PROBA

David Wiedemann

Table des matières

List of Theorems

Lecture 1: Introduction

Wed 22 Sep

1 Some historical models

1.1 Laplace Model

Definition 1 (Laplace Model)
Ω finite set, $|\Omega| = n$ is the set of outcomes.
We can observe whether $E \subset \Omega$ happens, and we define it's probability

$$\mathbb{P}(E) = \frac{|E|}{|\Omega|}$$

Question

Why should this have any meaning/content ?

Proposition 1
Consider laplace model for n coint tosses $\Rightarrow$ every sequence has probability 2^{-n}
Denote by H_n the number of heads in n tosses

$$\mathbb{P}(|\frac{H_n}{n} - \frac{1}{2}| > \epsilon) \to 0$$

More generally

Proposition 2
If you have a laplace modelfor some event E, and look at n repetitions, then

$$\forall \epsilon > 0 \mathbb{P}(|\frac{E_n}{n} - \mathbb{P}(E)| > \epsilon) \to 0$$

Limitations of Laplace Model

- All outcomes have equal probability ?
- Need $|\Omega| < \infty$, so what about infinite sets ?

What next ?

Definition 2 (Intermediate model)
Let Ω to be any set and $P : \Omega \to [0, 1]$, s.t. $\sum_{\omega\in\Omega} p(\omega) = 1$

Event : $E \subset \Omega$ and

$$\mathbb{P}(E) := \sum_{\omega\in E} p(\omega)$$

- More freedom
- If you take Ω finite, $p(\omega) = \frac{1}{|\Omega|} \Rightarrow$ Laplace model
- Price ? How to choose $p : \Omega \to [0, 1] \to$ collect data, do statistics
- keeps many nice properties

— For contable sets, this is equivalent to the standard model.
— For uncountable Ω ?
— Problem 1 : There is no function s.t.

$$p(\omega) > 0 \forall \omega \in \Omega \text{ and } \sum p(\omega) = 1$$

This intermediate model is in essence only for countable sets.

What about uncountable sets ?

— What about a random point int $[0,1]$ or $[0,1]^n$?

Intuitively, consider $[0,1]$, then we can set

$$\mathbb{P}(A) = \text{length}(A)$$

Definition 3 (Geometric probability)
Take $f : \mathbb{R} \to (0, \infty)$ to be a riemann-integrable function with total mass 1. For any $A \subset \mathbb{R}$,s.t. 1_A riemann-integrable, we set $\mathbb{P}(A) = \int_A f(x)dx$

— In general quite <u>ok</u>
BUT

— You would expect there is one framework for uncountable and countable sets.
— What about more complicated spaces (eg. space of continuous functions)
— $\mathbb{P}(\mathbb{Q})$ is undefined

2 Basic Formalism

2.1 Measure spaces : A notion of area

— Set + structure
— General setting to talk about area

Definition 4 (Measure space)
$(\Omega, \mathcal{F}, \mu)$ is called a measure space if :
— *Ω is some set*
— *$\mathcal{F} \subset P(\Omega)$ called a σ-algebra*
  — $\emptyset \in \mathcal{F}$
  — $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$
  — *$F_1, F_2, \ldots, \in \mathcal{F}$, then $\bigcup_{i \geq 1} F_i \in \mathcal{F}$ each F is called a measurable set.*
— *$\mu : \mathcal{F} \to [0, \infty)$ called the measure*
  — $\mu(\emptyset) = 0$

— *If $F_1, \dots,$ are disjoints sets of the σ-algebra, then*

$$\mu(\bigcup_{i\geq 1} F_i) = \sum_{i\geq 1} \mu(F_i)$$

— *Defined by Borel 1898 and Lebesgue 1901-1903*

2.2 Probability spaces

Given by Kolmogorov in 1933

Definition 5 (Probability space)
A triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space if it is a measure space and $\mathbb{P}(\Omega) = 1$

Interpretation

— Ω state space/universe
— $\mathcal{F}$ is the set of events you can observe/have access to
— $\mathbb{P}(E)$ is the probability of E

Lemme 3
Let $(\Omega, \mathcal{F}, \mu)$ be a measure space
— $\Omega \in \mathcal{F}$
— $F_1, F_2 \in \mathcal{F} \Rightarrow F_1 \setminus F_2 \in \mathcal{F}$
— $F_1, \dots \in \mathcal{F} \Rightarrow \bigcap F_i \in \mathcal{F}$
— $F_1, F_2, \dots \in \mathcal{F} \Rightarrow \bigcap_{i\geq 1} F_i$

Let us compare this definition with the prior ones
— Ω finite set, $\mathcal{F} = \mathcal{P}(\Omega), \mathbb{P}(F) = \frac{|F|}{|\Omega|}$ this is a probability space and a laplace model.
— For Ω countable, $\mathcal{F} = \mathcal{P}(\Omega), \mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\omega)$
— The really new part is $\mathcal{F}$ which restricts the sets we can measure

Lecture 2: ...

Wed 29 Sep

2.3 Basic properties

— $F_1, F_2, \dots, \in \mathcal{F}$ disjoint

$$\mu(\bigcup F_i) = \sum \mu(F_i)$$

— $F_1 \subset F_2 \in \mathcal{F}$ $\mu(F_1) \leq \mu(F_2)$
— $F_1 \subset F_2 \dots \in \mathcal{F}$

$$\mu(F_n) \to \mu(\bigcup F_i)$$

— $F_1, F_2, \dots, \mathcal{F}$

$$\mu(\bigcup F_i) \leq \sum \mu(F_i)$$

In addition, in probability spaces

— $\mathcal{P}(F^c) = 1 - \mathcal{P}(F)$

— $F_1 \supset F_2 \supset \dots \Rightarrow \mathcal{P}(F_n) \to \mathcal{P}(\bigcap F_i)$

2.4 Measurable and measure preserving maps

Definition 6

Let $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ two measure spaces.
$f : \Omega_1 \to \Omega_2$ is called measurable if for every $F \in \mathcal{F}_2$, $f^{-1}(F) \in \mathcal{F}_1$
A measurable function $f : (\Omega_1, \mathcal{F}_1) \to (\Omega_2, \mathcal{F}_2)$ is called measure preserving if $\forall F \in \mathcal{F}_2$ $\mu_1(f^{-1}(F)) = \mu_2(F)$.

Lemme 4 (Push-Forward measure)
Let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1), (\Omega_2, \mathcal{F}_2)$ be two measure spaces, and f measurable, then $\mathbb{P}_2(F) = \mathbb{P}_1(f^{-1}(F))$ is a probability measure.

3 Probability spaces

— Discrete probability spaces : Ω countable

— Continuous probability spaces : Ω uncountable.

3.1 Discrete probability spaces

Does introducing a σ-algebra $\mathcal{F}$ enlargen the generality ?

Proposition 5
If $(\Omega, \mathcal{F}, \mathbb{P})$ is a discret probability space, $\exists \Omega_2$ countable, $\mathbb{P}_2 : \mathcal{P}(\Omega_2) \to [0, 1]$ s.t. $(\Omega_2, P(\Omega_2), \mathbb{P}_2)$ is a probability space and $\exists f : (\Omega_1, \mathcal{F}_1) \to (\Omega_2, \mathcal{F}_2)$ is measure preserving

Still $\mathcal{F}$ is useful :

— can sequentially study a model/situation by taking $\mathcal{F}_1 \subset \mathcal{F}_2 \dots$

Lemme 6
There is no shift-invariant probability measure on $(\mathbb{Z}, P(\mathbb{Z}))$

Preuve

$$\mathbb{P}(\mathbb{Z}) = \mathbb{P}(\bigcup_n \{n\}) = \sum \mathbb{P}(\{n\}) = \infty$$

$\Rightarrow$ *cannot treat everyone on an equal ground !* $\square$

3.1.1 Symmetric simple random walk

A simple walk of length n s.t. $|s_n - s_{n-1}| = 1$.
Let Ω be the set of all walks of length n, and consider $(\Omega, P(\Omega), \mathbb{P})$.
What is the probability that S hits 0 ?
What does it look like, what is it's max ?

3.2 Continuous probability spaces

Can we define a probability measure on S^1 s.t. $(S^1, P(S^1))$ that is rotation invariant ?
Similarly to the countable case, but not the same as Ω is uncountable and setting $P(\{\omega\}) = 0$ gives no contradiction.

Proposition 7
You can not.

Preuve
Idea : decompose S^1 into countable many sets A_n st $\bigcup A_n = S^1$, they are disjoint and rotations of each other.
$\forall x \in S^1$, define S_x as $\{\ldots, T^{-2}x, T^{-1}x, x, Tx, \ldots\}$.
Note that either $S_x = S_y$ or $S_x \bigcap S_y = \emptyset$.

Lecture 3: Measurable maps

Wed 06 Oct

3.3 Borel σ-algebra

- Cannot define shift-invariant probability measure on $([0,1], \mathcal{P}([0,1]))$.
- What σ-algebra to choose on (X, τ) ?
- Want to know the siize of all open-sets

Definition 7 (Borel sigma-algebra)
On (X, τ) the borel σ-algebra $\mathcal{F}_\tau$ is the smallest σ-algebra containing τ.

This is well defined because, given a collection of σ-algebras, their intersection is too.

Two nice properties

- Continuous functions on a Borel σ-algebra are also measurable.

  Preuve
  Suffices to check that $f^{-1}(U) \in \mathcal{F}_{\tau_1}$ for $U \in \tau_2$ but this is immediate since f is continuous.

In $(\mathbb{R}^n, \tau_E)$, the Borel σ-algebra $\mathcal{F}_E$ is generated by $(a_1, b_1) \times ... \times (a_n, b_n)$. $\mathcal{F}_E$ is the smallest σ-algebra containing open intervalls. $\square$

3.4 Probability Measures on $\mathbb{R}^n$

Theorème 8 (Existence of Lebesgue-measure)
There exists a unique measure λ on $(\mathbb{R}^n, \mathcal{F}_E)$ s.t. $\lambda((a_1 \times b_1) \times \ldots \times (a_n, b_n)) = \prod_i |b_i - a_i|$

Theorème 9 (Uniforme Measure)
There exists a unique $\mathbb{P}$ measure on $([0,1]^n, \mathcal{F}_E)$ with the same property.

Both λ and $\mathbb{P}$ are shift-invariant in fact only shift invariant measures on $\mathbb{R}$ (up to a constant)

Preuve
Consider the case of $(\mathbb{R}^n, \mathcal{F}_E)$ and $f_r : x \to x + \tau, \tau \in \mathbb{R}^n$.
- *f_r continuous $\Rightarrow$ measurable*
- *$\tilde{\mathbb{P}}(A) = \mathbb{P}(f^{-1}(A))$ is a probability measure*
- *All boxes have the same measure* $\square$

3.5 Probability measures on $(\mathbb{R}, \mathcal{F}_E)$

We saw that we can put a uniform measure on $[0, 1]$.
All probability measures on $(\mathbb{R}, \mathcal{F}_E)$

1. $\mathbb{P} : \mathcal{F}_E \to [0, 1]$
2. These are actually only characterized by $\mathbb{P}((-\infty, x))$

Definition 8 (Cumulative distribution function)
$F : \mathbb{R} \to [0, 1]$ is called a c.d.f if
- *F is non-decreasing*
- *$F(x_n) \to 0$ then $x_n \to -\infty$*
- *$F(x_n) \to 1$ if $x_n \to 1$*
- *F is right-continuous.*

Theorème 10
Given a probability measure $\mathbb{P}$ on $(\mathbb{R}, \mathcal{F}_E)$, then $f(x) := \mathbb{P}((-\infty, x))$ is a c.d.f
Given a c.d.f, there exists a unique probability measure s.t. $\mathbb{P}(-\infty, x) = F(x)$

Preuve
Given $\mathbb{P}$ on $(\mathbb{R}, \mathcal{F}_E)$.
Let's show that $F(x) = \mathbb{P}((-\infty, x))$ is a c.d.f.

— $x < y \quad F(x) = \mathbb{P}((-\infty, x)) \leq \mathbb{P}(-\infty, y) = F(y)$

— $x_n \to -\infty \quad F(x_n) = \mathbb{P}(-\infty, x_n) \to \mathbb{P}(\bigcap_n (-\infty, x_n)) = 0$

— $x_n \to \infty \Rightarrow F(x_n) \to 1$ *is similar*

— *Also for right continuous* $x_n \to x$ *, we have that* $[x_n, \infty) \subset [x_{n+1}, \infty)$

How do we construct $\mathbb{P}$ *given* F *?*

Trick using push-forward measure.

Define $f : (0,1) \to \mathbb{R}$*, define*

$$f(x) = \inf_{y \in \mathbb{R}} \{F(y) \geq x\}$$

□

Define $\mathbb{P}(A) := \mathbb{P}_U(f^{-1}(A)) \forall A \in \mathcal{F}_E$ *Why is* f *measurable ?*

If f *is increasing* $\Rightarrow$ f *is measurable*

Lecture 4: ...

Wed 13 Oct

Each c.d.f gives rise to a unique $\mathbb{P}$.

A priori $\mathbb{P}_1 = \mathbb{P}_2$ means $\forall F \in \mathcal{F}_E \mathbb{P}_1(F) = \mathbb{P}_2(F)$.

We show that it suffices to show that $\mathbb{P}_1((-\infty, x]) = \mathbb{P}_2((-\infty, x]) \forall x \in \mathbb{R}$.

Lemme 11

Given $(\mathbb{R}, \mathcal{F}_E, \mathbb{P})$ *then* $\forall B \in \mathcal{F}_E, \forall \epsilon > 0$ *one can find disjoint intervals* $I_1, \ldots, I_n$ *s.t.* $\mathbb{P}(B \Delta (I_1 \cup \ldots \cup I_n)) < \epsilon$

Preuve

Consider the collection H *of all subsets* $H \in \mathcal{F}_E$ *s.t. the property above holds.*

We know that H *contains all intervalls, hence* $\sigma(H) = \mathcal{F}_E$.

So we only need to show that H *is a* $\sigma-$*algebra*

1. $\emptyset \in H$ *: Know that* $\forall x (-\infty, x] \in H$
2. *If* $B \in H \Rightarrow B^C \in H$.
   Given $\epsilon > 0$*, choose* $I_1, \ldots, I_n$ *s.t.* $\mathbb{P}(B \Delta (I_1 \cup \ldots)) < \epsilon$*, but* $(B \Delta A) = B^C \Delta A^C$ *, hence*
   $$\mathbb{P}(B^C \Delta (I_1 \cup \ldots)) < \epsilon$$
3. $H_1, \ldots \in H$*, we want* $\bigcup_i H_i \in H$ $\exists n \in \mathbb{N}$
   $$\mathbb{P}((\bigcup_{i=0}^{m} H_i) \Delta (\bigcup_i H_i)) < \frac{\epsilon}{2}$$
   $\forall i = 1, \ldots, m$ *, we have disjoint* $I_{i,1}, \ldots, I_{i,m_i}$ *s.t.*
   $$\mathbb{P}(H_i \Delta (I_{i,1} \cup \ldots)) < \frac{\epsilon}{2m}$$
   Now use that
   $$(\bigcup_{i=1}^{m} H_i) \Delta (\bigcup_{i=1}^{m} \bigcup_{j=1}^{m_i} I_{i,j}) \subseteq \bigcup_{i=1}^{m} (H_i \Delta \bigcup_{j=1}^{m_i} I_{i,j})$$
   Finally, we can write a finite union of disjoint intervals □

Corollaire 12
$\mathbb{P}_1, \mathbb{P}_2$ probability measure on $(\mathbb{R}, \mathcal{F}_E)$, then $\mathbb{P}_1 = \mathbb{P}_2$ as soon as

$$\mathbb{P}_1((-\infty, x]) = \mathbb{P}_2((-\infty, x])$$

or

$$\mathbb{P}_1(x, y) = \mathbb{P}_2(x, y)$$

Preuve
Notice $(-\infty, x)$ can be written as

$$(-\infty, x) = (\bigcup_n (x, x+n))^C$$

So it suffices to prove the first point.
Observe, for all intervalls $\mathbb{P}_1(I) = \mathbb{P}_2(I)$ since

$$\mathbb{P}_i(y, x) = \mathbb{P}_i(-\infty, x) - \mathbb{P}_i(-\infty, y)$$

The condition holds for B if $\forall \epsilon > 0$, we can pick $I_1, \ldots, I_n$ s.t.

$$\mathbb{P}_1(B\Delta(I_1 \cup \ldots)) < \epsilon$$

and

$$\mathbb{P}_2(B\Delta(I_1 \cup \ldots)) < \epsilon$$

So we need to check again that this is a $\sigma-$ algebra and we are done.
Now we can conclude that

$$|\mathbb{P}_1(B) - \mathbb{P}_1(I_1 \cup \ldots)| = |\mathbb{P}_1(B) - \mathbb{P}_2(I_1 \cup \ldots)| < \epsilon$$

and

$$|\mathbb{P}_2(B) - \mathbb{P}_1(I_1 \cup \ldots)| = |\mathbb{P}_2(B) - \mathbb{P}_2(I_1 \cup \ldots)| < \epsilon \qquad \square$$

An abstract uniqueness result follows from a similar strategy.

Theorème 13 (Dynkin)
$\mathbb{P}_1$ and $\mathbb{P}_2$ two probability measures on $(\Omega, \mathcal{F})$, suppose $\mathbb{P}_1(H) = \mathbb{P}_2(H)$ for all $H \in \mathcal{H} \subset \mathcal{F}$ and

- $\sigma(H) = \mathcal{F}$
- $H_1 \in \mathcal{H}, H_2 \in \mathcal{H} \Rightarrow H_1 \cap H_2 \in \mathcal{H}$

Then $\mathbb{P}_1 = \mathbb{P}_2$

3.6 Probability measures on $\mathbb{R}^n$

Definition 9 (Joint c.d.f.)
$F : \mathbb{R}^n \to [0, 1]$

— *F non-decreasing in each coordinate*
— *$F(x_1, \ldots, x_n) \to 1$ if all $x_i \to -\infty$*
— *right-continuous*

Theorème 14
Joint c.d.f $\iff$ $\mathbb{P}$ on $(\mathbb{R}^n, \mathcal{F}_E)$

3.7 Product probability measures on $\mathbb{R}^n, \mathbb{R}^{\mathbb{N}}$

— Related to independence
— Natural mathematically

2 steps

— product σ-algebra
— product measure

3.7.1 Product σ-algebra

Definition 10 (Product algebra)
Let $(\Omega_i, \mathcal{F}_i)_{i\geq 1}$ measurable spaces, then the product σ-algebra $\mathcal{F}_\pi$ on $\prod_i \Omega_i$ is the σ-algebra generated by sets $F = E_1 \times \ldots \times E_n \times \Omega_{n+1} \times \ldots$, $E_i \in \mathcal{F}_i$

Remarque
— *Projections are measurable*
— *In fact, product σ-algebra s.t. all projections are measurable*

Notice on $\mathbb{R}^n$, we now have two ways to define a $\sigma-$algebra.
— Take $(\mathbb{R}^n, \tau_E)$ and induce a Borel σ-algebra
— Take n copies of $(\mathbb{R}, \mathcal{F}_E)$ and consider $\mathcal{F}_\pi$ on $\mathbb{R}^n$

3.8 Product measures

Definition 11
Given $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)_{i\geq 1}$ probability spaces $\mathbb{P}_\pi$ on $(\prod_i \Omega_i, \mathcal{F}_\pi)$ is called the product measure of $\mathbb{P}_i$.
If $\forall n \geq 1$, all sets $E = E_1 \times E_2 \times \ldots \times E_n \times \Omega_{n+1} \times \ldots$

$$\mathbb{P}_\pi(E) = \prod_{i=1}^{n} \mathbb{P}_i(E_i)$$

Lecture 5: Conditional probability

Wed 20 Oct

3.9 Infinite product spaces

Case of $(\mathbb{R}, \mathcal{F}_E, \mathbb{P}_i)_{i\geq 1}$.

Space of infinite faire coin tosses

We want the infinite product of $(\{0,1\}, P(\{0,1\}), \mathbb{P})$.
We use the uniform measure $([0,1], \mathcal{F}_E, \mathbb{P})$, for $x \in [0,1)$, $x = 0.x_1x_2\ldots$, we send $f : x \to (x_1, x_2, \ldots)$

Lemme 16
f as defined above is measurable

Preuve
Note that

— *$\mathcal{F}_\pi$ generated by $F_1 \times \ldots, F_n \times \{0,1\} \times \{0,1\}$ with $|F_i| = 1$*
— *$\mathcal{F}_E$ is generated by sets of the forme $(2^{-n}j, 2^{-n}(j+1))$.*

Moreover, $(j2^{-n}, (j+1)2^{-n})$ is in correspondence with $F_1 \times \ldots \times F_n \times \{0,1\} \times \ldots$.

Proposition 17
There existes a product probability measure on $(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\pi)$

Preuve
Consider $f : ([0,1], \mathcal{F}_E) \mapsto (\{0,1\}^{\mathbb{N}}, \mathcal{F}_\pi)$.
We define $\mathbb{P}_\pi$ as the pushforward of $\mathbb{P}_U$ under f $\square$

Lecture 6: Random Variables

Wed 27 Oct

4 Random Variables

Definition 12 (Random Variables)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.
Then $X : \Omega\mathbb{R}$ measurable as a map $(\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{F}_E)$ is called a (real) random variable.

The pushforward measure $\mathbb{P}_X(F) = \mathbb{P}(X^{-1}(F)) \forall F \in \mathcal{F}_E$ is called the law of X

Remarque
There is a more general notion of $(\Omega_2, \mathcal{F}_2)$ valued random variable.

Definition 13 (Equality of RV)
X, Y two random variables are called equal in law if

$$\mathbb{P}_X(F) = \mathbb{P}_Y(F) \forall F \in \mathcal{F}_E$$

Definition 14
X is a R.V. we call the c.d.f. of $\mathbb{P}_X$ F_X

$$F_X(s) = \mathbb{P}_X(X \le s)$$

Proposition 19

Each R.V. X gives rise to a unique c.d.f. $F_X(s) = \mathbb{P}_X(X \leq s)$ and conversely, each c.d.f. gives rise to a unique law of a probability measure

Preuve

Follows directly from the proposition relating probability measures and c.d.f. $\square$

Lemme 20

1. $\mathbb{P}_X < s = F(S^-)$
2. $\mathbb{P}_X(X = s) = F(s) - F(s^-)$
3. $\mathbb{P}_X(X \in (a, b)) = F(b^-) - F(a)$