

# MIPS: Mesure, Intégration

Pr. Lacroix

SEATECH  
Université de Toulon

Promo 2018

## 1 Introduction

## 2 Mesure

- Principe fondamental
- Algèbre
- Tribu
- Mesure
- Lebesgue
- Mesurable
- Image
- Produit
- Plus de deux termes

## 3 Intégrales

- Construction
- Calcul 1D
- Presque partout
- Calcul  $>1D$  : Fubini

- Dérivation dans  $\mathbb{R}^n$

## 4 Convergences

- CVPS
- Beppo-Lévi & Fatou
- TCDL
- Continuité de l'intégrale
- Dérivation de l'intégrale
- Convergence faible

## 5 Inégalités

- Hölder
- Minkowski
- Inégalité générique
- Jensen

## 6 Conditionnement

- Absolue continuité
- Radon-Nikodym

- Mesure  $\implies$  Intégrale = Valeur moyenne d'une fonction
- Mesure  $\implies$  Probabilité
- Probabilité  $\implies$  Modèles probabilistes
- Statistiques  $\implies$  Inférence, Qualité, Aide à la décision, Risque...

- La mesure est donc à la base de tout ce qui suit.
- Émergence de la théorie moderne de la mesure au 20<sup>ième</sup> siècle.

On note  $\Omega$  un ensemble non vide. On note  $\mathcal{P}(\Omega)$  l'ensemble des parties de  $\Omega$ .

La définition ou construction d'une mesure  $\mu$  sur  $\Omega$  va se dérouler en deux étapes :

- sélectionner les sous-ensembles de  $\Omega$  auxquels on va assigner un poids ou mesure ;
- définir pour chacun de ceux qui ont une mesure la valeur exacte de celle-ci de sorte qu'elle soit cohérente au regard du principe fondamental suivant :

La mesure du tout est égale à la somme des mesures des parties.

### Exemple

Longueur, poids, volume, surface, richesse, quantités...

## Définition

Une **algèbre** est un sous-ensemble  $\mathcal{A} \subset \mathcal{P}(\Omega)$  qui est stable au sens suivant :

- $\emptyset \in \mathcal{A}$  ;
- $A_1, \dots, A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^n A_i \in \mathcal{A}$  ;
- $A \in \mathcal{A} \Rightarrow A^c := \Omega \setminus A \in \mathcal{A}$ .

## Exemple

L'ensemble des unions finies d'intervalles de  $\mathbb{R}$ . L'ensemble des parties d'un ensemble fini.

## Définition

Une **tribu** ou  **$\sigma$ -algèbre** est un sous-ensemble  $\mathcal{T} \subset \mathcal{P}(\Omega)$  qui est stable au sens suivant :

- $\emptyset \in \mathcal{T}$  ;
- $A_1, \dots, A_n, \dots \in \mathcal{T} \Rightarrow \cup_{i \geq 1} A_i \in \mathcal{T}$  ;
- $A \in \mathcal{T} \Rightarrow A^c := \Omega \setminus A \in \mathcal{T}$ .

Les éléments de la tribu  $\mathcal{T}$  sont appelés des évènements, parfois qualifiés d'observables ou de mesurables.

## Exemple

$\mathcal{P}(\Omega)$ .  $\{\emptyset, \Omega\}$ .



## Définition (Mesure et Probabilité)

On appelle **espace mesurable** un couple  $(\Omega, \mathcal{T})$  constitué d'un ensemble et d'une tribu de ses parties.

Une **mesure** (ou mesure positive) est alors une application  $\mu : \mathcal{T} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$  telle que

- $\mu(\emptyset) = 0$  ;
- $A_i \in \mathcal{T}$  et  $A_i \cap A_j = \emptyset \implies \mu(\cup_i A_i) = \sum_i \mu(A_i)$ .

Une **probabilité** n'est autre qu'une mesure de masse totale 1, i.e. telle que  $\mu(\Omega) = 1$ .

## Exemple

Comment faire pour un dé à six faces ??? Sept faces ???  $\mathbb{N}$  ???

Construction de la **Mesure de Lebesgue**  $\lambda_n$  sur  $(\Omega, \mathcal{T}) = (\mathbb{R}^n, \mathcal{B}_n)$  :

- pavé : sous-ensemble du type  $I_1 \times \dots \times I_n$  avec  $I_j \in \mathcal{B}_1$  ;
- $\mathcal{B}_1$  est la plus petite tribu contenant les intervalles de  $\mathbb{R}$  ;
- $\mathcal{B}_n$  est la plus petite tribu contenant les pavés ;
- $\lambda_n(I_1 \times \dots \times I_n) = \times_{j=1}^n \lambda_1(I_j)$  ;
- $\lambda_1(\text{Intervalle}) = \mathbf{Longueur}(\text{Intervalle})$ .

$\lambda_1$  mesure les longueurs,  $\lambda_2$  les aires,  $\lambda_3$  les volumes .

C'est à Lebesgue que l'on doit l'existence de ces mesures, les preuves modernes utilisent les travaux de Kolmogorov et de Carathéodory.

## Définition

$(\Omega_1, \mathcal{T}_1)$  et  $(\Omega_2, \mathcal{T}_2)$  deux espaces mesurables,  $f : \Omega_1 \rightarrow \Omega_2$  une application.

- $f^{-1}(\mathcal{T}_2) := \{f^{-1}(T) : T \in \mathcal{T}_2\}$  est une tribu de  $\Omega_1$  ;
- $f$  est dite  $(\mathcal{T}_1, \mathcal{T}_2)$ -**mesurable** si  $f^{-1}(\mathcal{T}_2) \subset \mathcal{T}_1$ .

Des exemples : toute application ayant un ensemble fini ou dénombrable de discontinuités  $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$  est  $(\mathcal{B}_p, \mathcal{B}_q)$ -mesurable.

Difficile de construire  $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$  non  $(\mathcal{B}_p, \mathcal{B}_q)$ -mesurable.

## Définition

Soit  $f : \Omega_1 \rightarrow \Omega_2$   $(\mathcal{T}_1, \mathcal{T}_2)$ -mesurable. Soit  $\mu$  une mesure positive sur  $(\Omega_1, \mathcal{T}_1)$ . On appelle alors **mesure image de  $\mu$  par  $f$** , et on note  $\mu_f$ , la mesure définie par

$$\mu_f(T_2) = \mu(f^{-1}(T_2)), \quad T_2 \in \mathcal{T}_2.$$

Soit  $\mu_1$  (reps.  $\mu_2$ ) une mesure sur  $(\Omega_1, \mathcal{T}_1)$  (resp.  $(\Omega_2, \mathcal{T}_2)$ ).

### Définition

- On appelle *tribu produit* de  $\mathcal{T}_1$  par  $\mathcal{T}_2$  la plus petite tribu contenant tous les “pavés” de la forme  $T_1 \times T_2$ ,  $T_i \in \mathcal{T}_i$ . On la note  $\mathcal{T}_1 \otimes \mathcal{T}_2$ .
- On appelle *mesure produit* de  $\mu_1$  par  $\mu_2$  la seule mesure définie sur  $\mathcal{T}_1 \otimes \mathcal{T}_2$ , notée  $\mu_1 \otimes \mu_2$ , et telle que

$$\mu_1 \otimes \mu_2(T_1 \times T_2) = \mu_1(T_1)\mu_2(T_2),$$

en adoptant la convention  $0 \times \infty = 0$ .

## Définition

*Par associativité, on peut généraliser à plus de deux termes :*

$$\bigotimes_{i=1}^n \mu_i \left( \prod_{i=1}^n T_i \right) := \prod_{i=1}^n \mu_i(T_i).$$

Par construction de la mesure de Lebesgue, on a les relations :

$$\lambda_n = \lambda_1^{\otimes n}, \quad \lambda_p \otimes \lambda_q = \lambda_{p+q}.$$

Ces propriétés vont jouer un rôle fondamental dans la définition et le calcul des intégrales multiples.

Soit  $(\Omega, \mathcal{T}, \mu)$  un espace mesuré.

- Soit  $A \in \mathcal{T}$ , alors  $1_A : \Omega \rightarrow \mathbb{R}$  définie par  $1_A(\omega) = 1$  si  $\omega \in A$  et sinon  $1_A(\omega) = 0$  est appelée **fonction indicatrice de  $A$** . On appelle **intégrale** de  $1_A$  contre  $\mu$  la quantité notée  $\int_{\Omega} 1_A d\mu$  et définie par

$$\int_{\Omega} 1_A d\mu := \mu(A).$$

- $g : \Omega \rightarrow \mathbb{R}$  est dite **simple positive** si elle s'écrit sous la forme

$$g = \sum_{i=1}^n \alpha_i 1_{A_i} \text{ pour certains } A_i \in \mathcal{T} \text{ et } \alpha_i \geq 0.$$

Son **intégrale** contre  $\mu$ , notée  $\int_{\Omega} g d\mu$ , est définie par

$$\int_{\Omega} g d\mu := \sum_{i=1}^n \alpha_i \mu(A_i).$$

Soit  $f : \Omega \rightarrow \mathbb{R}^+$   $(\mathcal{T}, \mathcal{B}_1)$ -mesurable (on dira simplement “mesurable”).

Alors il existe une suite simplement croissante de fonctions simples  $g_i$  telle que

$$\forall \omega \in \Omega, \lim_{i \rightarrow \infty} \uparrow g_i(\omega) = f(\omega).$$

Alors l'**intégrale de  $f$**  contre  $\mu$  est définie par

$$\int_{\Omega} f d\mu := \lim_{i \rightarrow \infty} \uparrow \int_{\Omega} g_i d\mu.$$

On dit que  $f$  est **intégrable** si  $\int_{\Omega} f d\mu < +\infty$ .



Soit  $f : \Omega \rightarrow \mathbb{R}$  mesurable mais non nécessairement positive. On appelle

- partie positive de  $f$  la fonction mesurable  $f^+ = \max(f, 0)$ ;
  - partie négative de  $f$  la fonction mesurable  $f^- = -\min(f, 0)$ .
- Alors

$$f = f^+ - f^-, |f| = f^+ + f^-.$$

**$f$  intégrable contre  $\mu$  ssi  $|f|$  intégrable contre  $\mu$**

**ssi  $\int_{\Omega} |f| d\mu < +\infty$**

$$\int_{\Omega} f d\mu = \int_{\Omega} f^+ d\mu - \int_{\Omega} f^- d\mu.$$

L'intégrale est linéaire et vérifie les propriétés habituelles de l'intégrale que vous connaissiez. Elle est également monotone, i.e.  $f \leq g \Rightarrow \int f d\mu \leq \int g d\mu$ .

L'intégrale contre la mesure de Lebesgue  $\lambda_1$  coïncide en général avec l'intégrale "classique",

$$\int_{\mathbb{R}} f d\lambda_1 = \int_{-\infty}^{+\infty} f(x) dx.$$

donc changements de variable, primitives, et autres méthodes d'intégration restent les mêmes.

Attention aux intégrales dites semi-convergentes qui peuvent être définies au sens de Riemann généralisé, mais pas au sens de Lebesgue.

Si  $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$  alors on prend le plus souvent  $\mathcal{T} = \mathcal{P}(\Omega)$ , et  $\mu$  est définie par la donnée de chaque  $\mu(\{\omega_i\})$ . Alors toute application  $f : \Omega \rightarrow \mathbb{R}$  est mesurable et

$$\int_{\Omega} f d\mu = \sum_i f(\omega_i) \mu(\{\omega_i\}).$$

## Définition

- Soit  $(\Omega, \mathcal{T}, \mu)$  un espace mesuré. Une propriété  $P$  est dite vérifiée ou vraie **presque partout**, ou **presque sûrement**, ssi il existe  $\Omega_1 \in \mathcal{T}$  tel que  $\mu(\Omega \setminus \Omega_1) = 0$ , et que  $P$  soit vraie pour chaque  $\omega \in \Omega_1$ .
- Un ensemble est dit **négligeable** s'il est inclus dans un ensemble mesurable ( $\in \mathcal{T}$ ) qui lui-même est de mesure nulle.

Deux fonctions mesurables définies sur  $\omega$  sont dites égales presque partout si l'ensemble des points où elles diffèrent est négligeable.

On note  $f = g$  p.p. . Alors, leurs intégrales, si elles existent, sont égales :

$$f = g \text{ p.p.} \Rightarrow \int f d\mu = \int g d\mu .$$

## Théorème (de Fubini - Tonelli)

Soient  $(\Omega_i, \mathcal{T}_i, \mu_i)$ ,  $i = 1, 2$ , deux espaces mesurés positifs, et  $(\Omega_1 \times \Omega_2, \mathcal{T}_1 \otimes \mathcal{T}_2, \mu_1 \otimes \mu_2)$  leur produit.

- Si  $f : \Omega_1 \times \Omega_2$  est mesurable, alors p.s.,  $f(\omega_1, \cdot)$  et  $f(\cdot, \omega_2)$  le sont.
- Si  $f$  est p.s. positive, alors

$$\begin{aligned} & \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \otimes \mu_2) \\ = & \int_{\Omega_1} \left( \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right) d\mu_1(\omega_1) \\ = & \int_{\Omega_2} \left( \int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2). \end{aligned}$$

## Théorème (de Fubini)

- $f$  est  $\mu_1 \otimes \mu_2$ -intégrable ssi  $\omega_1 \mapsto \int_{\Omega_2} |f|(\omega_1, \omega_2) d\mu_2(\omega_2)$  est  $\mu_1$ -intégrable ssi  $\omega_2 \mapsto \int_{\Omega_1} |f|(\omega_1, \omega_2) d\mu_1(\omega_1)$  est  $\mu_2$ -intégrable.
- Si  $f$  est  $\mu_1 \otimes \mu_2$ -intégrable, alors

$$\begin{aligned} & \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \otimes \mu_2) \\ = & \int_{\Omega_1} \left( \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right) d\mu_1(\omega_1) \\ = & \int_{\Omega_2} \left( \int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2). \end{aligned}$$

Ceci se généralise de la façon la plus intuitive au cas de plus de deux variables.

- Soient  $U$  et  $V$  deux sous-ensembles de  $\mathbb{R}^n$  et  $\mathbb{R}^p$ , et soit  $\phi : U \rightarrow V$  une application. On dit que  $\phi$  est **dérivable** en  $u \in U$  ssi il existe une application linéaire continue  $D\phi_u : \mathbb{R}^n \rightarrow \mathbb{R}^p$  telle que

$$\phi(u + h) = \phi(u) + D\phi_u(h) + o(h).$$

- Une telle application linéaire, si elle existe, est unique.
- On dit que  $\phi$  est dérivable sur  $U$  si elle l'est en chaque point  $u \in U$ . On dit qu'elle est  $\mathcal{C}^1$  si  $u \mapsto D\phi_u$  est continue.
- Les applications  $\phi \mapsto D\phi_u$  et  $\phi \mapsto D\phi$  sont linéaires.
- Si  $W \subset \mathbb{R}^q$  et  $\psi : V \rightarrow W$ , si  $\phi$  est dérivable en  $u$  et  $\psi$  en  $\phi(u)$ , alors  $\psi \circ \phi$  l'est en  $u$  et

$$D(\psi \circ \phi)_u = D\psi_{\phi(u)} \circ D\phi_u.$$

- Si  $\phi$  et  $\psi$  sont dérivables, alors  $\psi \circ \phi$  l'est.
- Si elles sont  $\mathcal{C}^1$ ,  $\psi \circ \phi$  aussi.
- $\phi = (\phi_1, \dots, \phi_p)$  : on appelle les  $\phi_i$  les **applications composantes** de  $\phi$ .
- Si  $\phi$  est dérivable en  $u$ , alors **la matrice de  $D\phi_u$  relativement aux bases canoniques, appelée Jacobienne de  $\phi$  en  $u$ , notée  $Jac(\phi, u)$ , est donnée par**

$$Jac(\phi, u) = \left( \frac{\partial \phi_i}{\partial x_j} \right)_{\substack{1 \leq i \leq p, \\ 1 \leq j \leq n}}$$

## Théorème (Changement de variable intégrales multiples)

Soit  $f : V \rightarrow \mathbb{R}$  mesurable, et  $\phi : U \rightarrow V$  bijective (ici  $n = p$ ) de classe  $\mathcal{C}^1$ . On appelle **Jacobien** de  $\phi$  en  $u$  la quantité notée  $J(\phi, u)$  et définie par

$$J(\phi, u) = |\det \text{Jac}(\phi, u)|.$$

Alors  $f1_V$  est  $\lambda_p$ -intégrable ssi  $f \circ \phi 1_U J(\phi, \cdot)$  l'est, et alors

$$\int_U f \circ \phi J(\phi, \cdot) d\lambda_p = \int_V f d\lambda_p.$$

### Exemple

Calculer l'aire d'un disque, le volume d'une sphère, d'une ellipse....



## Définition

- Une suite  $f_n : \Omega \rightarrow \mathbb{R}^p$  est dite converger p.s. vers  $f : \Omega \rightarrow \mathbb{R}^p$  (on note  $f_n \rightarrow_{p.s.} f$ ) ssi  $f_n$  et  $f$  sont mesurables et
$$\mu(\{f_n \not\rightarrow f\}) = 0.$$
- Si  $p = 1$ , on appelle limite supérieure (reps. inférieure) des  $f_n$  la quantité notée  $\limsup f_n$  (resp.  $\liminf f_n$ ) définie par
$$\limsup f_n = \inf_n (\sup_{p \geq n} f_p) \text{ (resp. } \liminf f_n = \sup_n (\inf_{p \geq n} f_p)).$$
Les deux sont mesurables si les  $f_n$  le sont.
- $f_n \rightarrow_{p.s.} f$  ssi  $f = \liminf f_n = \limsup f_n$ .

On retrouvera la CVPS dans la loi des grands nombres, un des grands théorèmes de convergence de la statistique.

Intervertir limite et intégrale.

### Théorème (Beppo-Lévi)

*Si  $0 \leq f_1 \leq \dots \leq f_n \leq \dots$ , alors intégrables ou pas,*

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n d\mu = \int_{\Omega} (\lim_{n \rightarrow \infty} f_n) d\mu$$

### Lemme (de Fatou)

*Si les  $f_n$  sont positives,*

$$\int_{\Omega} \liminf f_n d\mu \leq \liminf \int_{\Omega} f_n d\mu$$

Intervertir toujours.

### Théorème (de Convergence Dominée de Lebesgue - TCDL)

*Si  $f_n \rightarrow_{p.s.} f$ , et s'il existe  $g$  intégrable telle que  $|f_n| \leq g$ , alors*

- *$f$  est intégrable.*
- $\int_{\Omega} |f_n - f| d\mu \rightarrow 0$ .
- $\lim_n \int_{\Omega} f_n d\mu = \int_{\Omega} \lim_n f_n d\mu$ .

## Théorème (de Continuité de l'intégrale paramétrée)

Soit  $I \subset \mathbb{R}$  un intervalle,  $t_0 \in I$ , et  $f : \Omega \times I \rightarrow \mathbb{R}$  telle que

- pour tout  $t$  voisin de  $t_0$ ,  $f(\cdot, t)$  est intégrable ;
- $\mu$ -p.s.,  $f(\omega, \cdot)$  est continue en  $t_0$  ;
- il existe  $h$  intégrable telle que  $\mu$ -p.s., pour tout  $t$  voisin de  $t_0$ ,  
 $|f(\omega, t)| \leq h(\omega)$  ;

Alors  $t \mapsto \int_{\Omega} f(\omega, t) d\mu(\omega)$  est continue en  $t_0$ , i.e.

$$\lim_{t \rightarrow t_0} \int_{\Omega} f(\omega, t) d\mu(\omega) = \int_{\Omega} f(\omega, t_0) d\mu(\omega) .$$

## Théorème (de Dérivation sous le signe somme)

Soit  $I \subset \mathbb{R}$  un intervalle,  $t_0 \in I$ , et  $f : \Omega \times I \rightarrow \mathbb{R}$  telle que

- pour tout  $t$  voisin de  $t_0$ ,  $f(\cdot, t)$  est intégrable ;
- $\mu$ -p.s.,  $f(\omega, \cdot)$  est dérivable au voisinage de  $t_0$  ;
- il existe  $h$  intégrable telle que  $\mu$ -p.s., pour tout  $t$  voisin de  $t_0$ ,  
 $|\frac{\partial f}{\partial t}(\omega, t)| \leq h(\omega)$  ;

Alors  $t \mapsto \int_{\Omega} f(\omega, t) d\mu(\omega)$  est dérivable et en  $t_0$ ,

$$\frac{\partial}{\partial t} \left( \int_{\Omega} f(\omega, t) d\mu(\omega) \right) = \int_{\Omega} \left( \frac{\partial f}{\partial t} \right) (\omega, t) d\mu(\omega) .$$

Une notion essentielle pour le théorème limite centrale de la statistique :

### Définition

Sur  $(\mathbb{R}^p, \mathcal{B}_p)$ , on se donne une suite de mesures positives  $(\mu_n)_{n \geq 1}$ , et une mesure positive  $\mu$ . On suppose que les  $\mu_n$  et  $\mu$  assignent des mesures finies aux parties bornées mesurables de  $\mathbb{R}^p$ .

On dit que  $(\mu_n)$  converge faiblement vers  $\mu$  ssi quelle que soit  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  continue à support compact,

$$\int_{\mathbb{R}^p} f d\mu_n \rightarrow_{n \rightarrow \infty} \int_{\mathbb{R}^p} f d\mu.$$

On note  $\mu_n \rightarrow_{w*} \mu$ .

Soient  $p, q \geq 1$ . On dit que  $p$  et  $q$  sont conjugués si  $\frac{1}{p} + \frac{1}{q} = 1$ .

### Théorème (Hölder)

*Soient  $f, g : \Omega \rightarrow \mathbb{R}$  telles que  $|f|^p$  et  $|g|^q$  soient intégrables, avec  $p$  et  $q$  conjugués. Alors  $fg$  est intégrable et*

$$\int |fg| d\mu \leq \left( \int |f|^p d\mu \right)^{\frac{1}{p}} \left( \int |g|^q d\mu \right)^{\frac{1}{q}}.$$

### Corollaire (Inégalité de Cauchy-Schwarz)

$$\sum_{k=1}^n |x_k y_k| \leq \sqrt{\sum_{k=1}^n x_k^2} \sqrt{\sum_{k=1}^n y_k^2}.$$

## Théorème (Minkowski)

Soit  $p \geq 1$  et  $f, g$  telles que  $|f|^p$  et  $|g|^p$  soient intégrables. Alors  $|f + g|^p$  est intégrable et

$$\left( \int |f + g|^p d\mu \right)^{\frac{1}{p}} \leq \left( \int |f|^p d\mu \right)^{\frac{1}{p}} + \left( \int |g|^p d\mu \right)^{\frac{1}{p}}.$$

On note souvent  $\|f\|_p := \left( \int |f|^p d\mu \right)^{\frac{1}{p}}$ , et on l'appelle la **norme  $p$  de  $f$** . C'est presque une norme au sens strict. Il s'agit d'un outil fondamental de l'analyse fonctionnelle.

On dit que  $(f_n)$  **converge vers  $f$  en norme  $p$** , et on note

$$f_n \xrightarrow{\|\cdot\|_p} f, \text{ si } \|f_n - f\|_p \rightarrow 0.$$



### Lemme

Soit  $f : \Omega \rightarrow \mathbb{R}$  mesurable et  $\psi : \mathbb{R} \rightarrow \mathbb{R}^+$  mesurable positive. Soit  $A \in \mathcal{B}_1$  et soit

$$i_A = \inf\{\psi(x) : x \in A\}.$$

Alors  $i_A \mu(f^{-1}(A)) \leq \int_{\Omega} \psi \circ f d\mu$ .

### Corollaire

- Markov :  $r \geq 1$  et  $\alpha > 0$  alors  $\mu\{|f| \geq \alpha\} \leq \frac{\int |f|^r d\mu}{\alpha^r}$  ;
- Bienaymé - Tchebychev : si  $f$  et  $f^2$  sont intégrables alors  $\mu\{|f - \int f d\mu| \geq \alpha\} \leq \frac{\int (f - \int f d\mu)^2 d\mu}{\alpha^2}$ .

## Théorème (Jensen)

Soit  $f : \Omega \rightarrow ]a, b[$  intégrable et  $\psi : ]a, b[ \rightarrow \mathbb{R}$  convexe telle que  $\psi \circ f$  soit intégrable. Si  $\mu$  est une probabilité, alors

$$\psi\left(\int f d\mu\right) \leq \int \psi \circ f d\mu.$$

## Corollaire

On suppose toujours  $\mu(\Omega) = 1$ . Alors

- Lyapunov : si  $0 < s \leq t$ , alors  $\|f\|_s \leq \|f\|_t$  ;
- Cantelli : si  $f$  est d'intégrale nulle et  $f^2$  intégrable, alors si  $a > 0$ ,  
$$\mu\{f > a\} \leq \frac{\int f^2 d\mu}{\int f^2 d\mu + a^2}.$$

### Définition (Absolue Continuité)

Une mesure  $\nu$  sur  $(\Omega, \mathcal{T})$  est dite **absolument continue** **relativement à une mesure**  $\mu$  sur  $(\Omega, \mathcal{T})$  (on note  $\nu \ll \mu$ ) ssi

$$\forall A \in \mathcal{T}, \mu(A) = 0 \Rightarrow \nu(A) = 0.$$

On dit que  $\mu$  est  **$\sigma$ -finie** si il existe une suite croissante  $\Omega_n \in \mathcal{T}$  telle que  $\mu(\Omega_n) < +\infty$  et  $\Omega = \cup_n \Omega_n$ .

## Théorème (Radon-Nikodym)

*Si  $\mu$  est  $\sigma$ -finie et si  $\nu \ll \mu$ , alors il existe  $h : \Omega \rightarrow \mathbb{R}$  mesurable positive telle que*

$$\forall A \in \mathcal{T}, \nu(A) = \int_A h d\mu.$$

*On appelle  $h$  la **dérivée de Radon-Nikodym de  $\nu$  par rapport à  $\mu$** , et on note*

$$h = \frac{d\nu}{d\mu}.$$

*Si  $g : \Omega \rightarrow \mathbb{R}$  est mesurable positive, alors*

$$\int g d\nu = \int g \frac{d\nu}{d\mu} d\mu.$$