Intégration et probabilités

Introduction

Références :

— Billingsley, *Probabiliy and measure*

— Kolmogorov & Fomin, tome 2

Motivations :

— Définir la longueur d'une partie de $\mathbb{R}$

— Définir l'aire d'une partie de $\mathbb{R}^2$

— Définir $\int f \mathrm{d}x$ pour $f : \mathbb{R}^d \to \mathbb{R}$

— Définir, préciser la notion mathématique décrivant une suite infinie de jets de dés

Par exemple :

— Si $f : \mathbb{R} \to \mathbb{R}$, on peut définir $\int f$ comme l'aire algébrique définie par le graphe de f. Ainsi, définir une aire permet de définir une intégrale

— De même, $\lambda(A) = \mathbb{1}_A$ avec $\mathbb{1}_A(x) = 1$ ssi $x \in A$. Donc définir une intégrale revient à définir une mesure.

— Tirer un nombre au hasard dans $[0,1]$, cela revient à tirer au hasard la suite de ses décimales au D10, car on mesure une partie de $\{0,1,\dots 9\}^{\mathbb{N}}$

On se demande alors comment définir la surface d'une partie du plan.

Méthode 1 : à la Riemann. On approxime avec un quadrillage. On compte le nombre de carrés qui intersectent l'ensemble considéré, puis on conclut en passant à la limite quand le côté du quadrillage tend vers 0.

Méthode 2 : on pose $\lambda(A) := \inf\limits_{(R_i)} \sum_{i=1}^{\infty} \lambda(R_i)$ où R_i est une suite de rectangles recouvrant A.

À noter : les deux méthodes ont des cas pathologiques différents.

Ensembles dénombrables

Définition : Un ensemble est dénombrable ssi il est en bijection avec $\mathbb{N}$

Propriété : Toute partie d'un ensemble dénombrable est au plus dénombrable

Démonstration : On pose $x : \mathbb{N} \to X, Y \subset X$. Si Y n'est pas fini :

$$i_1 = \min\{i \in \mathbb{N}, x_i \in Y\}$$
$$\dots$$
$$i_n = \min\{i \in \mathbb{N}, x_i \in Y \setminus \{x_1, \dots, x_{n-1}\}\}$$

Ainsi, $k \mapsto x_{n_k}$ est une bijection de $\mathbb{N}$ vers Y.

□

Propriété : L'image d'une suite est au plus dénombrable.

Démonstration : On note $x : \mathbb{N} \to X$ une suite. On crée de manière analogue une sous-suite injective de x de même image que x (sauf si $f(x(\mathbb{N}))$ est fini).

□

Propriété : $\mathbb{N} \times \mathbb{N}$ est dénombrable.

Démonstration : $(n_1, n_2) \mapsto 2^{n_1}(2n_2 + 1) - 1$ est une bijection $\mathbb{N}^2 \to \mathbb{N}$.

□

Propriété : Une réunion au plus dénombrables d'ensembles au plus dénombrable est au plus dénombrable.

Démonstration : On traite le cas "union dénombrable d'ensembles dénombrables".

Soit A_i des parties dénombrables d'un ensemble X. Pour tout i, il existe $b_i : \mathbb{N} \to A_i$ bijection. (nb : ceci requiert en fait l'axiome du choix dénombrable)

Alors $\begin{array}{rcl} \mathbb{N}^2 & \to & \bigcup\limits_i A_i \\ (i,j) & \mapsto & b_i(j) \end{array}$ est surjective.

Donc $\bigcup\limits_i A_i$ est au plus dénombrable.

Or $\bigcup\limits_i A_i \supset A_i$.

Donc $\bigcup\limits_i A_i$ est dénombrable.

□

Propriété : Si X est dénombrable, $\mathcal{P}(X)$ ne l'est pas.
Plus généralement, quel que soit X, X et $\mathcal{P}(X)$ ne sont jamais en bijection (théorème de Cantor).

Démonstration : Supposons qu'il existe $x : \begin{array}{rcl} X & \to & \mathcal{P}(X) \\ x & \mapsto & A_x \end{array}$ une bijection.

Considérons $B := \{x, x \notin A_x\}$. Comme x est une bijection, il existe $y \in X$ tel que $B = A_y$.

Question : a-t-on $y \in B$. On arrive à un paradoxe type Russel.

□

Exercice :
— $\{0,1\}^{\mathbb{N}}$ est non dénombrable.
— $\mathbb{R}$ est non dénombrable.

$\limsup$ et $\liminf$

Définition :
Soit $(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ (plus généralement $\in \bar{\mathbb{R}}^{\mathbb{N}}$). Alors $s_n := \sup\limits_{k\geq n} x_k$.
s_n est décroissante (donc a une limite dans $\bar{\mathbb{R}}$).
Alors $\lim s_n =: \limsup x_n = \inf s_n$.
De même pour $\liminf x_n$.

Propriété : $\lim x_n$ existe ssi $\liminf x_n = \limsup x_n$. Dans ce cas, $\lim x_n = \limsup x_n = \liminf x_n$.

Démonstration : $\Leftarrow$: $i_n \leq x_n \leq s_n$. On conclut par théorème d'encadrement.
$\Rightarrow$: Si $x_n \to l$ alors : $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, l - \varepsilon \leq i_n \leq l \leq s_n \leq l + \varepsilon$. Donc $s_n \to l$ et $i_n \to l$.

□

Propriété : Si y_n est une sous-suite de x_n, alors $\liminf x_n \leq \liminf y_n \leq \limsup y_n \leq \limsup x_n$

Ainsi, si l est valeur d'adhérence de x_n, alors $\liminf x_n \leq l \leq \limsup x_n$.

Propriété : $\limsup x_n = -\liminf(-x_n)$

Propriété : Il existe une sous-suite de x_n qui converge vers $\limsup x_n$. Idem pour $\liminf x_n$.

Démonstration : On choisit $k_n \geq n$ tel que $s_n - \frac{1}{n} \leq x_{k_n} \leq s_n$. $n \mapsto x_{k_n}$ converge vers $\limsup x_n$.

□

Familles sommables

On pose $(a_i)_{i\in I}$ famille de nombres positifs.
Définition : $\sum_{i\in I} a_i := \sup\limits_{F\subset I \text{fini}} \sum_{i\ \in F} a_i$

Propriété : Si $\sum_{i\in I} a_i$ est fini, alors $\{i \in I, a_i \neq 0\}$ est au plus dénombrable.

Démonstration : $\{i \in I, a_i \in \mathbb{R} \setminus \{0\}\} \subset \bigcup_{k\in\mathbb{N}} \underbrace{\{i \in I, a_i \geq \frac{1}{k}\}}_{\# \leq k \sum_{i\in I} a_i}$

$\square$

À partir de maintenant, on considérera I dénombrable.

Propriété : Si $\sigma : \mathbb{N} \to I$ est une bijection, alors $\sum_{i\in I} a_i = \lim\limits_{n\to+\infty} \sum_{k=1}^{n} a_{\sigma(k)} =: \sum_{k=1}^{+\infty} a_{\sigma(k)}$

Démonstration : $\forall F \subset I$ fini, $\sigma^{-1}(F)$ est fini donc majoré par un entier N.

$\sum_{i\in F} a_i = \sum_{k\in\sigma^{-1}(F)} a_{\sigma(k)} \leq \sum_{k=1}^{N} a_{\sigma(k)} \leq \sum_{k=1}^{+\infty} a_{\sigma(k)}$

Donc par passage au sup : $\sum_{i\in I} a_i \leq \sum_{k=1}^{+\infty} a_{\sigma(k)}$.

Réciproquement, $\sum_{k=1}^{N} a_{\sigma(k)} = \sum_{i\in\sigma([\![1,N]\!])} a_i \leq \sum_{i\in I} a_i$. On conclut par passage à la limite.

$\square$

Corollaire : Si $(a_k) \in \mathbb{R}_+^{\mathbb{N}}, \sum_{k=1}^{+\infty} a_k = \sum_{k=1}^{+\infty} a_{\sigma(k)}$ et ce quel que soit $\sigma : \mathbb{N} \to \mathbb{N}$ bijection.

En particulier dans le cas $I = \mathbb{N}^2, (a_{i,j})_{(i,j)\in I} \in \mathbb{R}_+^I$:

Propriété : $\sum_{(i,j)\in I} a_{i,j} = \sum_{i=1}^{+\infty} \left(\sum_{j=1}^{+\infty} a_{i,j}\right) = \sum_{j=1}^{+\infty} \left(\sum_{i=1}^{+\infty} a_{i,j}\right)$

Démonstration : $F \subset I$ fini. Il existe $N \in \mathbb{N}$ tel que $F \subset [\![1,N]\!]^2$. Donc $\sum_{(i,j)\in F} a_{i,j} \leq \sum_{i=1}^{N} \sum_{j=1}^{N} a_{i,j} \leq \sum_{i=1}^{N} \sum_{j=1}^{+\infty} a_{i,j} \leq \sum_{i=1}^{+\infty} \sum_{j=1}^{+\infty} a_{i,j}$.

Réciproquement, $\forall N \in \mathbb{N}, \forall M \in \mathbb{N}, \sum_{i=1}^{N} \sum_{j=1}^{M} a_{i,j} \leq \sum_{(i,j)\in\mathbb{N}^2} a_{i,j}$.

Donc $(M \to +\infty)$, $\sum_{i=1}^{N} \sum_{j=1}^{+\infty} a_{i,j} \leq \sum_{(i,j)\in\mathbb{N}^2} a_{i,j}$.

Donc $(N \to +\infty)$, $\sum_{i=1}^{+\infty} \sum_{j=1}^{+\infty} a_{i,j} \leq \sum_{(i,j)\in\mathbb{N}^2} a_{i,j}$.

$\square$

Séries absolument convergentes

Soit $(a_i)_{i\in I}$ une famille de réels tels que $\sum_{i\in I} |a_i|$ soit finie.
On définit $a_i^+ := \max(a_i, 0)$, $a_i^- := \max(-a_i, 0)$.
Donc $a_i^+ - a_i^- = a_i$ et $a_i^+ + a_i^- = |a_i|$.

Propriété : $\sum_{i\in I} a_i^+ - \sum_{i\in I} a_i^- = \sum_{k=1}^{+\infty} a_{\sigma(k)}$ et ce quel que soit $\sigma : \mathbb{N} \to I$ bijection.

Démonstration : $\sum_{i\in I} a_i^+ \leq \sum_{i\in I} |a_i|$ donc la somme est finie. Idem pour $\sum_{i\in I} a_i^-$.

$\sum_{k=1}^{n} a_{\sigma(k)} = \sum_{k=1}^{n} a_{\sigma(k)}^+ - \sum_{k=1}^{n} a_{\sigma(k)}^- \underset{n\to+\infty}{\to} \sum_{k=1}^{+\infty} a_{\sigma(k)}^+ - \sum_{k=1}^{+\infty} a_{\sigma(k)}^-$

□

Corollaire : Sous réserve de convergence absolue, on a :

$$\sum_{k=1}^{+\infty} a_k = \sum_{k=1}^{+\infty} a_{\sigma(k)}$$

$$\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} a_{i,j} = \sum_{j=1}^{+\infty}\sum_{i=1}^{+\infty} a_{i,j}$$

Vocabulaire

Définition : Soit X un ensemble. On dit que $\mathcal{A} \subset \mathcal{P}(X)$ est :
- une algèbre (d'ensembles) si elle est stable par union finie, intersection finie et passage au complémentaire, contient $\emptyset$ et X.
- une tribu (ou σ-algèbre) si c'est une algèbre stable par réunion/intersection dénombrable.

Exemple :
- $\mathcal{P}(X)$ est une tribu.
- $\{\emptyset, X\}$ est une tribu.

Si on se donne une partition finie de X : $X = X_1 \sqcup X_2 \cdots \sqcup X_k$, alors l'ensemble des $A \subset X$ de la forme $A = \bigcup_{n\in I\subset[\![1,k]\!]} X_n$ est une tribu finie.

Lemme : Toute algèbre finie est associée à une partition finie.

Démonstration : Soit $\mathcal{A}$ une algèbre finie.

$\forall x \in X, A(x) := \bigcap_{\substack{A\in\mathcal{A}\\ x\in A}} A$.

Pour x et y donnés, soit $A(x) = A(y)$, soit $A(x) \cap A(y) = \emptyset$.

Fixons $x \in X, B \in \mathcal{A}$.
- Soit $x \in B$ et alors $A(x) \subset B$.
- Soit $x \in {}^cB$ et alors $A(x) \subset {}^cB$ i.e. $A(x) \cap B = \emptyset$

On conclut avec $B = A(y)$.

□

Définition : Si $\mathcal{A}$ est une algèbre de X et $m : \mathcal{A} \to [0, +\infty]$ une fonction.
On dit que m est une *mesure additive* si :

— $m(\emptyset) = 0$
— $m(A \sqcup B) = m(A) + m(B) \qquad (A \cap B = \emptyset)$
Définition : Si $\mathcal{T} \subset \mathcal{P}(X)$ est une tribu, $m : \mathcal{T} \to [0, +\infty]$ est une *mesure* si :
— $m(\emptyset) = 0$
— $m(\bigsqcup_{i \in I} A_i) = \sum_{i \in I} m(A_i)$ pour $(A_i)_{i \in I}$ famille dénombrable disjointe.
Remarque : Toute mesure est une mesure additive.
Remarque : On appelle parfois les mesures "mesures σ-additives".
Remarque : Lorsque $m : \mathcal{A} \to [0, +\infty]$ est une mesure additive sur une algèbre, les propriétés suivantes sont équivalentes :

1. Si $A_i \in \mathcal{A}$ sont disjoints, (A_i) dénombrable, $\bigsqcup_{i \in I} A_i \in \mathcal{A}$, alors $m(\bigsqcup_{i \in I} A_i) = \sum_{i \in I} m(A_i)$
2. Si $A, A_i \in \mathcal{A}$, $A \subset \bigcup_{i \in I} A_i$, alors $m(A) \leq \sum_{i \in I} m(A_i)$.

Dans ce cas, on dit que m est σ-additive.
Démonstration : $(1) \Rightarrow (2)$:
Soit $A_i \in \mathcal{A}$. On définit $\tilde{A}_i$ par : $\tilde{A}_1 = A_1, \ldots \tilde{A}_n = A_n \setminus \tilde{A}_{n-1} \quad \forall n \geq 1$
Alors $\bigcup A_i = \bigsqcup \tilde{A}_i$.
Si $A \subset \bigcup A_i$, alors $A \subset \bigsqcup \tilde{A}_i$. Alors $A = \bigsqcup (A \cap \tilde{A}_i)$.
Donc $m(A) = m(\bigsqcup (\tilde{A}_i \cap A)) \leq \sum m(A_i)$.
$(2) \Rightarrow (1)$:
Si $A = \bigsqcup A_i \overset{(2)}{\Rightarrow} m(A) \leq \sum m(A_i)$.
$A \supset \bigsqcup_{i=1}^{n} A_i$ quel que soit n.
Donc $m(A) \geq \sum_{i=1}^{n} m(A_i)$. Donc $(n \to +\infty)$, $m(A) \geq \sum_{i=1}^{+\infty} m(A_i)$.

□

Définition : Soit $f : \Omega \to X$ une application. Si $\mathcal{A}$ est une algèbre (ou une tribu) sur Ω, alors on définit l'algèbre (tribu) image par :

$$f_* \mathcal{A} = \{A \subset X, f^{-1}(A) \in \mathcal{A}\}$$

Si $\mathcal{A}$ est une algèbre (tribu) sur X, alors

$$f^* \mathcal{A} = \{f^{-1}(A), A \in \mathcal{A}\}$$

est une algèbre (tribu) sur Ω.

La vérification du fait que $f^* \mathcal{A}$ et $f_* \mathcal{A}$ est une algèbre (tribu) découle des propriétés des préimages :

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$$
$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$$
$$f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$$

Définition : Si $f : (\Omega, \mathcal{A}, m) \to X$ est une application, on définit la mesure image (ou la loi) comme la mesure :

$$(f_* m)(Y) := m(f^{-1}(Y))$$

définie sur $f_*\mathcal{A}$.

Définition : m est dite finie ssi $m(X) < +\infty$

Définition : m est dite de probabilité ssi $m(X) = 1$

Définition : $f : (\Omega, \tau) \to (X, \mathcal{T})$ est dite mesurable si :

$$\forall Y \in \mathcal{T}, f^{-1}(Y) \in \tau$$

i.e.

$$f_*\tau \supset \mathcal{T}$$
$$f^*\mathcal{T} \subset \tau$$

Exercice : Soit Ω, X des ensembles, $\mathcal{T}$ une tribu sur X. Soit $f : \Omega \to X$ une application, $g : \Omega \to X$ une application à valeurs dans un ensemble fini Y. Alors g est $f^*\mathcal{T}$ mesurable ssi $\exists h : (X, \mathcal{T}) \to (Y, \mathcal{P}(Y))$ mesurable telle que $g = h \circ f$. i.e. "g est f-mesurable ssi g ne dépend que de f".

Modélisation d'une expérience aléatoire finie (ex : jets de dés)

Soit Y un ensemble fini représentant les issues possibles. Il y a 2 manières de représenter un tirage aléatoire sur Y.

1. On se donne une mesure de probabilité sur $(Y, \mathcal{P}(Y))$. Pour ceci, il suffit de donner $p : Y \to [0,1]$ tel que $\sum_{y \in Y} p(y) = 1$. On note P la mesure de probabilité ainsi créée.
2. On se donne un espace de probabilité abstrait $(\Omega, \mathcal{T}, \mathbb{P})$ et une application mesurable $f : \Omega \to Y$ telle que $f_*\mathbb{P} = P$.

Pour passer de 1. à 2., il suffit de prendre $\Omega = Y$, $\mathcal{T} = \mathcal{P}(Y)$, $\mathbb{P} = P$, $f = \mathrm{id}$.

L'expérience aléatoire consistant à jeter un nombre fini k de dés de valeurs possibles $Y_1, \ldots, Y_k$ est simplement une expérience aléatoire à valeurs dans le produit $Y = Y_1 \times Y_2 \times \cdots \times Y_k$.

La description en termes de variables aléatoires consiste donc à se donner une application mesurable $f : (\Omega, \mathcal{T}, P) \to Y$, c'est à dire k applications mesurables $f_i : (\Omega, \mathcal{T}, P) \to Y_i$, définies sur *le même espace de probabilités*.

Définition : La loi de f (qui est une probabilité sur Y) est dite *loi jointe*. Les lois des f_i (qui sont des probabilités sur Y_i) sont dites *lois marginales*.

Remarque : La loi jointe détermine les lois marginales, qui peuvent se décrire explicitement par $m_i(y_i) = \sum\limits_{y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_k} m(y_1, \ldots, y_k$.

Plus abstraitement, ce soint les mesures images $m_i = (\Pi_i)_* m_i$ où $\Pi_i : Y \to Y_i$ est la projection.

Remarque : La loi jointe est déterminée par $|Y_1| \times \cdots \times |Y_k| - 1$ nombres réels (-1 à cause de la contrainte $\sum p = 1$).

Les lois maginales sont déterminées par $|Y_1| + \cdots + |Y_k| - k$ nombres réels, ce qui est beaucoup moins.

Si on se donnes les marginales $m_1, \ldots, m_k$, ilm existe de nombreuses lois jointes qui engendrent ces marginales. L'une d'entre elles est particulièrement intéressante : la loi produit $m((y1_, \dots, y_k)) = m_1(y_1) \cdot \cdots \cdot m_k(y_k)$, qui correspond (par définition) à des expériences indépendantes.

Définition :

— Les événement A, B dans un espace de probabilité $(\Omega, \mathcal{T}, P)$ sont dits indépendants si $P(A \cap B) = P(A)P(B)$.

— Si $(X_i, \mathcal{T}_i)_{1 \leq i \leq k}$ sont des espaces mesurables (c'est à dire munis de tribus $\mathcal{T}_i$), les variables aléatoires (applications mesurables) $f_i : (\Omega, \mathcal{T}, P) \to (X_i, \mathcal{T}_i)$ sont dites *indépendantes* si $\forall Z_i \in \mathcal{T}_i, P(f_1 \in Z_1, \dots, P_k \in Z_k) = P(f_1 \in Z_i) \cdot \cdots \cdot P(f_k \in Z_k)$

Propriété : Les événements A et B sont indépendants ssi les variables aléatoires $\mathbb{1}_A, \mathbb{1}_B : (\Omega, \mathcal{T}, P) \to \{0, 1\}$ le sont.

Démonstration : Il suffit de montrer que ${}^c A$ et B sont indépendants (le reste est évident ou vient par symétrie).

$$
\begin{aligned}
P({}^c A \cap B) &= P((\Omega \setminus A) \cap B) \\
&= P(B \setminus A \cap B) \\
&= P(B) - P(A \cap B) \\
&= P(B) - P(A)P(B) \\
&= (1 - P(A))P(B) \\
&= P({}^c A)P(B)
\end{aligned}
$$

□

Définition : Les événements $A_1, \dots, A_k$ sont dits indépendants si $\mathbb{1}_{A_1} \dots \mathbb{1}_{A_k} : \Omega \to \{0, 1\}$ le sont.

Remarque : Il ne suffit pas d'avoir l'indépendance deux à deux ou $P(A_1 \cap \dots \cap A_k) = P(A_1) \cdot \cdots \cdot P(A_k)$.

Propriété : Il suffit d'avoir $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot \cdots \cdot P(A_{i_k})$, et ce $\forall \{i_1, \dots, i_k\} \subset [\![1, k]\!]$.

Démonstration : Il faut montrer que

$$
(*) \quad P(B_1 \cap \dots \cap B_k) = P(B_1) \cdot \cdots \cdot P(B_k) \forall B_i \in \{\emptyset, A_i, {}^c A_i, \Omega\}
$$

Il découle de l'hypothèse que c'est vrai pour $B_i \in \{\emptyset, A_i, \Omega\}$.

Il suffit donc de constater que $(*)$ implique $P({}^c B_1 \cap B_2 \cap \dots \cap B_k) = P({}^c B_1)P(B_2) \cdot \cdots \cdot P(B_k)$, ce qui se montre comme ce-dessus. On conclut par récurrence finie.

□

Exemple : Tirage non indépendant :

On tire $-$ chiffres dans $[\![1, 6]\!]$, en leur imposant d'être distincts. La loi jointe est donc : $P(y_1, \dots, y_6) = \begin{cases} 0 & \text{si non distincts} \\ \frac{1}{6!} & \text{si distincts} \end{cases}$.

Les lois marginales sont : $P_1(y_1) := \sum\limits_{y_2, \dots, y_k} P(y_1, \dots, y_k) = \frac{5!}{6!} = \frac{1}{6}$. Les lois marginales sont donc les mêmes que pour un tirage indépendant !

Définition : On dit que $f_i,\ i \in I$ sont indépendantes si $f_i,\ i \in F$ le sont pour tout $F \subset I$ fini.

Modélisation d'une suite infinie de jets dés indépendants

Donnons-nous une suite infinie d'espaces de probabilités finis (Y_i, P_i) (la tribu est $\mathcal{P}(Y_i)$).

Pour chaque n, on a vu que l'on peut trouver des variables aléatoires indépendantes $f_i : (\Omega_n, \mathcal{T}_n, P_n) \to Y_i$ de loi P_i.

Question : peut-on prendre $(\Omega_n, \mathcal{T}_n, P_n)$ indépendant de n ?

> **Théorème 1**
> Il existe un espace de probabilité $(\Omega, \mathcal{T}, P)$ et une suite de variables aléatoires $f_i : \Omega \to Y_i$ qui sont indépendantes et de loi P_i.

Remarque : Les variables aléatoires $f_i,\ i \in \mathbb{N}$ sont indépendantes ssi $f_1, \dots, f_n$ le sont pour tout n.

L'hypothèse d'indépendance consiste donc à dire que, pour tout n et pour tout $(y_1, \dots, y_n) \in Y_1 \times \dots \times Y_n$, l'événement $\{f_1 = y_1, \dots, f_n = y_n\}$ est mesurable $() \in \mathcal{T})$ et de mesure $P(f_1 = y_1, \dots f_n = y_n) = P_1(y_1) \cdot \dots \cdot P_n(y_n)$.

En termes de loi, ceci implique que $\{y_1\} \times \dots \{y_n\} \times Y_{n+1} \times \dots$ est mesurable sur $X := \prod Y_i$ et que sa mesure est $m(\{y_1\} \times \dots \times Y_{n+1}) = P_1(y_1) \cdot \dots \cdot P_n(y_n)$.

Introduction de l'algèbre $\mathcal{A}_\infty$ engendrée par les cylindres finis

Sur le produit $X = \prod Y_i$, pour n fixé, les ensembles de la forme $\{y_1\} \times \dots \times \{y_n\} \times Y_{n+1} \times \dots$ forment une partition finie (ce sont les cylindres finis), qui engendre une algèbre finie $\mathcal{A}_n$ (qui est donc aussi une tribu).

C'est l'algèbre engendrée par les n premières coordonnées. En effet si $\Pi : X \to Y_1 \times \dots \times Y_n$ est la projection, alors $\mathcal{A}_n = \Pi^*(\mathcal{P}(Y_1 \times \dots \times Y_n))$.

Cette algèbre décrit les parties de X qui peuvent être décrites en termes des n premières coordonnées.

On a $\mathcal{A}_n \subset \mathcal{A}_{n+1}$. On note $\mathcal{A}_\infty = \bigcup_{n \geq 1} \mathcal{A}_n$.

$\mathcal{A}_\infty$ est donc l'algèbre des parties de X qui dépendent d'un nombre fini de coordonnées. C'est l'algèbre engendrée par les cylindres finis.

Contrairement aux $\mathcal{A}_n$, $\mathcal{A}_\infty$ est infinie et ce n'est pas une tribu !

L'hypothèse d'indépendance des f_i implique que la loi m doit être définie sur $\mathcal{A}_\infty$, et qu'elle y est déterminée par la relation

$$(*) \quad m(\{y_1\} \times \dots \times \{y_n\} \times Y_{n+1} \times \dots) = P_1(y_1) \cdot \dots \cdot P(y_n)$$

Théorème 2
Il existe sur $X = \prod Y_i$ une tribu τ, qui contient $\mathcal{A}_\infty$, et une mesure m sur $\mathcal{T}$ qui vérifie $(*)$.

On vient en fait de voir que le théorème 1. implique le théorème 2. Réciproquement, il suffit de prendre $\Omega = X, \mathcal{T} = \tau, P = m, f =$ projection.

Pour démontrer l'utilité du théorème 2., donnons des exemples d'ensembles qu'il est naturel de considérer et qui sont dans τ mais pas dans $\mathcal{A}_\infty$. On suppose $Y_i \subset \mathbb{R}$

Exemple : L'ensemble $\{(y_i) \in X, \frac{y_1 + \ldots y_n}{n} \to l\}$ est mesurable. En effet, il s'écrit : $\bigcap_{k\geq 1} \bigcup_{n\in\mathbb{N}} \bigcap_{m\geq n} \{\left|\frac{y_1+\ldots y_n}{n} - l\right| \leq \frac{1}{k}\}$, i.e. $\forall k \geq 1, \exists n \in \mathbb{N}, \forall m \geq n, \ldots$. Chacun des ensembles est dans $\mathcal{A}_\infty$ donc l'ensemble considéré est dans τ.

$\liminf$ et $\limsup$ d'ensembles

Si A_n est une suite d'ensembles, on note :

$$\liminf A_n = \bigcup_n \bigcap_{m\geq n} A_m = \{A_i \text{ APCR}\}$$

$$\limsup A_n = \bigcap_n \bigcup_{m\geq n} A_m = \{A_i \text{ infinitely often (i.o.)}\}$$

Si τ est une tribu, que les $A_n \in \tau$, alors $\limsup A_n \in \tau$ et $\liminf A_n \in \tau$.

Propriété :
— $\liminf A_n \subset \limsup A_n$
— $\liminf^c A_n =^c (\limsup A_n)$

Démonstration : $\forall m, M,\quad \bigcap_{n\geq m} A_n \subset \underset{n\geq M}{A}_n$. Donc $\bigcap_{n\geq m} A_n \subset \limsup A_n$, donc $\liminf A_n \subset limsupA_n$

□

Exercice : $\limsup \mathbb{1}_{A_n} = \mathbb{1}_{\limsup A_n}$

Exemple : On considère un tirafe aléatoire indépendant $f_n \in -1, 1^{\mathbb{N}}$, ce que l'on voit comme un jeu de hasard (le joueur gagne ou perd 1 à chaque étape). Étant donnée la richesse initiale r_0 et un objectif R, on considère l'événement $\{$le joueur atteint la richesse R avant de se ruiner$\}$.

Il s'écrit $\bigcup_{n\geq 1} \{y_1 + \ldots y_n \geq -r_0 \quad \forall k < n \text{ et } y_1 + \cdots + y_k = R - r_0\}$.

C'est une réunion dénombrable d'éléments de $\mathcal{A}_\infty$

Le théorème 2 sera déductible du théorème suivant :

Théorème 3 Hahn-Kolmogorov
Soit $\mathcal{A}$ une algèbre d'ensembles sur X. Soit $\underline{m}$ une mesure de probabilité additive sur $\mathcal{A}$, qui vérifie la propriété de σ-additivité.
Alors il existe une tribu τ contenant $\mathcal{A}$, et une mesure de proba m sur τ qui prolonge $\underline{m}$. De plus, on peut prendre : $m(B) = {}_{B\subset\bigcup A_i}\sum_{i\in\mathbb{N}} \underline{m}(A_i)$, où le inf est pris sur les recouvrements dénombrables de B par des éléments de $\mathcal{A}$.