \) yields a surjection of \( 2^\mathbb{N} \) onto \([0, 1]\).
- The image of \([\sigma]\) is the dyadic interval
  \[
  \left[ \sum_{k=0}^{n-1} \sigma(k)/2^{k+1}, 2^n + \sum_{k=0}^{n-1} \sigma(k)/2^{k+1} \right].
  \]
- The Lebesgue measure (in \( \mathbb{R} \)) of this interval is \( 2^{-n} \).
Lebesgue Measure on Cantor Space

Yet another view:

- $X \in 2^\mathbb{N}$ represents outcome of an infinite sequence of coin tosses – 0 is Heads, 1 is Tails.
- If the coin is fair, each outcome has probability $1/2$.
- A finite string $\sigma$ represents the outcome of a finite number of independent coin tosses.
- The probability of outcome $\sigma$ is $(1/2)^{|\sigma|}$. 
Nullsets

Nullsets are sets that are measure theoretically small, just as countable sets are small with respect to cardinality.

Intuitively, a nullset is a set that can be covered by open sets of arbitrary small measure.

**Definition**

A subset $A \subseteq 2^\mathbb{N}$ is a nullset for Lebesgue measure (or has Lebesgue measure zero) if for every $\varepsilon > 0$ there exists an open set $U = \bigcup_{\sigma \in W} [\sigma]$ such that

$$A \subseteq U \quad \text{and} \quad \sum_{\sigma \in W} \lambda[\sigma] = \sum_{\sigma \in W} 2^{-|\sigma|} < \varepsilon.$$
Nullsets

To define Martin-Löf tests, it is convenient to reformulate this a little.

**Proposition**

A set $A \subseteq 2^\mathbb{N}$ is a nullset iff there exists a set $W \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ such that, if we let $W_n = \{\sigma: (n, \sigma) \in W\}$, for all $n \in \mathbb{N}$,

$$A \subseteq [W_n] \quad \text{and} \quad \sum_{\sigma \in W_n} 2^{-|\sigma|} < 2^{-n}.$$ 

$\cap_n [W_n]$ is itself a nullset. It is an intersection of a sequence of open sets. Such sets are called $G_\delta$ or $\Pi_2^0$-sets.

$\Rightarrow$ Every nullset is contained in a $G_\delta$ nullset.
Nullsets

Remarks

- We can always assume the sequence \((W_n)\) is nested. (Why?)
- \(G_\delta\) sets can be easily effectivized. What ‘codes’ a \(G_\delta\) set in Cantor space is a subset of \(\mathbb{N} \times 2^{<\mathbb{N}}\).
- On such sets we can easily impose definability/effectivity conditions, e.g. require that they are recursively enumerable.
Martin-Löf Tests and Randomness

**Definition**

- A **Martin-Löf (ML) test** (for Lebesgue measure) is a recursively enumerable set \( W \subseteq \mathbb{N} \times 2^{<\mathbb{N}} \) such that, if we let \( W_n = \{\sigma: (n, \sigma) \in W\} \), for all \( n \in \mathbb{N} \),

  \[
  \sum_{\sigma \in W_n} 2^{-|\sigma|} < 2^{-n}.
  \]

- A set \( A \subseteq 2^\mathbb{N} \) is **Martin-Löf null** if it is covered by a Martin-Löf test, i.e. if there exists a Martin-Löf test \( W \) such that \( A \subseteq \bigcap_n [W_n] \).
- A sequence \( X \in 2^\mathbb{N} \) is **Martin-Löf random** if \( \{X\} \) is not Martin-Löf null.
Existence of Random Sequences

- Every ML-test $W$ describes a $G_δ$ nullset, with the additional requirement that it is effectively presented ($W$ is r.e.).
- There are only countably many r.e. sets, and hence only countably many many ML-tests.
- Being random means not being contained in the union of all $G_δ$ sets defined by any ML test.
- A basic result of measure theory says that a countable union of nullsets is again a nullset (the standard “$ε/2^n$-proof”).
- Therefore, the set of all non-random sequences is a nullset, and consequently, $\lambda$-almost every sequence is ML random.
Universal Tests

In the last argument, we used that a countable union of nullsets is a nullset.

It turns out that even more is true: The union of all ML-tests is again a ML-test, a universal test.

- There exists a ML-test \((\mathcal{U}_n)\) such that \(X\) is ML-random iff \(X\) is not covered by \((\mathcal{U}_n)\).
- In other words, the ML-random sequences are precisely the ones in the complement of \(\bigcap_n [\mathcal{U}_n]\).
- The ML-random sequences form the largest effective (in the sense of Martin-Löf) set of measure 1.
Universal Tests

Construction of a universal test

- Start uniformly enumerating all r.e. subsets \( W^{(e)} \) of \( \mathbb{N} \times 2^{\mathbb{N}} \).
- Once we see that the measure condition of some \( W_{n}^{(e)} \) is violated, we stop enumerating it.
- Given a uniform enumeration of all tests \( (\tilde{W}_{n}^{(e)}) \) (with possible repetitions), we can define a universal test \( (U_{n}) \) by letting

\[
U_{n} = \bigcup_{e} \tilde{W}_{n+e+1}^{(e)}
\]

Note that this test has the nice property that it is nested, i.e. \( [U_{n+1}] \subseteq [U_{n}] \). We will always assume this from now on.

Later we will encounter other ways to define universal tests.
Basic Properties of Random Sequences

- The set of Martin-Löf random reals is **invariant under prefix operations** (adding, deleting, replacing a finite prefix).
- If $Z \subseteq \mathbb{N}$ is **computably enumerable**, then the sequence given by the characteristic function of $Z$ is **not Martin-Löf random**.
- **Any finite string appears** somewhere in a Martin-Löf random real, in fact it appears **infinitely often** in a Martin-Löf random real.
- For every Martin-Löf random sequence $X \in 2^{\mathbb{N}}$,
  $$\lim_{n} \frac{\sum_{k=0}^{n-1} X(i)}{n} = \frac{1}{2}.$$  

These assertions can be proved directly by defining a suitable test. (Exercise!) But we will prove different characterizations of random sequences which may make this easier.
Betting Games and Martingales

Betting strategies

A betting strategy $b$ is a function $b : 2^{\mathbb{N}} \to [0, 1] \times \{0, 1\}$.

Interpretation:

- A string $\sigma$ represents the outcomes of a 0-1-valued (infinite) process (e.g. a coin toss).
- $b(\sigma) = (\alpha, i)$ then tells the gambler on which outcome to bet next, $i$, and what percentage of his current capital to bet on this outcome, $\alpha$.
- When the next bit of the process is revealed and agrees with $i$, the capital is multiplied by $(1 + \alpha)$. If it is different from $i$, the gambler loses his bet, i.e. his capital is multiplied by $(1 - \alpha)$.
Betting Games and Martingales

We can keep track of the player’s capital through a function $F : 2^{\mathbb{N}} \rightarrow [0, \infty)$. 

$F$ satisfies

$$F(\sigma) = \frac{F(\sigma 0) + F(\sigma 1)}{2} \quad \text{for all } \sigma.$$ (*)

This reflects the property that the game is fair – the expected value of the capital after the next round is the same as the player’s capital before he makes his bet.

Any function satisfying (*) is called a martingale.

Given a martingale, we can reconstruct the accordant betting function from it.
Betting Games and Martingales

Successful martingales

A martingale is successful on an infinite sequence $X$ if

$$\limsup_{n \to \infty} F(X | n) = \infty,$$

We can actually replace $\limsup$ by $\lim$:

- For every martingale $F$ there exists a martingale $G$ such that for all $X$,

  $$\limsup_{n} F(X | n) = \infty \quad \text{implies} \quad \lim_{n} G(X | n) = \infty.$$

(Set some money aside regularly.)
Betting Games and Martingales

A martingale succeeds only on very few sequences.

Martingale Convergence Theorem [Ville, Doob]

For any martingale $F$, the set of sequences $X \in 2^\mathbb{N}$ such that

$$\limsup_{n \to \infty} F(X | n) = \infty$$

has $\lambda$-measure zero.

We will prove an effective version of this theorem.
From ML-tests to Martingales

Goal: Given a ML-test \((u_n)\), define a martingale succeeding on the sequences covered by \((u_n)\).

Basic Idea: Whenever a string appears at level \(n\) of the test, \(F\) reaches a value of at least \(n\).

- For a single string \(\sigma\), define the following martingale.

  \[
  F_\sigma(\tau) = \begin{cases} 
  2^{-|\sigma| - |\tau|} & \text{if } \tau \subseteq \sigma, \\
  1 & \text{if } \tau \supseteq \sigma, \\
  0 & \text{otherwise}.
  \end{cases}
  \]

- \(F_\sigma\) starts out with a capital of \(2^{-|\sigma|}\) and doubles its capital every step along \(\sigma\), then stops betting.
- If an outcome is not compatible with \(\sigma\), its capital is lost.
From ML-tests to Martingales

- Now, for one level $U_n$ of the ML-test, blend the individual “string”-martingales into one.
- If $(F_n)$ is a sequence of martingales and $\sum_n F(\epsilon) < \infty$, then
  \[ F = \sum_n F_n \]
  is a martingale.
- Hence define
  \[ F_n(\tau) = \sum_{\sigma \in U_n} F_\sigma(\tau). \]
  and check that the sum of the $F_\sigma(\epsilon)$ is finite.
  - $F_\sigma(\epsilon) = 2^{-|\sigma|}$.
  - Hence $F_n(\epsilon) = \sum_{\sigma \in U_n} F_\sigma(\epsilon) = \sum_{\sigma \in U_n} 2^{-|\sigma|} \leq 2^{-n}$. 
From ML-tests to Martingales

- The inequality $F_n(\varepsilon) \leq 2^{-n}$ further lets us combine the martingales for each $U_n$ into one martingale $F$,

$$F(\sigma) = \sum_{n} F_n(\sigma).$$

- If $X \in \bigcap_n [U_n]$, there exists a sequence $(\sigma_n)$ such that for all $n$, $\sigma_n \in U_n$ and $\sigma_n \subset X$.

- It follows that $F_n(\sigma_n) \geq 1$.

- More importantly, by the definition of $F_n$, $F_n(\tau) \geq 1$ for all $\tau \supseteq \sigma$, hence in particular for all $\sigma_m$ where $m \geq n$.

- It follows that for all $n$, $F(\sigma_n) \geq \sum_{k=1}^{n} F_k(\sigma_n) \geq n$, that is, $F$ is unbounded along $X$. 
Left.enumerable Martingales

What is the computational complexity of $F$?

- A function $F : 2^{<\mathbb{N}} \rightarrow \mathbb{R}$ is enumerable from below or left-enumerable if there exists, uniformly in $\sigma$, a recursive nondecreasing sequence $(q_k^{(\sigma)})$ of rational numbers such that $q_k^{(\sigma)} \rightarrow F(\sigma)$.

- Equivalently, the left cut of $F(\sigma)$ is uniformly enumerable, i.e. the set

  $$\{(q, \sigma) : q < F(\sigma)\}$$

  is recursively enumerable.

It is not hard to see that $F$ defined above is left-enumerable.
Left-enumerable Martingales

We have proved the following:

For any ML-test \((u_n)\) there exists a left-enumerable martingale \(F\) such that if \(X \in \bigcap_n [u_n]\), then \(F\) succeeds on \(X\).

In other words, if \(X\) is not ML-random, we can find a left-enumerable martingale that succeeds on \(X\).
From Martingales to ML-Tests

Does a converse of this hold? Can we transform a left-enumerable martingale $F$ into a ML-test?

Basic idea: Whenever $F$ first reaches a capital of $2^n$ on some string $\sigma$, enumerate $\sigma$ into $U_n$.

- Since $F$ is enumerable from below, this is an r.e. event.
- We only have to make sure that there are not too many such $\sigma$.
- This is guaranteed by Kolmogorov’s inequality (actually due to Ville).

Suppose $F$ is a martingale. For any string $\sigma$ and any prefix-free set $W$ of strings extending $\sigma$,

$$F(\sigma) \geq \sum_{\tau \in W} 2^{||\sigma|-|\tau||} F(\tau)$$

- Prefix-free: No two strings are comparable by $\subset$. 

From Martingales to ML-Tests

From Kolmogorov’s inequality we get the desired result:

Given a martingale $F$, let $C_k(F) = \{\sigma: F(\sigma) \geq k\}$. Then

$$\lambda[\mathcal{C}_k(F)] \leq F(\varepsilon)/k.$$ 

- Let $W$ be a prefix-free set such that $\llbracket W \rrbracket = \llbracket C_k(F) \rrbracket$. (This can be found effectively.)
- Then $\lambda[\mathcal{C}_k(F)] = \lambda[\mathcal{W}] = \sum_{\tau \in W} 2^{-|\tau|}$.
- By Kolmogorov’s inequality, $F(\varepsilon) \geq \sum_{\tau \in W} 2^{-|\tau|} F(\tau) \geq \sum_{\tau \in W} 2^{-|\tau|} k$.
- Hence $\lambda[\mathcal{C}_k(F)] \leq F(\varepsilon)/k$, as required.
From Martingales to ML-Tests

We have proved the first main theorem of algorithmic randomness, due to Schnorr and independently Levin.

**Theorem**

A sequence $X$ is ML-random if and only if no left-enumerable martingale succeeds on it.
Alternative Randomness Concepts

Of course, ML-tests are not the only possible way to effectivize nullsets.

ML-randomness is the most prominent concept because it shows a rather strong robustness with respect to the different approaches.

We will briefly discuss a few other notions – some based on martingales, others based on tests.
Alternative Randomness Concepts

Test-based concepts

- Weak 2-randomness
- Schnorr randomness

Martingale-based concepts

- Computable randomness
- Resource-bounded randomness
Weak 2-Randomness

Martin-Löf test has to fulfill two effectivity requirements.

- uniform recursive enumerability of $(W_n)$,
- measure of the $W_n$ converges to 0 effectively, $\lambda[W_n] \leq 2^{-n}$.

For a weak 2-test we only require that $(W_n)$ is uniformly r.e. and that $\lambda \bigcap_n [W_n] = 0$.

One can show that weak 2-randomness is strictly stronger than ML-randomness. There exists an $X$ that is ML-random but not weak 2-random.

We will encounter such an example later.
Schnorr Randomness

On the other hand, one might argue that the effectivity requirement for ML-tests is too weak. Test should be computable in some form, not merely r.e.

Schnorr tests

A ML-test \( (W_n) \) is a Schnorr test if the real number

\[
\sum_{\sigma \in W_n} 2^{-|\sigma|}
\]

is computable uniformly in \( n \).

A real number \( \alpha \) is computable if there exists a computable function \( g : N \rightarrow \mathbb{Q} \) such that for all \( n \), \( |\alpha - g(n)| \leq 2^{-n} \).

Note: If \( (W_n) \) is a Schnorr test then the sets \( W_n \) are uniformly computable.
Computable Randomness

The same criticism applies to the martingale characterisation of randomness. Betting strategies should be computable [Schnorr].

**Definition**

A sequence $X$ is **computably random** if no computable martingale succeeds on it.

A function $F : 2^{<N} \rightarrow \mathbb{R}$ is **computable** if there exists a computable function $g : 2^{<N} \times N \rightarrow \mathbb{Q}$ such that for all $\sigma, n$,

$$|F(\sigma) - g(\sigma, n)| \leq 2^{-n}.$$ 

One can refine the computability requirement even further, by imposing a time-bound on $F$. This leads to the theory of **resource-bounded measure**, which has successfully been used in computational complexity.
Relations between Randomness Concepts

The following **strict implications** hold:

\[ X \text{ weak 2-random} \quad \Downarrow \quad X \text{ ML-random} \quad \Downarrow \quad X \text{ computably random} \quad \Downarrow \quad X \text{ Schnorr random} \]