

# PROBA

David Wiedemann

## Table des matières

<b>1</b>	<b>Some historical models</b>	<b>4</b>
1.1	Laplace Model . . . . .	4
<b>2</b>	<b>Basic Formalism</b>	<b>5</b>
2.1	Measure spaces : A notion of area . . . . .	5
2.2	Probability spaces . . . . .	6
2.3	Basic properties . . . . .	6
2.4	Measurable and measure preserving maps . . . . .	7
<b>3</b>	<b>Probability spaces</b>	<b>7</b>
3.1	Discrete probability spaces . . . . .	7
3.1.1	Symmetric simple random walk . . . . .	8
3.2	Continuous probability spaces . . . . .	8
3.3	Borel $\sigma$ -algebra . . . . .	8
3.4	Probability Measures on $\mathbb{R}^n$ . . . . .	9
3.5	Probability measures on $(\mathbb{R}, \mathcal{F}_E)$ . . . . .	9
3.6	Probability measures on $\mathbb{R}^n$ . . . . .	11
3.7	Product probability measures on $\mathbb{R}^n, \mathbb{R}^{\mathbb{N}}$ . . . . .	12
3.7.1	Product $\sigma$ -algebra . . . . .	12
3.8	Product measures . . . . .	12
3.9	Infinite product spaces . . . . .	12
<b>4</b>	<b>Random Variables</b>	<b>13</b>
4.1	Independence of RV . . . . .	16
4.2	Example of continuous random variables . . . . .	16
4.3	Transformation of random variables . . . . .	17
<b>5</b>	<b>Random Vectors</b>	<b>17</b>

## List of Theorems

1	Definition (Laplace Model) . . . . .	4
---	--------------------------------------	---

1	Proposition . . . . .	4
2	Proposition . . . . .	4
2	Definition (Intermediate model) . . . . .	4
3	Definition (Geometric probability) . . . . .	5
4	Definition (Measure space) . . . . .	5
5	Definition (Probability space) . . . . .	6
3	Lemme . . . . .	6
6	Definition . . . . .	7
4	Lemme (Push-Forward measure) . . . . .	7
5	Proposition . . . . .	7
6	Lemme . . . . .	7
7	Proposition . . . . .	8
7	Definition (Borel sigma-algebra) . . . . .	8
8	Theorème (Existence of Lebesgue-measure) . . . . .	9
9	Theorème (Uniforme Measure) . . . . .	9
8	Definition (Cumulative distribution function) . . . . .	9
10	Theorème . . . . .	9
11	Lemme . . . . .	10
12	Corollaire . . . . .	11
13	Theorème (Dynkin) . . . . .	11
9	Definition (Joint c.d.f.) . . . . .	11
14	Theorème . . . . .	12
10	Definition (Product algebra ) . . . . .	12
11	Definition . . . . .	12
16	Lemme . . . . .	13
17	Proposition . . . . .	13
12	Definition (Random Variables) . . . . .	13
13	Definition (Equality of RV) . . . . .	13
14	Definition . . . . .	13
19	Proposition . . . . .	14
20	Lemme . . . . .	14
15	Definition . . . . .	14
21	Lemme . . . . .	14
16	Definition . . . . .	14
22	Proposition . . . . .	14
17	Definition . . . . .	15
23	Proposition . . . . .	15
24	Proposition . . . . .	15
18	Definition (Independence of RV) . . . . .	16
25	Proposition . . . . .	16
19	Definition (Random variables with density) . . . . .	16

26	Theorème (Version of central limit theorem) . . . . .	17
27	Lemme . . . . .	17
28	Proposition . . . . .	17
20	Definition (Random Vector) . . . . .	17
30	Lemme . . . . .	17
31	Proposition . . . . .	18
32	Lemme . . . . .	18
33	Proposition . . . . .	18
34	Corollaire . . . . .	18
21	Definition (Random vectors with density) . . . . .	18

# 1 Some historical models

## 1.1 Laplace Model

### Definition 1 (Laplace Model)

$\Omega$  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  $n$  coin tosses  $\Rightarrow$  every sequence has probability  $2^{-n}$

Denote by  $H_n$  the number of heads in  $n$  tosses

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

More generally

#### Proposition 2

If you have a laplace model for some event  $E$ , and look at  $n$  repetitions, then

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

## Limitations of Laplace Model

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

What next ?

### Definition 2 (Intermediate model)

Let  $\Omega$  to be any set and  $P : \Omega \rightarrow [0, 1]$ , 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  $\Omega$  finite,  $p(\omega) = \frac{1}{|\Omega|} \Rightarrow$  Laplace model
- Price ? How to choose  $p : \Omega \rightarrow [0, 1] \rightarrow$  collect data, do statistics
- keeps many nice properties

- For countable sets, this is equivalent to the standard model.
- For uncountable  $\Omega$ ?
- 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 in  $[0, 1]$  or  $[0, 1]^n$ ?

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

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

### Definition 3 (Geometric probability)

Take  $f : \mathbb{R} \rightarrow (0, \infty)$  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 ok  
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 :

- $\Omega$  is some set
- $\mathcal{F} \subset P(\Omega)$  called a  $\sigma$ -algebra
  - $\emptyset \in \mathcal{F}$
  - $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$
  - $F_1, F_2, \dots \in \mathcal{F}$ , then  $\bigcup_{i \geq 1} F_i \in \mathcal{F}$  each  $F$  is called a measurable set.
- $\mu : \mathcal{F} \rightarrow [0, \infty)$  called the measure
  - $\mu(\emptyset) = 0$

— If  $F_1, \dots$ , are disjoint sets of the  $\sigma$ -algebra, then

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

— 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

- $\Omega$  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

- $\Omega \in \mathcal{F}$
- $F_1, F_2 \in \mathcal{F} \Rightarrow F_1 \setminus F_2 \in \mathcal{F}$
- $F_1, \dots \in \mathcal{F} \Rightarrow \bigcap F_i \in \mathcal{F}$
- $F_1, F_2, \dots \in \mathcal{F} \Rightarrow \bigcap_{i \geq 1} F_i \in \mathcal{F}$

Let us compare this definition with the prior ones

- $\Omega$  finite set,  $\mathcal{F} = \mathcal{P}(\Omega)$ ,  $\mathbb{P}(F) = \frac{|F|}{|\Omega|}$  this is a probability space and a laplace model.
- For  $\Omega$  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

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

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

$$\mu(F_n) \rightarrow \mu\left(\bigcup F_i\right)$$

—  $F_1, F_2, \dots, \mathcal{F}$

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

In addition, in probability spaces

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

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

## 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.

$f : \Omega_1 \rightarrow \Omega_2$  is called measurable if for every  $F \in \mathcal{F}_2$ ,  $f^{-1}(F) \in \mathcal{F}_1$

A measurable function  $f : (\Omega_1, \mathcal{F}_1) \rightarrow (\Omega_2, \mathcal{F}_2)$  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 :  $\Omega$  countable

— Continuous probability spaces :  $\Omega$  uncountable.

### 3.1 Discrete probability spaces

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

### Proposition 5

If  $(\Omega, \mathcal{F}, \mathbb{P})$  is a discrete probability space,  $\exists \Omega_2$  countable,  $\mathbb{P}_2 : \mathcal{P}(\Omega_2) \rightarrow [0, 1]$  s.t.  $(\Omega_2, \mathcal{P}(\Omega_2), \mathbb{P}_2)$  is a probability space and  $\exists f : (\Omega_1, \mathcal{F}_1) \rightarrow (\Omega_2, \mathcal{F}_2)$  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  $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}))$

**Preuve**

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

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

□

### 3.1.1 Symmetric simple random walk

A simple walk of length  $n$  s.t.  $|s_n - s_{n-1}| = 1$ .

Let  $\Omega$  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 its 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  $\Omega$  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  $\{\dots, T^{-2}x, T^{-1}x, x, Tx, \dots\}$ .*

*Note that either  $S_x = S_y$  or  $S_x \cap S_y = \emptyset$ .*

## Lecture 3: Measurable maps

Wed 06 Oct

### 3.3 Borel $\sigma$ -algebra

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

#### Definition 7 (Borel sigma-algebra)

*On  $(X, \tau)$  the borel  $\sigma$ -algebra  $\mathcal{F}_\tau$  is the smallest  $\sigma$ -algebra containing  $\tau$ .*

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

### Two nice properties

- Continuous functions on a Borel  $\sigma$ -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  $\sigma$ -algebra  $\mathcal{F}_E$  is generated by  $(a_1, b_1) \times \dots \times (a_n, b_n)$ .  
 $\mathcal{F}_E$  is the smallest  $\sigma$ -algebra containing open intervals.  $\square$

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

#### Theorème 8 (Existence of Lebesgue-measure)

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

#### Theorème 9 (Uniform Measure)

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

Both  $\lambda$  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  $f_r : x \rightarrow x + \tau, \tau \in \mathbb{R}^n$ .

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

#### Definition 8 (Cumulative distribution function)

$F : \mathbb{R} \rightarrow [0, 1]$  is called a c.d.f if

- $F$  is non-decreasing
- $F(x_n) \rightarrow 0$  then  $x_n \rightarrow -\infty$
- $F(x_n) \rightarrow 1$  if  $x_n \rightarrow 1$
- $F$  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.

- $x < y \quad F(x) = \mathbb{P}((-\infty, x)) \leq \mathbb{P}((-\infty, y)) = F(y)$
- $x_n \rightarrow -\infty \quad F(x_n) = \mathbb{P}((-\infty, x_n)) \rightarrow \mathbb{P}(\bigcap_n (-\infty, x_n)) = 0$
- $x_n \rightarrow \infty \Rightarrow F(x_n) \rightarrow 1$  is similar
- Also for right continuous  $x_n \rightarrow x$ , 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  $f : (0, 1) \rightarrow \mathbb{R}$ , 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  $I_1, \dots, I_n$  s.t.  $\mathbb{P}(B \Delta (I_1 \cup \dots \cup I_n)) < \epsilon$

### 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 intervals, 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  $I_1, \dots, I_n$  s.t.  $\mathbb{P}(B \Delta (I_1 \cup \dots)) < \epsilon$ , but  $(B \Delta A) = B^C \Delta A^C$ , hence

$$\mathbb{P}(B^C \Delta (I_1 \cup \dots)) < \epsilon$$

3.  $H_1, \dots \in H$ , we want  $\bigcup_i H_i \in H \exists n \in \mathbb{N}$

$$\mathbb{P}((\bigcup_{i=0}^m H_i) \Delta (\bigcup_i H_i)) < \frac{\epsilon}{2}$$

$\forall i = 1, \dots, m$ , we have disjoint  $I_{i,1}, \dots, I_{i,m_i}$  s.t.

$$\mathbb{P}(H_i \Delta (I_{i,1} \cup \dots)) < \frac{\epsilon}{2m}$$

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

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

So it suffices to prove the first point.

Observe, for all intervals  $\mathbb{P}_1(I) = \mathbb{P}_2(I)$  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  $I_1, \dots, I_n$  s.t.

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

and

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

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

Now we can conclude that

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

and

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

□

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

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

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

**3.6 Probability measures on  $\mathbb{R}^n$** **Definition 9 (Joint c.d.f.)**

$F : \mathbb{R}^n \rightarrow [0, 1]$

- $F$  non-decreasing in each coordinate
- $F(x_1, \dots, x_n) \rightarrow 1$  if all  $x_i \rightarrow -\infty$
- 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  $\sigma$ -algebra
- product measure

#### 3.7.1 Product $\sigma$ -algebra

##### Definition 10 (Product algebra )

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

##### Remarque

- Projections are measurable
- In fact, product  $\sigma$ -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  $\sigma$ -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  $E = E_1 \times E_2 \times \dots \times E_n \times \Omega_{n+1} \times \dots$

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

## 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 fair 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  $x \in [0, 1]$ ,  $x = 0.x_1x_2\dots$ , we send  $f : x \rightarrow (x_1, x_2, \dots)$

**Lemme 16**

*f as defined above is measurable*

**Preuve**

*Note that*

—  $\mathcal{F}_\pi$  generated by  $F_1 \times \dots, F_n \times \{0, 1\} \times \{0, 1\}$  with  $|F_i| = 1$

—  $\mathcal{F}_E$  is generated by sets of the form  $(2^{-n}j, 2^{-n}(j+1))$ . □

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

**Proposition 17**

*There exists 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  $f$*  □

## 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  $X : \Omega \mapsto \mathbb{R}$  measurable as a map  $(\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{F}_E)$  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  $\mathbb{P}_X$   $F_X$*

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

**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.  $\mathbb{P}_X(X \in (a, b)) = F(b^-) - F(a)$

**Definition 15**

$X$  a R.V.,  $s \in \mathbb{R}$ .

If  $F(s) - F(s^-) > 0 \iff \mathbb{P}_X(X = s) > 0$ , then  $s$  is a atom of  $X$

**Lemme 21**

A R.V. can have at most countably many atoms or in other words, a c.d.f. can have at most countably many jumps.

**Definition 16**

$X$  a R.V.

If  $F_X$  increases by jumps, we call  $X$  a discrete R.V.

If  $F_X$  is cts, we call  $X$  a cts R.V.

**Proposition 22**

$X$  a R.V. Then we can write  $F(X) = aF_Y + bF_Z$  s.t.  $a + b = 1$  and  $Y$  discrete,  $Z$  cts R.V.

**Preuve**

If  $F_X$  is discrete or cts, we are done.

$\exists S = \{s_1, s_2, \dots\}$  s.t.  $F_X(s_i) - F_X(s_i^-) > 0$  iff  $s_i \in S$ .

Consider

$$\hat{F}_Y(s) = \sum 1_{\{S \geq s_i\}} (F(s_i) - F(s_i^-))$$

and

$$\hat{F}_Z(s) = F_X(s) - \hat{F}_Y(s)$$

We now show that  $\hat{F}_Z$  continuous.

Finally, define

$$F_Y(s) = \frac{\hat{F}_Y(s)}{\hat{F}_Y(\infty)}$$

and similarly

$$F_Z(s) = \frac{\hat{F}_Z(s)}{\hat{F}_Z(\infty)}$$

$\square$

## Lecture 7: Example of RV

Wed 03 Nov

### Geometric R.V.

Let  $S = \mathbb{N}$  and  $0 < p \leq 1$ .

$$\mathbb{P}(X = k) = (1 - p)^{k-1}p$$

Corresponds to first succes if success rate is  $p$ .

### Definition 17

We call a rv with support  $\mathbb{N}$  memoryless if

$$\mathbb{P}(X > k + l | X > k) = \mathbb{P}(X > l)$$

### Proposition 23

*Geo( $p$ ) is memoryless and every memoryless RV with support on  $\mathbb{N}$  is a geometric rv.*

### Preuve

$$\mathbb{P}(X > k + l | X > k) \mathbb{P}(X > k) = \mathbb{P}(X > k + l) = (1 - p)^{k+l}$$

But also  $\mathbb{P}(X > l) = (1 - p)^l$

$$\mathbb{P}(X > k + l | X > k) = (1 - p)^l$$

Now suppose  $X$  is a memoryless RV with  $\mathbb{P}(X > 1) > 0$ , then

$$\mathbb{P}(X > l + 1 | X > 1) = \frac{\mathbb{P}(X > l + 1)}{\mathbb{P}(X > 1)} = \mathbb{P}(X > l) \quad \square$$

Inductively, it follows that  $\mathbb{P}(X > l) = \mathbb{P}(X > 1)^l$

### Poisson RV

Define

$$\mathbb{P}(Poi(\lambda) = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

### Proposition 24

$$Ber(n, \frac{\lambda}{n}) \mapsto Poi(\lambda) \text{ as } n \rightarrow \infty$$

in the sense that  $\forall k \in \mathbb{N}$

$$\mathbb{P}(\text{Ber}(n, \frac{\lambda}{n}) = k) \rightarrow \mathbb{P}(\text{Poi}(\lambda) = k)$$

**Preuve**

$$\begin{aligned} \mathbb{P}(\text{Bin}(n, \frac{\lambda}{n}) = k) &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \mathbb{P}(\text{Bin}(n, \frac{\lambda}{n}) = k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &= \frac{\lambda^k}{k!} e^{-\lambda} \left(\frac{n!}{(n-k)! n^k}\right) \left(1 - \frac{\lambda}{n}\right)^{-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

## 4.1 Independence of RV

**Definition 18 (Independence of RV)**

$(X_i)_{i \geq 1}$  RV defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  are called mutually independent if  $\forall J \subset \{1, 2, \dots\}$  finite ad  $\forall E_j \in \mathcal{F}_E \forall j \in J$ .

$$\mathbb{P}\left(\bigcap_{j \in J} \{X_j \in E_j\}\right) = \prod \mathbb{P}(X_j \in E_j)$$

**Proposition 25**

$(X_i)_{i \geq 1}$  RV with laws  $\mathbb{P}_{X_i}$  then we can find a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and RV  $\tilde{X}_i$  s.t.

- $X_i \simeq \tilde{X}_i$
- $(\tilde{X}_i)$  are mutually independent.

**Preuve**

Consider the product probability space of  $(\mathbb{P}_{X_i})$  i.e.  $(\mathbb{R}^n, \mathcal{F}_\pi, \mathbb{P}_\pi)$ .

Let  $\tilde{X}_i$  be the projection on the  $i$ -th coordinate.

Are  $(\tilde{X}_i)$  independent?

$$\mathbb{P}_\pi\left(\bigcap_{j \in J} \{\tilde{X}_j \in E_j\}\right) = \mathbb{P}_\pi\left(\bigcap_i F_i\right)$$

□

With  $F_i = \mathbb{R}$  if  $i \notin J$  and  $F_i = E_i$  if  $i \in J$

## 4.2 Example of continuous random variables

We will mainly work with a subclass of continuous rv :

**Definition 19 (Random variables with density)**

We call a continuous rv  $X$  with c.d.f.  $F_x$  a r.v. with densite if  $\exists f : \mathbb{R} \rightarrow [0, \infty)$  which is integrable,  $\int_{\mathbb{R}} f = 1$  st

$$\mathbb{P}(X \leq t) = F_X(t) = \int_{-\infty}^t f_X(s) ds$$



## Lecture 8: rv with density

Wed 10 Nov

The gaussian random variable describes sums of independent errors

### Theorème 26 (Version of central limit theorem)

Let  $X_1, \dots$  be iid random variables s.t.  $\mathbb{P}(|X_I| < C) = 1$  for some  $C > 0$  and such that  $-X_i$  and  $X_i$  have the same law.

Then  $S_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}} \rightarrow \mathcal{N}(0, \sigma^2)$  in the sense that  $\mathbb{P}(S_n \in (a, b)) \rightarrow \mathbb{P}(\mathcal{N}(0, \sigma^2) \in (a, b))$

## 4.3 Transformation of random variables

### Lemme 27

If  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  continuous and r.v. on  $(\omega, \mathcal{F}, \mathbb{P})$  then  $\phi(X)$  is also r.v. on  $(\Omega, \mathcal{F}, \mathbb{P})$

### Preuve

Need to check that  $\phi \circ X$  is measurable.

- Continuous functions are measurable
- Composition of measurable functions is measurable □

### Proposition 28

Let  $X$  be a continuous random variable with density  $f_X : \mathbb{R} \rightarrow [0, \infty)$ .

Let  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  bijective, continuously differentiable with  $\phi'(x) \neq 0 \forall x \in \mathbb{R}$ .

Then  $\phi(X)$  is a r.v. with density given by

$$f_{\phi(X)}(x) = \frac{1}{\phi' \circ \phi^{-1}(X)} f_X(\phi^{-1}(x))$$

## 5 Random Vectors

### Definition 20 (Random Vector)

$(\Omega, \mathcal{F}, \mathbb{P})$   $X_1, \dots, X_n$  random variables then  $\overline{X} = (X_1, \dots, X_n)$  is called a random vector.

### Remarque

Marginal laws on their own do not describe the behavior of  $\overline{X}$ .

### Lemme 30

If  $X_1, \dots, X_n : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{F}_E)$  measurable  $\overline{X}$  is measurable from  $(\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^n, \mathcal{F}_E)$

**Preuve**

Suffices to check that  $\bar{X}^{-1}$  of each  $E = F_1 \times \dots \times F_n$  is in  $\mathcal{F}$

$$\bar{X}^{-1}(E) = \bigcap X_i^{-1}(F_i) \in \mathcal{F}$$

Hence we can define

$$\mathbb{P}_{\bar{X}}(E) := \mathbb{P}(\bar{X}^{-1}(E)) \forall E \in \mathcal{F}_E$$

Which is a probability law on  $(\mathbb{R}^n, \mathcal{F}_E)$  called the joint law of  $\bar{X}$  □

**Proposition 31**

The joint law of a random vector  $\bar{X}$  is uniquely characterised by the joint cdf

**Preuve**

Restatement of probability measure on  $\mathbb{R}^n$  are in correspondence with joint cdf □

**Lemme 32**

$X_1, \dots, X_n$  random variables  $(\Omega, \mathcal{F})$  are independent if and only if

$$F_{\bar{X}}(x_1, \dots, x_n) = \prod_i F_i(x_i)$$

**Transformations of random vectors****Proposition 33**

$\bar{X}$  is a  $\mathbb{R}^n$  valued random vector and  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$  continuous then  $\phi(\bar{X})$  is a  $\mathbb{R}^m$  valued random vector with values in  $\mathbb{R}^m$

**Corollaire 34**

Let  $X_1, \dots, X_n$  be random variables on  $\Omega$  then  $\sum a_i X_i$  is a random variable

**Definition 21 (Random vectors with density)**

Let  $\bar{X} = (X_1, \dots, X_n)$  random vector then

$$f_{\bar{X}} : \mathbb{R}^n \rightarrow [0, \infty)$$

Riemann integrable with  $\int_{\mathbb{R}^n} f_{\bar{X}}(y) dy = 1$  if  $\forall [a_0, b_1) \times \dots [a_n, b_n) = B$  we have

$$\mathbb{P}(\{X_1 \in [a_0, b_1)\} \cap \dots) = \int_B f_{\bar{X}}(y) dy$$

**Gaussian vectors**

Let  $\bar{\mu} \in \mathbb{R}^n$  and  $C$  positive definite  $n \times n$  matrix called covariance matrix then the density

$$f_{\bar{\mu}, C}(\bar{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \det C^{\frac{1}{2}}} \exp(-2(\bar{x} - \bar{\mu})^T C^{-1}(\bar{x} - \bar{\mu}))$$