UNE NOTE DE PROBABILITÉS, FORMALISME DE BASE

Cette note regroupe un matériel d'origines assez diverses : notions enseignées, notions qu'on voudrait pouvoir enseigner, notions simples mais plus avancées, notions plus avancées dont on a oublié la simplicité. Elle constitue pour nous l'occasion de faire quelques révisions et d'abuser d'un style assez formel.

Nous emploierons l'expression « au plus dénombrable » pour qualifier un ensemble fini ou équipotent à l'ensemble des entiers naturels $\mathbb{N}$ même si pour nous l'expression « dénombrable », plus courte, a le même sens. Pour désigner une collection au plus dénombrable, nous noterons $(A_n)_n$, par exemple, sans préciser l'ensemble des indices considérés (fini ou infini dénombrable). L'expression « collection quelconque » désigne une collection dont le cardinal est quelconque (au plus dénombrable ou plus que dénombrable) et sera indexée par des ensembles notés I ou J — dont le cardinal est le cardinal de la collection. Enfin une réunion vide de parties d'un ensemble Ω est égale à l'ensemble vide et une intersection vide à l'ensemble total Ω.

1. Espaces mesurables, probabilisés

1.1. Espaces mesurables

Définition 1. — Un *espace mesurable* est un couple $(E, \mathcal{E})$ où :

(i) E est un ensemble ;

(ii) $\mathcal{E}$ est une tribu, ou σ-algèbre, sur E, c'est-à-dire une partie de $\mathcal{P}(E)$ contenant E, stable par passage au complémentaire et stable par intersections au plus dénombrables de ses éléments.

Les éléments de la tribu $\mathcal{E}$, qui sont des sous-ensembles de E, sont appelés *ensembles mesurables.*

Notons que si $(E, \mathcal{E})$ est un espace mesurable, par stabilité par passage au complémentaire, $\mathcal{E}$ contient $\varnothing$ et est stable par réunions au plus dénombrables de ses éléments. Ces propriétés peuvent se substituer respectivement aux axiomes d'appartenance de E à $\mathcal{E}$ et de stabilité par intersections au plus dénombrables.

Exemples. — Soit E un ensemble. L'ensemble de parties $\{\varnothing, E\}$ est la tribu grossière (la moins fine), ou triviale, sur E, l'ensemble de parties $\mathcal{P}(E)$ est la tribu discrète sur E, l'ensemble de parties $\{B \subset E : B \text{ ou } B^c \text{ est au plus dénombrable}\}$ est une tribu sur E que nous nommerons ponctuelle. Notons que si E est au plus dénombrable, sa tribu ponctuelle est alors sa tribu discrète.

1.2. Tribus trace

Définition 2. — Soient $(E, \mathcal{E})$ un espace mesurable et $B \subset E$ une partie quelconque de E. On appelle *tribu trace* de $\mathcal{E}$ sur B la tribu $\mathcal{E}_B = \{B \cap C : C \in \mathcal{E}\}$.

Vérifions qu'il s'agit d'une tribu : $B = B \cap E \in \mathcal{E}_B$; si $C = B \cap A$ alors $\complement_B C = B \cap \complement_E A \in \mathcal{E}_B$; si $C_n = B \cap A_n$, alors $\bigcap_n C_n = B \cap \bigcap_n A_n \in \mathcal{E}_B$.

Lorsque B est un sous-ensemble mesurable de E, il est clair que la tribu trace $\mathcal{E}_B$ est l'ensemble des éléments de $\mathcal{E}$ qui sont inclus dans B.

1.3. Espaces probabilisés

Définition 3. — Un *espace probabilisé* est un triplet $(E, \mathcal{E}, \mu)$ où :

(i) E est un ensemble non vide et $(E, \mathcal{E})$ est un espace mesurable ;

(ii) μ est une mesure de probabilité sur $(E, \mathcal{E})$, c'est-à-dire que μ est une application de $\mathcal{E}$ dans $[0, 1]$ telle que $\mu(E) = 1$ et qui est σ-additive, *i.e.*, qui à une réunion au plus dénombrable d'éléments disjoints de $\mathcal{E}$ affecte pour probabilité la somme des probabilités de ces éléments.

Si $(E, \mathcal{E}, \mu)$ est un espace probabilisé, par additivité $\mu(B^c) = 1 - \mu(B)$ pour tout $B \in \mathcal{E}$, en particulier $\mu(\varnothing) = 0$. Notamment, $1 = \mu(E) \neq \mu(\varnothing) = 0$ et ainsi E ne peut être l'ensemble vide. En revanche, $(\varnothing, \{\varnothing\})$ est un espace mesurable tout à fait convenable — même si la seule mesure qu'il peut porter est la mesure nulle.

Lorsque l'espace probabilisé considéré modélise une expérience ou un phénomène aléatoire, on le note le plus souvent $(\Omega, \mathcal{A}, \mathbb{P})$. Les éléments de Ω sont appelés dans ce cas *épreuves*, et Ω est appelé *ensemble des épreuves* (ou « ensemble fondamental » ou encore « univers ») bien que dans le langage courant des probabilistes on l'appelle simplement « grand Omega ». Les éléments de la tribu $\mathcal{A}$, qui sont des sous-ensembles de Ω, sont alors appelés *événements*.

Exemples. — *a*) Soient Ω un ensemble non vide, $\mathcal{A} = \{\varnothing, \Omega\}$ la tribu grossière, $\mathbb{P}$ définie sur $\mathcal{A}$ par $\mathbb{P}(\varnothing) = 0$ et $\mathbb{P}(\Omega) = 1$. Alors $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace probabilisé (grossier).

b) Soient Ω un ensemble non vide, $\mathcal{A}$ une tribu quelconque sur Ω, et $(x_n)_n \subset \Omega$ une partie au plus dénombrable (les $(x_n)_n$ sont distincts). À chaque x_n on associe $p_n \in \mathbb{R}_+$ de telle manière que la somme des p_n soit égale à 1. La fonction $\mathbb{P} : \mathcal{A} \to [0, 1]$, $A \mapsto \sum_n p_n \mathbb{1}_A(x_n)$, où $\mathbb{1}_A(x_n) = 1$ si $x_n \in A$ et $\mathbb{1}_A(x_n) = 0$ si $x_n \notin A$, est une mesure de probabilité sur $(\Omega, \mathcal{A})$. Une telle mesure est qualifiée de purement atomique.

Une telle mesure de probabilité est appelée *mesure de probabilité discrète* lorsque chacun des singletons $\{x_n\}$ est mesurable. On dit alors que ces points portent la mesure de probabilité.

Habituellement, on rencontre ce type de situation avec Ω au plus dénombrable, $\mathcal{A} = \mathcal{P}(\Omega)$ sa tribu discrète, et $(x_n)_n$ la liste de ses points ; le cas le plus classique étant celui où Ω est fini et $p_n = 1/\operatorname{Card}\Omega$ pour tout n, alors $\mathbb{P}(A) = \operatorname{Card} A/\operatorname{Card}\Omega$ pour tout $A \subset \Omega$.

c) Soient Ω un ensemble non vide de cardinal non dénombrable, $\mathcal{A}$ sa tribu ponctuelle. En posant pour $A \in \mathcal{A}$, $\mathbb{P}_0(A) = 0$ si A est au plus dénombrable et $\mathbb{P}_0(A) = 1$ sinon, on définit une application $\mathbb{P}_0 : \mathcal{A} \to [0, 1]$ qui est une mesure de probabilité.

Il est possible de montrer que toute mesure de probabilité sur $(\Omega, \mathcal{A})$ s'écrit de manière unique comme une combinaison barycentrique de $\mathbb{P}_0$ et d'une mesure de probabilité portée par une partie au plus dénombrable de Ω, c'est-à-dire une mesure de probabilité discrète.

1.4. Probabilités conditionnelles

Définition 4. — Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $B \in \mathcal{A}$ un sous-ensemble mesurable de probabilité strictement positive. On appelle mesure de probabilité conditionnelle sachant B l'application $\mathbb{P}(\,.\,|\,B) : \mathcal{A} \to [0,1]$, $A \mapsto \mathbb{P}(A\,|\,B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$. Pour $A \in \mathcal{A}$, le nombre $\mathbb{P}(A\,|\,B)$ est appelé probabilité (conditionnelle) de A sachant B.

On constate aisément que $\mathbb{P}(\,.\,|\,B)$ est une mesure de probabilité sur $(\Omega, \mathcal{A})$. En notant $\mathbb{P}_B$ la restriction de cette mesure à la tribu trace $\mathcal{A}_B$ — qui n'est autre que l'ensemble des éléments de $\mathcal{A}$ inclus dans B du fait de la mesurabilité de celui-ci —, on constate que $(B, \mathcal{A}_B, \mathbb{P}_B)$ est un espace probabilisé.

Remarque. — Contrairement au cadre des espaces mesurables, nous devons devons supposer ici le sous-ensemble B mesurable afin de pouvoir considérer sa probabilité et la restriction de la mesure de probabilité à celui-ci.

2. Applications mesurables, variables aléatoires

2.1. Applications mesurables

Définition 5. — Soient $(E, \mathcal{E})$ et $(F, \mathcal{F})$ deux espaces mesurable. Une application $f : E \to F$ de E dans F est *mesurable* si et seulement si, pour tout $C \in \mathcal{F}$, $f^{-1}(C) \in \mathcal{E}$. On note alors $f : (E, \mathcal{E}) \to (F, \mathcal{F})$.

Théorème 1. — *Soient E un ensemble, $(F, \mathcal{F})$ un espace mesurable et $f : E \to F$ une application. Le sous-ensemble $\sigma(f)$ de $\mathcal{P}(E)$ défini par $\sigma(f) = f^{-1}(\mathcal{F}) = \{f^{-1}(C) : C \in \mathcal{F}\}$ est une tribu sur E, et $f : (E, \sigma(f)) \to (F, \mathcal{F})$ est mesurable. Si $(E, \mathcal{E})$ est un espace mesurable et $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ est mesurable, alors $\sigma(f) \subset \mathcal{E}$.*

Démonstration. — On a $E = f^{-1}(F) \in \sigma(f)$. Puis, si $B = f^{-1}(C) \in \sigma(f)$, alors $B^c = f^{-1}(C)^c = f^{-1}(C^c) \in \sigma(f)$. Enfin, si $B_n = f^{-1}(C_n) \in \sigma(f)$, $\bigcap_n B_n = \bigcap_n f^{-1}(C_n) = f^{-1}(\bigcap_n C_n) \in \sigma(f)$. Ainsi $\sigma(f)$ est une tribu sur E. Les autres affirmations sont alors évidentes. Notons qu'elles signifient que $\sigma(f)$ est la plus petite tribu sur E pour laquelle f est mesurable et qu'elle dépend en outre de la structure mesurable sur l'espace d'arrivée $(F, \mathcal{F})$. □

Théorème 2. — *Soient $(E, \mathcal{E})$, $(F, \mathcal{F})$, $(G, \mathcal{G})$ trois espaces mesurables et $f : (E, \mathcal{E}) \to (F, \mathcal{F})$, $g : (F, \mathcal{F}) \to (G, \mathcal{G})$ deux applications mesurables. Alors, $g \circ f : (E, \mathcal{E}) \to (G, \mathcal{G})$ est mesurable.*

Démonstration. — Soit $D \in \mathcal{G}$. Puisque g est mesurable, $C = g^{-1}(D) \in \mathcal{F}$, et f est mesurable, $B = f^{-1}(C) \in \mathcal{E}$. Donc, $B = f^{-1}(g^{-1}(D)) = (g \circ f)^{-1}(D) \in \mathcal{E}$. On en déduit que $g \circ f$ est mesurable. □

Exemple. — Soient $(E, \mathcal{E})$ un espace mesurable et $B \in \mathcal{E}$. Nous notons $\mathbb{1}_B$ l'application de E dans $\{0, 1\}$ telle que $\mathbb{1}_B(x) = 1$ si $x \in B$ et $\mathbb{1}_B(x) = 0$ si $x \notin B$. Il n'y a que deux tribus sur $\{0, 1\}$ qui sont les tribus grossière et discrète et on constate que pour toutes deux $\mathbb{1}_B$ est mesurable.

2.2. Restrictions des espaces de départ et d'arrivée

Théorème 3. — *Soient $(E,\mathcal{E})$, $(F,\mathcal{F})$ deux espaces mesurables et $f:(E,\mathcal{E})\to(F,\mathcal{F})$ une application mesurable.*

(i) *Pour toute partie $B\subset E$, $f|_B:(B,\mathcal{E}_B)\to(F,\mathcal{F})$ est mesurable.*

(ii) *Pour toute partie $C\subset F$, et si on note $B=f^{-1}(C)$, alors $f|_B:(B,\mathcal{E}_B)\to(C,\mathcal{F}_C)$ est mesurable.*

Démonstration. — (i) Si $C\in\mathcal{F}$, on a $f|_B^{-1}(C)=f^{-1}(C)\cap B$ qui est l'intersection d'un élément de $\mathcal{E}$ avec C et appartient donc à la tribu trace $\mathcal{E}_B$.

(ii) Soit $D=D'\cap C\in\mathcal{F}_C$, où $D'\in\mathcal{F}$. On a $f|_B^{-1}(D)=f^{-1}(D'\cap C)\cap B=f^{-1}(D')\cap f^{-1}(C)\cap B=f^{-1}(D')\cap B$. Puisque $f^{-1}(D')\in\mathcal{E}$ par hypothèse sur f, alors $f^{-1}(D)\in\mathcal{E}_B$ ce qui montre la mesurabilité affirmée de $f|_B$. □

Remarque. — Il est fréquent que les restrictions considérées sur l'espace d'arrivée se limitent à des sous-ensembles contenant l'image de l'application f considérée, image qui n'est pas nécessairement mesurable. Ce que l'on peut énoncer est : *soit $C\subset F$ une partie quelconque telle que $f(E)\subset C$. Alors $f:(E,\mathcal{E})\to(F,\mathcal{F})$ est mesurable si et seulement si $f:(E,\mathcal{E})\to(C,\mathcal{F}_C)$ est mesurable.*

2.3. Variables aléatoires

Si $(\Omega,\mathcal{A},\mathbb{P})$ est un espace probabilisé modélisant une expérience ou un phénomène aléatoire et $(E,\mathcal{E})$ un espace mesurable, les applications mesurables de Ω dans E sont appelées *variables aléatoires*, désignées par une majuscule, par exemple X (le plus souvent par une lettre parmi $S,\dots,Z$), et on note $X:(\Omega,\mathcal{A},\mathbb{P})\to(E,\mathcal{E})$.

Par stabilité de la mesurabilité par composition, si $X:(\Omega,\mathcal{A},\mathbb{P})\to(E,\mathcal{E})$ est une variable aléatoire et $f:(E,\mathcal{E})\to(F,\mathcal{F})$ une application mesurable, $Y=f\circ X:(\Omega,\mathcal{A},\mathbb{P})\to(F,\mathcal{F})$ est une variable aléatoire.

Théorème 4 et définition. — *Soit $X:(\Omega,\mathcal{A},\mathbb{P})\to(E,\mathcal{E})$ une variable aléatoire.*

(i) *Le sous-ensemble $\sigma(X)$ de $\mathcal{A}$ défini par $\sigma(X)=X^{-1}(\mathcal{E})=\{X^{-1}(B):B\in\mathcal{E}\}$ est une sous-tribu de $\mathcal{A}$, elle est nommée* tribu engendrée par X.

(ii) *L'application $P_X:\mathcal{E}\to[0,1]$ définie par $P_X(B)=\mathbb{P}(X^{-1}(B))$, $B\in\mathcal{E}$, est une mesure de probabilité sur $(E,\mathcal{E})$, elle est nommée* loi de la variable aléatoire X.

Démonstration. — (i) Ce point est une application immédiate de l'énoncé similaire pour les applications mesurables.

(ii) Il est clair que l'application P_X est définie sur $\mathcal{E}$ et à valeurs dans $[0,1]$. On a $P_X(E)=\mathbb{P}(X^{-1}(E))=\mathbb{P}(\Omega)=1$. Si $(B_n)_n\subset\mathcal{E}$ est une collection au plus dénombrable d'ensembles mesurables disjoints, il en est de même de $(X^{-1}(B_n))_n\subset\mathcal{A}$ et on a $X^{-1}(\bigcup_n B_n)=\bigcup_n X^{-1}(B_n)$ et ainsi $P_X(\bigcup_n B_n)=\mathbb{P}(X^{-1}(\bigcup_n B_n))=\mathbb{P}(\bigcup_n X^{-1}(B_n))=\sum_n\mathbb{P}(X^{-1}(B_n))=\sum_n P_X(B_n)$. Ainsi, P_X est une mesure de probabilité sur $(E,\mathcal{E})$. □

Remarques. — *a*) On reverra dans la sous-section sur la génération de tribus la notation $\sigma(X)$ dans un contexte un peu plus général. On peut déjà se rappeler que c'est la plus petite tribu sur Ω telle que $X:\Omega\to(E,\mathcal{E})$ est mesurable.

b) D'autre part, il est à noter que si $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{E})$ est une variable aléatoire, l'espace $(E, \mathcal{E}, P_X)$ est un espace probabilisé qui peut servir de modèle probabiliste dès lors qu'on ne s'intéresse qu'à la variable X ou qu'à des fonctions de celle-ci.

2.4. Restrictions des espaces de départ et d'arrivée

Théorème 5. — *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé,* $(E, \mathcal{E})$ *un espace mesurable et* $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{E})$ *une variable aléatoire.*

(i) *Pour toute partie* $A \in \mathcal{A}$ *mesurable de probabilité strictement positive,* $X|_A : (A, \mathcal{A}_A, \mathbb{P}_A) \to (E, \mathcal{E})$ *est une variable aléatoire.*

(ii) *Soit* $B \subset E$ *une partie telle que* $A = X^{-1}(B) = \{X \in B\} \in \mathcal{A}$ *et* $\mathbb{P}(A) > 0$. *Alors,* $X|_A : (A, \mathcal{A}_A, \mathbb{P}_A) \to (B, \mathcal{E}_B)$ *est une variable aléatoire.*

Démonstration. — (i) La mesurabilité de $X|_A : (A, \mathcal{A}_A) \to (E, \mathcal{E})$ découle immédiatement de l'énoncé similaire portant sur les applications mesurables. Les exigences de mesurabilité et de probabilité strictement positive de A permettent simplement de considérer la mesure de probabilité $\mathbb{P}_A$ sur $(A, \mathcal{A}_A)$.

(ii) Là encore cet énoncé est une application immédiate de l'énoncé similaire pour les applications mesurables. Les conditions de mesurabilité de $A = X^{-1}(B)$ et de stricte positivité de $\mathbb{P}(A)$ ont même raison d'être qu'au point précédent. □

Remarques. — *a*) Tout comme pour les applications mesurables, des restrictions fréquentes que l'on fait sur l'espace d'arrivée le sont à des sous-ensembles contenant l'image de X. En fait, on va même plus loin : soit $B \subset E$ tel que $A = X^{-1}(B) \in \mathcal{A}$ et $\mathbb{P}(A) = 1$; le point (ii) s'applique alors mais, abusivement, on identifie A à Ω, $\mathcal{A}_A$ à $\mathcal{A}$ et $\mathbb{P}_A$ à $\mathbb{P}$, de sorte que $X|_A$ est elle aussi identifiée à X, l'idée étant qu'on a jeté un ensemble A^c de mesure nulle et donc négligeable.

b) Assez étrangement c'est l'apparemment plus simple type de variables aléatoires qu'il est le plus délicat à définir, celui de *variable aléatoire discrète*. Une première définition consiste à dire qu'une variable aléatoire $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{E})$ est discrète si E est un ensemble au plus dénombrable et si $\mathcal{E}$ est la tribu discrète. La loi d'une telle variable aléatoire est alors clairement une mesure de probabilité discrète sur $(E, \mathcal{E})$ puisque toute mesure de probabilité sur un tel espace est de ce type. La bonne définition consiste à dire qu'une variable aléatoire $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{E})$ est discrète si sa loi P_X de la variable aléatoire X est une mesure de probabilité discrète sur $(E, \mathcal{E})$: il existe une partie au plus dénombrable $(x_n)_n \subset E$ telle que les singletons $\{x_n\} \in \mathcal{E}$ pour tout n et $P_X((x_n)_n) = \sum_n P_X\{x_n\} = \sum_n \mathbb{P}\{X = x_n\} = \mathbb{P}\{X \in (x_n)_n\} = 1$. Dans ce cas il est courant de restreindre l'espace d'arrivée à d'autres sous-ensembles contenant $(x_n)_n$ voire à $(x_n)_n$ lui-même.

3. **Génération de tribus**

3.1. Généralités

Il est très fréquent qu'une tribu ne soit pas donnée explicitement mais définie comme étant la plus petite tribu possédant certaines propriétés. C'est ce que précise le théorème qui suit.

THÉORÈME 6 ET DÉFINITION. — *Soient E un ensemble et $(B_i)_{i\in I}$ une collection quelconque de parties de E. Il existe une unique tribu sur E appelée tribu engendrée par $(B_i)_{i\in I}$ et notée $\sigma((B_i)_{i\in I})$ contenant $(B_i)_{i\in I}$ telle que toute autre tribu sur E contenant elle aussi $(B_i)_{i\in I}$ contient la tribu $\sigma((B_i)_{i\in I})$.*

LEMME 1. — *Soit E un ensemble et $(\mathcal{B}_j)_{j\in J}$ une collection non vide de tribus sur E. Alors*

$$\bigcap_{j\in J} \mathcal{B}_j = \{B \subset E : B \in \mathcal{B}_j, \forall\, j \in J\}$$

est une tribu sur E.

Démonstration du lemme. — C'est immédiat : puisque pour tout $j \in J$, $E \in \mathcal{B}_j$, alors $E \in \mathcal{B}$; si $B \in \mathcal{B}$, alors, pour tout $j \in J$, B et $B^c \in \mathcal{B}_j$, donc $B^c \in \mathcal{B}$; enfin, si $(B_n)_n \subset \mathcal{B}$, alors, pour tout $j \in J$, $(B_n)_n \subset \mathcal{B}_j$ et $\bigcap_n B_n \in \mathcal{B}_j$, donc $\bigcap_n B_n \in \mathcal{B}$. □

Remarque. — Nous ne définissons pas l'intersection d'une collection vide de tribus sur un ensemble E ; celle-ci devrait être choisie comme étant la tribu discrète $\mathcal{P}(E)$.

Démonstration du théorème. — Pour l'existence, considérons l'ensemble des tribus contenant $(B_i)_{i\in I}$ que nous notons sous la forme d'une collection $(\mathcal{B}_j)_{j\in J}$. Cet ensemble est non vide puisqu'il contient $\mathcal{P}(E)$. Posons $\mathcal{B} = \bigcap_{j\in J} \mathcal{B}_j$. Il est alors clair que $\mathcal{B}$ répond à la question. Quant à l'unicité, soit $\mathcal{B}'$ une seconde tribu satisfaisant aux mêmes propriétés. Puisque $\mathcal{B}'$ fait partie des tribus contenant $(B_i)_{i\in I}$, alors $\mathcal{B} \subset \mathcal{B}'$. Mais, puisque $\mathcal{B}$ contient $(B_i)_{i\in I}$, on a par hypothèse $\mathcal{B}' \subset \mathcal{B}$. D'où l'égalité de ces tribus. □

THÉORÈME 7. — *Soit $(B_i)_{i\in I_j, j\in J}$ une collection quelconque de parties d'un ensemble E. Alors*

$$\sigma\big((B_i)_{i\in I_j, j\in J}\big) = \sigma\Big(\sigma\big((B_i)_{i\in I_j}\big)_{j\in J}\Big).$$

Démonstration. — Puisque, pour tout $j \in J$, tout $i \in I_j$, B_i appartient à la seconde tribu et donc la première tribu est incluse dans la seconde. Réciproquement, pour tout $j \in J$, $\sigma((B_i)_{i\in I_j})$ est inclus dans la première tribu et ainsi on a l'inclusion réciproque. Notons que ce théorème signifie que l'opération σ est associative. □

Remarque. — Lorsque $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux tribus sur un même ensemble E, on note $\mathcal{B}_1 \vee \mathcal{B}_2 = \sigma(\mathcal{B}_1, \mathcal{B}_2)$ la tribu qu'elles engendrent. L'opération binaire $\vee$ est l'opération de maximum sur l'ensemble des tribus d'un même ensemble. Nous avons vu qu'elle est associative et il est clair qu'elle est commutative. Ainsi, il n'y a aucune ambiguïté à considérer des tribus $\mathcal{B}_1 \vee \ldots \vee \mathcal{B}_n$ ou encore $\bigvee_{i\in I} \mathcal{B}_i$. L'opération de minimum correspondante ne pose aucune difficulté puisqu'il s'agit de l'intersersection telle qu'elle est définie dans le lemme précédent. Elle est, bien entendu, elle aussi associative et commutative.

Le lemme suivant sera assez pratique pour la suite. La notion sous-jacente est celle de plus grande tribu sur l'image rendant une application mesurable. Précisons que la notation $\tau(\,.\,)$ employée par la suite n'a rien de standard.

LEMME 2. — *Soient $(E, \mathcal{E})$ un espace mesurable, F un ensemble et $f : E \to F$ une application. Alors, $\tau(f) = \{C \subset F : f^{-1}(C) \in \mathcal{E}\}$ est une tribu sur F.*

Démonstration. — Cette preuve ne repose que sur les propriétés de l'image réciproque : puisque $E = f^{-1}(F)$, on a $F \in \tau(f)$; puis, si $C \in \tau(f)$, puisque

$f^{-1}(C^c) = f^{-1}(C)^c$, on a $C^c \in \tau(f)$; enfin, si $(C_n)_n \subset \tau(f)$, comme $f^{-1}(\bigcap_n C_n) = \bigcap_n f^{-1}(C_n)$, on a $\bigcap_n C_n \in \tau(f)$. □

THÉORÈME 8. — *Soient $(\Omega, \mathcal{A})$ et $(E, \mathcal{E})$ deux espaces mesurables, $(B_i)_{i\in I}$ une collection quelconque de parties de E engendrant la tribu $\mathcal{E}$ — notamment $(B_i)_{i\in I} \subset \mathcal{E}$ — et X une application de Ω dans E. Alors, l'application X est mesurable si et seulement si, pour tout $i \in I$, $X^{-1}(B_i) \in \mathcal{A}$.*

Démonstration. — La condition est évidemment nécessaire. Elle est suffisante car par hypothèse la tribu $\tau(X)$ sur $\mathcal{E}$ contient $(B_i)_{i\in I}$ et donc la tribu $\mathcal{E}$. □

Exemple. — La tribu sur un ensemble E engendrée par ses points est la tribu ponctuelle.

3.2. Tribus boréliennes

DÉFINITION 6. — Soit E un espace topologique. La *tribu borélienne* de E est la tribu sur E engendrée par la collection de ses ouverts. Elle est notée $\mathcal{B}(E)$.

Une application mesurable entre deux espaces topologiques munis de leurs tribus boréliennes est dite *borélienne*.

Remarque. — Par passage au complémentaire, une tribu borélienne est aussi engendrée par les fermés.

THÉORÈME 9. — *Soient E et F deux espaces topologiques, et $f : E \to F$ une application continue. Alors $f : (E, \mathcal{B}(E)) \to (F, \mathcal{B}(F))$ est mesurable,* i.e.*, f est borélienne.*

Démonstration. — Puisque f est continue, les images réciproques des ouverts de F par f sont ouverts dans E et sont donc dans $\mathcal{B}(E)$. La tribu $\tau(f)$ sur F définie au lemme 2 contient donc tous les ouverts, elle contient alors $\mathcal{B}(F)$, ce qui montre que f est mesurable pour les tribus boréliennes. □

Exemple. — Soit E un ensemble. La plus petite topologie sur E telle que les points soient fermés sera nommée topologie ponctuelle ; ses fermés sont les parties finies, $\varnothing$ et E et ses ouverts sont donc les complémentaires des parties finies, $\varnothing$ et E. On constate aisément que la tribu borélienne de la topologie ponctuelle est exactement la tribu ponctuelle sur E. En remplaçant « finies » par « dénombrables », on obtient encore une topologie sur E (topologie σ-ponctuelle ?) qui engendre aussi la tribu ponctuelle.

3.3. Tribus produit

DÉFINITION 7. — Soit $((E_i, \mathcal{E}_i))_{i\in I}$ une collection d'espaces mesurables. La *tribu produit* sur $\prod_{i\in I} E_i$ des tribus $(\mathcal{E}_i)_{i\in I}$ est la tribu engendrée par l'ensemble des $\prod_{i\in I} B_i$ où $B_i \in \mathcal{E}_i$ pour tout $i \in I$ et $B_i = E_i$ sauf pour un nombre fini d'indices. Elle est notée $\bigotimes_{i\in I} \mathcal{E}_i$.

Remarques. — *a*) Lorsque l'ensemble d'indice I est fini, la tribu produit est simplement engendrée par l'ensemble des pavés, ou cylindres, $\prod_{i\in I} B_i$ où $B_i \in \mathcal{E}_i$ pour tout $i \in I$.

b) On appellera pavé élémentaire dans le cas général tout produit d'ensembles mesurables égaux aux espaces entiers sauf pour un nombre fini (ou nul) d'indices.

c) En fait, la tribu produit est engendrée par les pavés élémentaires dont au plus un facteur diffère de l'espace total puisque les pavés élémentaires sont des intersections finies de tels pavés, pavés qui seront qualifiés de pavés basiques.

Le rôle des pavés élémentaires apparaît avec la considération de mesures sur un espace produit, de loi jointe de variables aléatoires et de la propriété d'indépendance, et ce parce que les réunions finies de pavés — qu'on peut supposer disjoints — forment une algèbre engendrant la tribu produit.

Exemples. — Les espaces $\{0,1\}^n$, $\{0,1\}^{\mathbb{N}^*}$ sont munis des tribus produit. Dans le premier cas la tribu est égale à la tribu discrète, dans le second elle est strictement plus petite. L'espace $\mathbb{R}^n$ peut être muni de sa tribu borélienne ou de sa tribu produit ; ces tribus sont égales car la métrique usuelle de $\mathbb{R}^n$ est équivalente à la métrique produit.

Théorème 10. — *Soient* $((E_i,\mathcal{E}_i))_{i\in I}$ *une collection d'espaces mesurables,*

$$(E,\mathcal{E}) = \left(\prod_{i\in I} E_i, \bigotimes_{i\in I} \mathcal{E}_i\right)$$

et $(\Omega,\mathcal{A})$ *un espace mesurable. Une application* $X : \Omega \to E$ *est mesurable si et seulement si chacune de ses coordonnées* $X_i : (\Omega,\mathcal{A}) \to (E_i,\mathcal{E}_i)$ *est mesurable.*

Lemme 3. — *Les applications coordonnées* $f_i : (E,\mathcal{E}) \to (E_i,\mathcal{E}_i)$ *sont mesurables.*

Démonstration du lemme. — Si $B_i \in \mathcal{E}_i$, $f_i^{-1}(B_i) = \prod_{j\in I} B_j$ où $B_j = E_j$ si $j \neq i$. C'est un pavé basique qui est donc mesurable dans la tribu produit. □

Démonstration du théorème. — Si $X : (\Omega,\mathcal{A}) \to (E,\mathcal{E})$ est mesurable, d'après le lemme et par composition, $X_i = f_i \circ X : (\Omega,\mathcal{A}) \to (E_i,\mathcal{E}_i)$ est mesurable. Réciproquement, supposons les coordonnées X_i mesurables. Alors la tribu $\tau(X)$ sur E contient les pavés basiques et donc la tribu produit. Ainsi, l'application X est mesurable. □

3.4. Tribus engendrées par des applications

Définition 8. — Soient Ω un ensemble, $((E_i,\mathcal{E}_i))_{i\in I}$ une collection d'espaces mesurables et $(X_i : \Omega \to E_i)_{i\in I}$ une collection d'applications. La *tribu engendrée par les applications* $(X_i)_{i\in I}$ est la tribu sur Ω engendrée par $\{X_i^{-1}(B_i) : B_i \in \mathcal{E}_i,\ i \in I\}$. Elle est notée $\sigma((X_i)_{i\in I})$.

La tribu $\sigma((X_i)_{i\in I})$ est la plus petite tribu sur Ω pour laquelle les applications $(X_i)_{i\in I}$ sont mesurables. En particulier, si Ω est muni d'une tribu $\mathcal{A}$ pour laquelle les applications $(X_i)_{i\in I}$ sont mesurables, alors $\sigma((X_i)_{i\in I})$ est une sous-tribu de $\mathcal{A}$.

Théorème 11. — (i) *Soient* Ω *un ensemble,* $(E,\mathcal{E})$ *un espace mesurable et* X *une application de* Ω *dans* E. *Alors* $\sigma(X) = X^{-1}(\mathcal{E}) = \{X^{-1}(B) : B \in \mathcal{E}\}$.

(ii) *Soient* Ω *et* E *deux ensembles,* $(B_i)_{i\in I}$ *une collection quelconque de parties de* E *et* X *une application de* Ω *dans* E. *Alors les tribus, sur* Ω, $X^{-1}(\sigma((B_i)_{i\in I}))$ *et* $\sigma((X^{-1}(B_i))_{i\in I})$ *sont égales.*

(iii) *Soient* Ω *un ensemble,* $((E_i,\mathcal{E}_i))_{i\in I}$ *une collection d'espaces mesurables et* $(X_i : \Omega \to E_i)_{i\in I}$ *une collection d'applications. Posons* $X = (X_i)_{i\in I}$ *vue comme*

une application de Ω à valeurs dans l'espace produit $E = \prod_{i\in I} E_i$ muni de sa tribu produit $\mathcal{E} = \bigotimes_{i\in I} \mathcal{E}_i$. On a alors $\sigma((X_i)_{i\in I}) = \sigma(X) = X^{-1}(\mathcal{E})$.

Démonstration. — (i) Il suffit de vérifier que $X^{-1}(\mathcal{E})$ est une tribu puisque $X^{-1}(\mathcal{E}) \subset \sigma(X)$ et que $X^{-1}(\mathcal{E})$ engendre $\sigma(X)$.

(ii) Nous en venons à une propriété importante qui est la commutation de l'opération de génération de tribu et celle d'image réciproque. Posons $\mathcal{E} = \sigma((B_i)_{i\in I})$ et considérons une tribu $\mathcal{A}$ sur Ω. Nous avons vu que X est mesurable de $(\Omega, \mathcal{A})$ dans $(E, \mathcal{E})$ si et seulement si $X^{-1}(B_i) \in \mathcal{A}$ pour tout $i \in I$. Notamment, si $\mathcal{A} = \sigma((X^{-1}(B_i))_{i\in I})$, alors $X : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$ est mesurable. Par conséquent $\sigma(X) = X^{-1}(\mathcal{E}) \subset \mathcal{A}$. Puisque par ailleurs $\sigma(X)$ est une tribu sur Ω contenant tous les $X^{-1}(B_i)$ pour $i \in I$, on a alors $\mathcal{A} = \sigma((X^{-1}(B_i))_{i\in I}) \subset \sigma(X)$. D'où l'égalité annoncée.

(iii) La tribu $X^{-1}(\mathcal{E})$ est la plus petite tribu sur Ω telle que X soit mesurable. Or X est mesurable si et seulement si chacune de ses coordonnées l'est, et la plus petite tribu sur Ω réalisant cette dernière condition est $\sigma((\sigma(X_i)_{i\in I}))$ qui est égale à $\sigma((X_i)_{i\in I})$, la conclusion s'ensuit. □

4. **Ensembles négligeables, complétion**

4.1. Ensembles négligeables

Définition 9. — Soit $(E, \mathcal{E}, \mu)$ un espace probabilisé.

(i) Une partie B de E est intérieurement négligeable si et seulement si quelque soit $B^- \in \mathcal{B}$, $B^- \subset B$, $\mu(B^-) = 0$.

(ii) Une partie B de E est extérieurement négligeable, ou plus simplement négligeable, si et seulement si il existe $B^+ \in \mathcal{B}$, $B^+ \supset B$, $\mu(B^+) = 0$.

Remarquons qu'il est possible de déclarer si une partie quelconque $B \subset E$ est intérieurement négligeable ou non : en effet le sous-ensemble de $\mathcal{E}$ des parties mesurables incluses dans B n'est pas vide puisqu'il contient l'ensemble vide ; nous pouvons définir sa mesure intérieure comme étant le supremum des probabilités de ces parties ; si ce supremum est nul, B est intérieurement négligeable, sinon, il ne l'est pas.

De même nous pouvons définir la mesure extérieure de $B \subset E$ comme étant l'infimum des probabilités des ensembles mesurables contenant B. S'il est clair qu'un ensemble est intérieurement négligeable si et seulement si sa mesure intérieure est nulle puisque cela signifie que tous ses sous-ensembles mesurables sont de probabilité nulle, il est un peu moins immédiat de montrer qu'un ensemble est extérieurement négligeable si et seulement si sa mesure extérieure est nulle : en effet, dans ce cas, choisissons (axiome du choix dénombrable), pour tout n, $B_n \supset B$ mesurable de probabilité inférieure ou égale à $1/n$; l'intersection de ces B_n est alors mesurable de mesure nulle et contient B.

Théorème 12. — (i) *Une intersection quelconque d'ensembles intérieurement négligeable est intérieurement négligeable.*

(ii) *Une intersection quelconque d'ensembles extérieurement négligeable est extérieurement négligeable.*

(iii) *Une réunion dénombrable d'ensembles extérieurement négligeable est extérieurement négligeable.*

Démonstration. — Laissée au lecteur. □

Lorsqu'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ modélise une expérience aléatoire, ses ensembles extérieurement négligeables sont dits ensembles presque impossibles. Les sous-ensembles de Ω qui sont de complémentaires extérieurement négligeables sont dits presque certains ou presque sûrs. Notons qu'un sous-ensemble presque certain de Ω contient un sous-ensemble mesurable de probabilité 1.

4.2. Complétion

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Nous ne nous intéresserons qu'aux ensembles extérieurement négligeables, et notons leur ensemble $\mathcal{N}(\mathcal{A}, \mathbb{P})$ et $\bar{\mathcal{A}} = \sigma(\mathcal{A}, \mathcal{N}(\mathcal{A}, \mathbb{P}))$ qui est appelée la tribu complétée de $\mathcal{A}$ par rapport à la mesure de probabilité $\mathbb{P}$.

Théorème 13. — *Une partie $A \subset \Omega$ appartient $\bar{\mathcal{A}}$ si et seulement si il existe A^- et $A^+ \in \mathcal{A}$ tels que $A^- \subset A \subset A^+$ et $\mathbb{P}(A^+ \backslash A^-)$,* i.e., $\mathbb{P}(A^-) = \mathbb{P}(A^+)$.

Démonstration. — Notons $\widehat{\mathcal{A}}$ la collection des parties de Ω vérifiant la condition précédente : $A \in \widehat{\mathcal{A}}$ si et seulement si $A = A' \cup N$ avec $A^- \in \mathcal{A}$ et N négligeable. On constate immédiatement que $A \in \bar{\mathcal{A}}$ et donc que $\widehat{\mathcal{A}} \subset \bar{\mathcal{A}}$. Pour montrer l'inclusion réciproque, il suffit de vérifier que $\widehat{\mathcal{A}}$ est une tribu puisque $\mathcal{A} \subset \widehat{\mathcal{A}}$ et $\mathcal{N}(\mathcal{A}, \mathbb{P}) \subset \widehat{\mathcal{A}}$:

(i) $\Omega, \varnothing \in \mathcal{A} \subset \widehat{\mathcal{A}}$;

(ii) Si $A \in \widehat{\mathcal{A}}$, $A', A'' \in \mathcal{A}$, $A' \subset A \subset A''$ et $\mathbb{P}(A'' \backslash A') = 0$, alors $(A'')^c \subset A^c \subset (A')^c$ et de plus $(A')^c \backslash (A'')^c = A'' \backslash A'$, ce qui montre que $A^c \in \widehat{\mathcal{A}}$;

(iii) Si $(A_n)_n \subset \widehat{\mathcal{A}}$, $A_n = A'_n \cup N_n$ union d'un élément de $\mathcal{A}$ et d'un négligeable, $\bigcup_n A_n = \left(\bigcup_n A'_n\right) \cup \left(\bigcup_n N_n\right)$ est encore la réunion d'un élément de $\mathcal{A}$ et d'un négligeable.

D'où la conclusion. □

Théorème 14. — *Il existe une unique mesure de probabilité $\bar{\mathbb{P}}$ sur $(\Omega, \bar{\mathcal{A}})$ prolongeant $\mathbb{P}$.*

Démonstration. — Si $A \in \bar{\mathcal{A}}$, $A', A'' \in \mathcal{A}$, $A' \subset A \subset A''$ et $\mathbb{P}(A'' \backslash A') = 0$, nous posons $\bar{\mathbb{P}}(A) = \mathbb{P}(A') = \mathbb{P}(A'')$.

Dans un premier temps, il nous faut vérifier que $\bar{\mathbb{P}}$ est bien définie, à savoir que $\bar{\mathbb{P}}(A)$ ne dépend pas du couple $(A', A'') \in \mathcal{A}^2$ choisi. Soient A^* et A^{**} satisfaisant les mêmes propriétés. On a $A' \subset A' \cup A^* \subset A''$, donc $\mathbb{P}(A') \leqslant \mathbb{P}(A' \cup A^*) \leqslant \mathbb{P}(A'')$ et comme $\mathbb{P}(A') = \mathbb{P}(A'')$, il y a donc égalité $\mathbb{P}(A') = \mathbb{P}(A' \cup A^*) = \mathbb{P}(A'')$. De même, $\mathbb{P}(A^*) = \mathbb{P}(A' \cup A^*) = \mathbb{P}(A^{**})$. Toutes ces probabilités sont donc égales et $\bar{\mathbb{P}}(A)$ est bien défini.

Dans un deuxième temps, il faut vérifier que $\bar{\mathbb{P}}$ est une mesure de probabilité. Évidemment, $\bar{\mathbb{P}}(\Omega) = \mathbb{P}(\Omega) = 1$. Soit $(A_n)_n \subset \bar{\mathcal{A}}$ disjoints. Pour chaque n, $A_n = A'_n \cup N_n$ où $A'_n \in \mathcal{A}$ et $N_n \in \mathcal{N}(\mathcal{A}, \mathbb{P})$. Les $(A'_n)_n$ sont disjoints, et $\bigcup_n N_n$ est extérieurement négligeable donc inclus dans un certain $N \in \mathcal{A}$ de probabilité nulle : $\mathbb{P}(N) = 0$. On a $\bigcup_n A'_n \subset \bigcup_n A_n = \left(\bigcup_n A'_n\right) \cup \left(\bigcup_n N_n\right) \subset \left(\bigcup_n A'_n\right) \cup N$, et aussi $\sum_n \mathbb{P}(A'_n) = \mathbb{P}\left(\bigcup_n A'_n\right) \leqslant \mathbb{P}\left(\left(\bigcup_n A'_n\right) \cup N\right) \leqslant \mathbb{P}\left(\bigcup_n A'_n\right) + \mathbb{P}(N) = \mathbb{P}\left(\bigcup_n A'_n\right) =$

$\sum_n \mathbb{P}(A'_n)$. Comme $\bar{\mathbb{P}}(A_n) = \mathbb{P}(A'_n)$ pour tout n, on en déduit que $\mathbb{P}(\bigcup_n A_n) = \sum_n \mathbb{P}(A_n)$. Ainsi $\bar{\mathbb{P}}$ est bien une mesure de probabilité sur $(\Omega, \bar{\mathcal{A}})$.

Enfin, il nous reste à vérifier l'unicité de ce prolongement : si $\widehat{\mathbb{P}}$ est une autre mesure de probabilité prolongeant $\mathbb{P}$ et si $A \in \bar{\mathcal{A}}$, alors pour tous A', $A'' \in \mathcal{A}$, $A' \subset A \subset A''$ et $\mathbb{P}(A') = \mathbb{P}(A'')$; on a $\mathbb{P}(A') = \widehat{\mathbb{P}}(A') \leqslant \widehat{\mathbb{P}}(A) \leqslant \widehat{\mathbb{P}}(A'') = \mathbb{P}(A'')$, il y a donc égalité ; ainsi, par définition de $\bar{\mathbb{P}}$, $\widehat{\mathbb{P}}(A) = \bar{\mathbb{P}}(A)$. □

5. Séparabilité, atomes, séparation

5.1. Séparabilité

Définition 10. — Une tribu $\mathcal{B}$ sur un ensemble Ω est *séparable* s'il existe une suite $(B_n)_n$ au plus dénombrable d'éléments de $\mathcal{B}$ telle que $\mathcal{B}$ est égale à la tribu engendrée par $(B_n)_n$.

Précisons aussi qu'une tribu séparable peut être extrèmement riche : la tribu borélienne de $\mathbb{R}$ est séparable comme c'est aussi le cas des tribus boréliennes des espaces topologiques métrisables et séparables. À l'opposé, les tribus engendrées par des partitions au plus dénombrables sont évidemment séparables. Notons qu'une tribu séparable a au plus la puissance du continu.

Théorème 15. — *La tribu borélienne d'un espace topologique métrisable séparable est séparable.*

Démonstration. — □

Théorème 16. — *Un produit dénombrable de tribus est séparable si et seulement si chaque tribu est séparable.*

Démonstration. — □

Théorème 17. — *Une tribu $\mathcal{B}$ sur un ensemble Ω est séparable si et seulement si il existe une application $X : \Omega \to \mathbb{R}$ telle que $\mathcal{B} = \sigma(X)$.*

Remarque. — Dans cet énoncé il est implicite que $\sigma(X)$ est la tribu engendrée par l'application X avec pour espace image $\mathbb{R}$ muni de sa tribu borélienne. C'est donc l'image réciproque de cette dernière tribu par l'application X.

Démonstration. — Supposons $\mathcal{B}$ séparable. Soient $(B_n)_n \subset \mathcal{B}$ une suite engendrant $\mathcal{B}$, où les n sont des entiers supérieurs ou égaux à 1, et

$$X = \sum_n \frac{1}{3^n} \mathbb{1}_{B_n}$$

(*voir* [1], p. 15). Nous constatons que X est $\mathcal{B}$-mesurable, et donc que $\sigma(X) \subset \mathcal{B}$. De plus, pour tout n, B_n est l'image réciproque par X de l'ensemble de réels de $[0, 1/2]$ dont le n-ième terme de l'écriture en base 3 est 1, et alors $\mathcal{B} = \sigma((B_n)_n) \subset \sigma(X)$. Par conséquent, $\mathcal{B} = \sigma(X)$.

Réciproquement, soient $X : \Omega \to \mathbb{R}$ et $\mathcal{B} = \sigma(X)$. La tribu borélienne de $\mathbb{R}$ étant séparable, soit $(C_n)_n$ une suite d'ensembles boréliens l'engendrant. Posons $B_n = X^{-1}(C_n) \in \mathcal{B}$. Il est alors clair que $\mathcal{B} = \sigma((B_n)_n)$. □

5.2. Atomes

Définition 11. — Soit $(E, \mathcal{E})$ un espace mesurable.

(i) Un *atome* de la tribu $\mathcal{E}$ est une classe d'équivalence pour la relation $\mathcal{R}_{\mathcal{E}}$ définie sur $E \times E$ par $x \mathcal{R}_{\mathcal{E}} y$ si et seulement si tout $B \in \mathcal{E}$ contenant $\{x\}$ contient $\{y\}$.

(ii) La tribu $\mathcal{E}$ est *séparée*, ou l'espace mesurable est *séparé*, si et seulement si les atomes de la tribu $\mathcal{E}$ sont réduits aux singletons.

Vérifions qu'il s'agit d'une relation d'équivalence : elle est réflexive car clairement $x \mathcal{R}_{\mathcal{E}} x$ pour tout $x \in E$; elle est symétrique car si $x \mathcal{R}_{\mathcal{E}} y$ alors, si $B \in \mathcal{E}$ contient $\{y\}$, ou bien x appartient à B, ou bien x appartient à B^c, mais si $x \in B^c$ et comme $y \notin B^c$ on n'aurait pas $x \mathcal{R}_{\mathcal{E}} y$, ainsi $x \in B$, et $y \mathcal{R}_{\mathcal{E}} x$; elle est transitive car si $x \mathcal{R}_{\mathcal{E}} y$ et $y \mathcal{R}_{\mathcal{E}} z$, alors tout $B \in \mathcal{E}$ contenant $\{x\}$ contient $\{y\}$ et contient alors $\{z\}$ et donc $x \mathcal{R}_{\mathcal{E}} z$.

On déduit du fait que $\mathcal{R}_{\mathcal{E}}$ est une relation d'équivalence que les atomes forment une partition de E. Plus généralement, si $(B_i)_{i \in I}$ est une collection quelconque de parties d'un ensemble E, nous notons $\mathcal{R}_{(B_i)_{i \in I}}$ la relation correspondant à cette collection qu'on constate, de la même manière que précédemment, être une relation d'équivalence.

Théorème 18. — *Soit $(E, \mathcal{E})$ un espace mesurable. Alors pour tous $B \in \mathcal{E}$, $x \in B$, $\dot{x} \in B$, où $\dot{x}$ est la classe d'équivalence de x pour la relation $\mathcal{R}_{\mathcal{E}}$. En particulier, tout ensemble mesurable est la réunion, disjointe, des atomes qu'il contient.*

Démonstration. — Soient $B \in \mathcal{E}$ et $x \in B$. Si $y \notin B$, alors $y \notin \dot{x}$ puisque B sépare x de y. Par contraposée, on en déduit que $\dot{x} \subset B$. Le reste s'ensuit puisque les atomes forment une partition de E. □

Théorème 19. — *Soient $(B_i)_{i \in I}$ une collection quelconque de parties d'un ensemble E et $\mathcal{E} = \sigma((B_i)_{i \in I})$. Alors, pour tous $x, y \in E$, $x \mathcal{R}_{\mathcal{E}} y$ si et seulement si $x \mathcal{R}_{(B_i)_{i \in I}} y$. Autrement dit, deux points de E ne sont pas séparables par les éléments de la tribu $\mathcal{E}$ si et seulement si ils ne sont pas séparables par une famille de parties qui engendre cette tribu.*

Démonstration. — Le sens direct est évident puisque $(B_i)_{i \in I} \subset \mathcal{E}$. Pour la réciproque, supposons que $x \mathcal{R}_{(B_i)_{i \in I}} y$ et posons $\bar{E} = E/\{x = y\}$, ce qui consiste à identifier les points x et y en laissant les autres points inchangés. Nous notons $\pi : E \to \bar{E}$ la projection canonique, $\bar{B}_i = \pi(B_i)$ pour tout $i \in I$, $\bar{\mathcal{E}} = \sigma((\bar{B}_i)_{i \in I})$, et $\bar{x} = \pi(x) = \pi(y)$. On constate que par construction et du fait que x et y ne soient pas séparés par $(B_i)_{i \in I}$, pour tout $i \in I$, $\pi^{-1}(\bar{B}_i) = B_i$. Ainsi la tribu $\pi^{-1}(\bar{\mathcal{E}})$ est une tribu sur E contenant $(B_i)_{i \in I}$. On a donc $\mathcal{E} \subset \pi^{-1}(\bar{\mathcal{E}})$ (en fait il y a égalité). Mais par construction $\pi^{-1}(\bar{\mathcal{E}})$ ne sépare pas x et y, il en est donc de même de la tribu $\mathcal{E}$. □

Théorème 20. — *Soient $(B_i)_{i \in I}$ une collection quelconque de parties d'un ensemble E et $\mathcal{E} = \sigma((B_i)_{i \in I})$. Alors pour tout $x \in E$, l'atome $\dot{x}$ de la tribu $\mathcal{E}$ contenant x s'écrit sous la forme*

$$\dot{x} = \bigcap_{i \in I} B_i^{\varepsilon_i(x)},$$

où $B_i^{\varepsilon_i(x)} = B_i$ si $x \in B_i$, et $B_i^{\varepsilon_i(x)} = B_i^c$ si $x \notin B_i$.

Démonstration. — D'après le théorème précédent, on a $y \in \dot{x}$ si et seulement si, pour tout $i \in I$, ou bien x, $y \in B_i$, ou bien x, $y \in B_i^c$. L'écriture de $\dot{x}$ en résulte. □

THÉORÈME 21. — *Si un espace mesurable $(E, \mathcal{E})$ est séparable, alors ses atomes sont mesurables.*

Démonstration. — Cet énoncé est une conséquence immédiate du théorème précédent. □

Remarques. — *a*) Les atomes d'une tribu ne sont généralement pas mesurables et ne sont donc pas dans la tribu considérée. Un exemple de cela est la tribu produit de $\{0,1\}^{\mathbb{R}_+}$ qui est séparée mais à laquelle n'appartiennent pas les singletons.

b) On trouve dans la littérature (*voir* [3]) une autre définition des atomes d'une tribu, atomes que nous nommerons *atomes mesurables* : soit $(E, \mathcal{E})$ un ensemble mesurable ; un sous-ensemble mesurable $A \in \mathcal{E}\backslash\{\varnothing\}$ est un atome mesurable si et seulement si $B \in \mathcal{E}$, $B \subset A$ implique $B = \varnothing$ ou $B = A$.

Nous allons montrer que les atomes mesurables sont les vrais atomes, mesurables. Soit $x \in E$ et supposons $\dot{x} \in \mathcal{E}$; si $B \in \mathcal{E}$, $B \subset \dot{x}$ alors nécessairement ou bien $B = \varnothing$, ou bien $B = \dot{x}$ sans quoi B séparerait des points de $\dot{x}$; ainsi $\dot{x}$ est un atome mesurable. Réciproquement, si $A \in \mathcal{E}$ est tel qu'il existe x, $y \in A$, $B \in \mathcal{E}$ tels que $x \in B$ et $y \in B^c$, alors $\varnothing \neq A \cap B \subsetneq A$, et donc A n'est pas un atome mesurable. Par contraposée deux points d'un atome mesurable A ne sont pas séparables et donc si $x \in A$, $\dot{x} \supset A$. Comme, par ailleurs, $\dot{x}$ est inclus dans tout ensemble mesurable contenant x, on a l'égalité $\dot{x} = A$.

Les atomes d'une tribu se séparent donc en deux catégories : les atomes mesurables et les autres. Dans [3], une tribu $\mathcal{E}$ est dite atomique si tout les éléments de $\mathcal{E}$ sont la réunion des atomes mesurables qu'ils contiennent, ce qui revient à demander que tous les atomes soient mesurables.

5.3. Mesures atomiques

5.4. Séparation

À tout espace mesurable $(E, \mathcal{E})$ est associé un espace mesurable $(\dot{E}, \dot{\mathcal{E}})$ séparé : $\dot{E}$ est l'ensemble quotient de E par la relation $\mathcal{R}_{\mathcal{E}}$, et $\dot{\mathcal{E}}$ est l'image de $\mathcal{E}$ par la surjection canonique π de E dans $\dot{E}$. Si on identifie $\dot{x}$ pour $x \in E$ à la partie de E correspondante, alors pour $B \in \mathcal{E}$, $\pi(B) = \bigcup_{x \in B} \dot{x}$. Il est clair que $B \subset \pi(B)$, mais réciproquement, pour tout $x \in B$, $\dot{x} \subset B$ et ainsi $\pi(B) \subset B$, d'où l'égalité. On constate qu'avec cette identification $\dot{\mathcal{E}}$ est égale à $\mathcal{E}$ et est donc une tribu. Il est de plus évident que la surjection canonique π est mesurable.

Si $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{E})$, alors X induit la variable aléatoire $\dot{X} : (\Omega, \mathcal{A}, \mathbb{P}) \to (\dot{E}, \dot{\mathcal{E}})$ de manière naturelle : si π est la surjection canonique de E dans $\dot{E}$, on a $\dot{X} = \pi \circ X$ qui est mesurable. De plus, du fait que les tribus $\mathcal{E}$ et $\dot{\mathcal{E}}$ sont isomorphes, les tribus $\sigma(X)$ et $\sigma(\dot{X})$ sont égales et on peut identifier les lois de X et de $\dot{X}$. C'est pourquoi on ne perd pas réellement de généralité en Théorie des Probabilités à supposer les variables aléatoires à valeurs dans des espaces mesurables séparés.

Pour être complet, il faut encore montrer qu'on peut passer au quotient sur l'espace de départ après être passé au quotient sur l'espace d'arrivée. Soit $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ où nous supposons, quitte à passer au quotient, $(F, \mathcal{F})$ séparé. Si x, $y \in E$ sont tels que $f(x) \neq f(y)$, par séparation de $(F, \mathcal{F})$ soit $C \in \mathcal{F}$ tel que

$f(x) \in C$ et $f(y) \in C^c$. On a ainsi $x \in f^{-1}(C)$ et $y \in f^{-1}(C^c)$. Les points x et y sont donc séparés l'un de l'autre dans $(E, \mathcal{E})$. Par contraposée, nous en déduisons que f est constante sur les atomes de $(E, \mathcal{E})$. Ainsi f induit une unique application $\bar{f} : (\dot{E}, \dot{\mathcal{E}}) \to (F, \mathcal{F})$ qui est bien évidemment mesurable puisque $\dot{\mathcal{E}}$ est identique à $\mathcal{E}$. La propriété universelle satisfaite par $\bar{f}$ est qu'alors pour tout $X : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$ on a

$$f \circ X = \bar{f} \circ \pi_E \circ X = \bar{f} \circ \dot{X}$$

où π_E est la projection canonique de $(E, \mathcal{E})$ dans $(\dot{E}, \dot{\mathcal{E}})$. Si $(F, \mathcal{F})$ n'est plus supposé séparé, on a alors

$$\widehat{f \circ X} = \pi_F \circ f \circ X = \dot{f} \circ X = \bar{\dot{f}} \circ \pi_E \circ X = \bar{\dot{f}} \circ \dot{X},$$

où π_F est la projection canonique de $(F, \mathcal{F})$ dans $(\dot{F}, \dot{\mathcal{F}})$, ce qui montre que du point de vue probabiliste tous les espaces mesurables — hormis peut-être $(\Omega, \mathcal{A})$ — peuvent être supposés séparés.

6. **Séparabilité essentielle**

Définition 12. — Une sous-tribu $\mathcal{B}$ sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ est *essentiellement séparable* s'il existe une suite $(B_n)_n$ au plus dénombrable d'éléments de $\mathcal{B}$ telle que

$$\sigma(B_n)_n \subset \mathcal{B} \subset \sigma(B_n)_n \vee \mathcal{N}(\mathcal{A}, \mathbb{P}).$$

Rappelons que $\mathcal{N}(\mathcal{A}, \mathbb{P})$ est l'ensemble des sous-ensembles $(\mathcal{A}, \mathbb{P})$-négligeables de Ω, c'est-à-dire les sous-ensembles $N \subset \Omega$ tel qu'il existe $A \in \mathcal{A}$, $N \subset A$, $\mathbb{P}(A) = 0$.

Théorème 22. — *Une tribu $\mathcal{B}$ est essentiellement séparable si et seulement si il existe une application $X : \Omega \to \mathbb{R}$ telle que $\sigma(X) \subset \mathcal{B} \subset \sigma(X) \vee \mathcal{N}(\mathbb{P})$).*

Démonstration. — Le cas de la séparabilité essentielle se déduit immédiatement du cas de la simple séparabilité. □

Théorème 23. — *Une tribu $\mathcal{B}$ est essentiellement séparable si et seulement si l'espace topologique $L^1(\Omega, \mathcal{B}, \mathbb{P})$ est séparable.*

Démonstration. — Supposons la tribu $\mathcal{B}$ essentiellement séparable, essentiellement engendrée par $(B_n)_n \subset \mathcal{B}$. Posons $\mathcal{B}_n = \sigma\{B_k : k \leqslant n\}$ qui est une tribu finie, V_n l'ensemble des variables aléatoires $\mathcal{B}_n$-mesurables à valeurs dans $\mathbb{Q}$ qui est un ensemble dénombrable, et V la réunion des ensembles $(V_n)_n$ qui est aussi dénombrable. Soit $X \in L^1(\Omega, \mathcal{B}, \mathbb{P})$. Choisissons pour tout n une variable aléatoire $X_n \in V_n$ uniformément (et donc aussi dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$) proche à $1/n$ près de $\mathbb{E}[X \,|\, \mathcal{B}_n]$ (calculée sans aucune ambiguïté sur les atomes de $\mathcal{B}_n$ de probabilité strictement positive et égale à 0, par exemple, sur les autres). Puisque, par convergence de martingales fermées, dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$, la suite $(\mathbb{E}[X \,|\, \mathcal{B}_n])_n$ converge vers $\mathbb{E}[X \,|\, \sigma((B_n)_n)] = X$ (p.s.), il en est de même de la suite $(X_n)_n$. Ainsi V est un sous-ensemble dense de $L^1(\Omega, \mathcal{B}, \mathbb{P})$ qui est donc séparable.

Réciproquement, supposons que $L^1(\Omega, \mathcal{B}, \mathbb{P})$ soit séparable, et soit $(X_n)_n$ une suite de variables aléatoires dense dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$. Chaque tribu $\sigma(X_n)$ est séparable et donc engendrée par une suite au plus dénombrable d'événements. Alors

$\sigma((X_n)_n)$ est engendrée par la réunion de ces événements qui sont en nombre dénombrable, cette dernière tribu est donc séparable. Comme

$$(X_n)_n \subset L^1\big(\Omega, \sigma\big((X_n)_n\big), \mathbb{P}\big) \subset L^1(\Omega, \mathcal{B}, \mathbb{P})$$

et par densité de la suite $(X_n)_n$, on a égalité des deux espaces. Ainsi, $\sigma((X_n)_n)$ et $\mathcal{B}$ sont essentiellement égales, d'où l'essentielle séparabilité de $\mathcal{B}$. □

Remarques. — *a*) On peut remplacer dans l'énoncé précédent l'espace L^1 par les espaces L^p, $p \in [1, \infty[$.

b) Ces espaces L^p ne permettant pas de distinguer deux tribus égales à des négligeables près, on ne peut pas remplacer dans l'énoncé précédent « essentiellement séparable » par « séparable ».

THÉORÈME 24. — *Une sous-tribu $\mathcal{B}$ d'une tribu $\mathcal{A}$ essentiellement séparable est essentiellement séparable.*

Démonstration. — Notons $\psi = \mathbb{E}[\,.\,|\,\mathcal{B}] : L^1(\Omega, \mathcal{A}, \mathbb{P}) \to L^1(\Omega, \mathcal{B}, \mathbb{P})$. Puisque $L^1(\Omega, \mathcal{A}, \mathbb{P})$ est séparable, soit $(f_n)_n$ une suite dense dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$. Soit $g \in L^1(\Omega, \mathcal{B}, \mathbb{P})$. Puisque $g \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, il existe un changement d'indices j tel que $(f_{j(n)})_n$ tend vers g dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$. Puisque ψ est continue, alors $(\psi(f_{j(n)}))_n$ tend vers $\psi(g) = g$ dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$. Ceci montre que la suite $(g_n)_n = (\psi(f_n))_n$ est dense dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$ qui est donc séparable. □

Remarques. — *a*) La preuve précédente est un cas particulier du fait général suivant : si E et F sont deux espaces topologiques, E séparable, alors, s'il existe $\psi : E \to F$ continu surjectif, F est séparable.

b) Il faut se garder de croire qu'une sous-tribu d'une tribu séparable puisse être automatiquement séparable. La remarque suivant l'un des exercices proposés ci-après le montre clairement.

Exercices

Exercice 1. — Soient Ω un ensemble infini non dénombrable ($\mathbb{R}$ par exemple), et $\mathcal{A}$ l'ensemble des parties de Ω qui sont au plus dénombrables ou de complémentaires au plus dénombrables, c'est-à-dire muni de sa tribu ponctuelle.

(i) Montrer que $\mathcal{A}$ est une tribu sur Ω.

(ii) Montrer que $\mathcal{A}$ est la plus petite tribu sur Ω contenant les singletons, c'est-à-dire que si $\mathcal{B}$ est une tribu sur Ω telle que $\{\omega\} \in \mathcal{B}$ pour tout $\omega \in \Omega$, alors $\mathcal{A} \subset \mathcal{B}$.

(iii) Soit $\mathbb{P}_0 : \mathcal{A} \to [0, 1]$ définie par : pour $A \in \mathcal{A}$, $\mathbb{P}_0(A) = 0$ si A est dénombrable et $\mathbb{P}_0(A) = 1$ sinon. Montrer que $\mathbb{P}_0$ est une mesure de probabilité sur $(\Omega, \mathcal{A})$.

Si $\mathbb{P}$ est une mesure de probabilité sur $(\Omega, \mathcal{A})$, on note $\mathrm{supp}(\mathbb{P}_{\mathrm{disc}})$ le support de la partie discrète de $\mathbb{P}$ qui est

$$\mathrm{supp}(\mathbb{P}_{\mathrm{disc}}) = \big\{\omega \in \Omega : \mathbb{P}\{\omega\} > 0\big\},$$

(on notera que les singletons sont dans $\mathcal{A}$ par hypothèse) et $\mathbb{P}_{\mathrm{disc}}$ l'application

$$\begin{aligned} \mathcal{A} &\longrightarrow [0, 1] \\ A &\longmapsto \mathbb{P}_{\mathrm{disc}}(A) = \sum_{\omega \in \mathrm{supp}(\mathbb{P}_{\mathrm{disc}}) \cap A} \mathbb{P}\{\omega\} \end{aligned}$$

qui est la partie discrète de $\mathbb{P}$. On admettra que le support de la partie discrète d'une mesure de probabilité $\mathbb{P}$ est un ensemble fini ou dénombrable. Enfin, on définit la partie diffuse $\mathbb{P}_{\text{diff}}$ de $\mathbb{P}$ par

$$\mathbb{P}_{\text{diff}}(A) = \mathbb{P}(A) - \mathbb{P}_{\text{disc}}(A), \qquad A \in \mathcal{A}.$$

(iv) Montrer que, pour tout $A \in \mathcal{A}$, $\mathbb{P}_{\text{disc}}(A) = \mathbb{P}(\text{supp}(\mathbb{P}_{\text{disc}}) \cap A)$. En déduire que $\mathbb{P}_{\text{disc}}$ est une sous-probabilité, c'est-à-dire qu'elle est dénombrablement additive mais $0 \leqslant \mathbb{P}_{\text{disc}}(\Omega) \leqslant 1$ (sa mesure totale de Ω est inférieure ou égale à 1).

(v) Montrer que $\mathbb{P}_{\text{diff}}$ est une sous-probabilité (on précisera notamment pourquoi $\mathbb{P}_{\text{diff}}(A) \in [0,1]$ pour tout $A \in \mathcal{A}$).

(vi) Montrer que le support de la partie discrète de $\mathbb{P}_{\text{diff}}$ est vide, *i.e.*, que $\mathbb{P}_{\text{diff}}$ est sans partie discrète. En déduire que $\mathbb{P}_{\text{diff}}(A) = 0$ pour tout A fini ou dénombrable.

(vii) Justifier que, pour tout $A \in \mathcal{A}$, $\mathbb{P}_{\text{diff}}(A^c) = \mathbb{P}_{\text{diff}}(\Omega) - \mathbb{P}_{\text{diff}}(A)$. En déduire que $\mathbb{P}_{\text{diff}} = t\,\mathbb{P}_0$ pour une constante réelle t qu'on précisera.

(viii) Déduire des questions précédentes que toute mesure de probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{A})$ distincte de $\mathbb{P}_0$ est un barycentre d'une unique mesure de probabilité discrète sur $(\Omega, \mathcal{A})$ et de $\mathbb{P}_0$.

Remarques. — *a*) Si dans l'exercice précédent $\Omega = \mathbb{R}$, alors $\mathcal{A}$ est une sous-tribu de $\mathcal{B}(\mathbb{R})$ (cette dernière tribu est séparable, *voir* point suivant). La décomposition des mesures de probabilités sur $(\mathbb{R}, \mathcal{A})$ s'apparente à la décomposition de Lebesgue (en beaucoup plus simple).

b) La tribu $\mathcal{A}$ de l'exercice précédent n'est pas séparable. Soit $(B_n)_n \subset \mathcal{A}$. Puisque nous nous intéressons à $\mathcal{B} = \sigma(B_n)$, alors quitte à remplacer pour certains indices B_n par B_n^c, ont peut supposer que chaque B_n est au plus dénombrable. Tout atome A de $\mathcal{B}$ s'écrit

$$A = \bigcap_{n\in I} B_n \cap \bigcap_{n\in J} B_n^c. \qquad \text{et posons} \qquad \Omega' = \bigcap_n B_n^c,$$

qui est un ensemble de complémentaire au plus dénombrable. La trace $\mathcal{B}'$ de $\mathcal{B}$ sur Ω' a pour atomes $A' = \bigcap_{n\in I} B_n \cap \Omega' \subset \bigcup_n B_n \cap \Omega'$ ensemble fixe qui est dénombrable. Ainsi, $\mathcal{B}'$ a au plus un nombre dénombrable d'atomes alors que l'ensemble des atomes de la trace $\mathcal{A}'$ de $\mathcal{A}$ sur Ω' a même cardinal que Ω'.

c) Considérons une mesure de probabilité sur une sur-tribu de $\mathcal{A}$. Elle induit sur $\mathcal{A}$ une mesure de probabilité dont nous connaissons la décomposition. On constate alors que la tribu $\mathcal{A}$ est essentiellement engendrée par les points du support de la partie discrète de cette mesure de probabilité.

Appendice : quelques mots de topologie

Définition 13. — Un *espace topologique* est un couple $(E, \mathcal{O})$ où :

(i) E est un ensemble ;

(ii) $\mathcal{O}$ est une *topologie* sur E, c'est-à-dire une toute famille de parties de E vérifiant

(O_1) $\varnothing \in \mathcal{O}$ et $E \in \mathcal{O}$;

(O_2) $\forall\, (A, B) \in \mathcal{O}^2$, $A \cap B \in \mathcal{O}$;

(O_3) $\forall\, I,\ \forall\, (A_i)_{i\in I},\ \bigcup_{i\in I} A_i \in \mathcal{O}$.

Les éléments de $\mathcal{O}$ sont appelés les *ouverts* de la topologie $\mathcal{O}$.

Exemple. — Soit E un ensemble. Notons $\mathcal{F} = \{B \subset E : \operatorname{Card} B < \infty \text{ ou } B = E\}$. On a : $\varnothing \in \mathcal{F}$, $E \in \mathcal{F}$; si A, $B \in \mathcal{F}$, $A \cup B \in \mathcal{F}$; si $(A_i)_{i\in I} \subset \mathcal{F}$, alors $\bigcap_{i\in I} A_i \in \mathcal{F}$. Ceci montre que $\mathcal{F}$ est l'ensemble des fermés d'une topologie sur E, et ainsi $\mathcal{O} = \{B \subset E : \operatorname{Card} B^c < \infty \text{ ou } B = \varnothing\}$ est une topologie sur E, topologie que nous avons qualifiée de ponctuelle.

Tous les fermés de cette topologie sont compacts : soit $F \in \mathcal{F}$ et $(U_i)_{i\in I} \subset \mathcal{O}$ un recouvrement de F par des ouverts non vides ; si F est fini il est alors clair qu'on peut extraire de $(U_i)_{i\in I} \subset \mathcal{O}$ une famille finie recouvrant F ; si F est infini, auquel cas nécessairement $F = E$, si l'un des $(U_i)_{i\in I}$ est égal à E on choisit ce dernier pour recouvrir E, sinon, soit U_{i_0} l'un de ces ouverts, on a $U_{i_0} = E\backslash F_0$ où F_0 est un ensemble fini qui est recouvert par $(U_i)_{i\in I\backslash\{i_0\}}$, on extrait de cette dernière famille un recouvrement fini de F_0, alors U_{i_0} et ce dernier recouvrement est un recouvrement fini de E, ce qui montre que E est compact.

Si E est un ensemble fini, cette topologie n'est rien d'autre que la topologie discrète pour laquelle toute partie est ouverte et fermée.

Nous supposerons pour la suite que E est un ensemble infini. La première conséquence de cette hypothèse est que quelques soient A, $B \in \mathcal{O}\backslash\{\varnothing\}$, $A \cap B \neq \varnothing$, c'est-à-dire que tout couple d'ouverts non vides ont une intersection non vide, ou encore que deux fermés différents de E ne peuvent recouvrir E. En particulier $(E, \mathcal{O})$ est alors un espace topologique connexe — alors que lorsque E est fini il est totalement discontinu.

Toujours lorsque E est infini, l'espace topologique $(E, \mathcal{O})$ satisfait peu d'axiomes de séparation :

T_0 : $\forall\, x,\ y \in E,\ x \neq y,\ \exists\, U \in \mathcal{O},\ x \in U,\ x \notin U$ (prendre $U = E\backslash\{y\}$;

T_1 : $\forall\, x,\ y \in E,\ x \neq y,\ \exists\, U_x,\ U_y \in \mathcal{O},\ x \in U_x,\ y \in U_y,\ U_x \not\subset U_y$ et $U_y \not\subset U_x$ (prendre $U_x = E\backslash\{y\}$ et $U_y = E\backslash\{x\}$).

En revanche les axiomes de séparation suivants ne sont pas satisfaits :

T_2 : $\forall\, x,\ y \in E,\ x \neq y,\ \exists\, U_x,\ U_y \in \mathcal{O},\ x \in U_x,\ y \in U_y,\ U_x \cap U_y = \varnothing$ (puisque deux ouverts non vides ne sont pas disjoints).

Références

[1] Dellacherie (C.) et Meyer (P.-A.), *Probabilités et Potentiel*, Chapitres I à IV, Hermann (1975).

[2] Hocking (J. G.) et Young (G. S.), *Topology*, Dover (1988).

[3] Nadkarni (M. G.), Ramachandran (D.), Bhaskara Rao (K. P. S.), « On sequences of σ-algebras », *Colloquium Mathematicum*, vol. xxxii (1975), p. 219–226.

Table des matières

1. Espaces mesurables, probabilisés 1

1.1. *Espaces mesurables* 1

1.2. *Tribus trace* 1

1.3. *Espaces probabilisés* . . . 2
1.4. *Probabilités conditionnelles* . . . 3
2. Applications mesurables, variables aléatoires . . . 3
2.1. *Applications mesurables* . . . 3
2.2. *Restrictions des espaces de départ et d'arrivée* . . . 4
2.3. *Variables aléatoires* . . . 4
2.4. *Restrictions des espaces de départ et d'arrivée* . . . 5
3. Génération de tribus . . . 5
3.1. *Généralités* . . . 5
3.2. *Tribus boréliennes* . . . 7
3.3. *Tribus produit* . . . 7
3.4. *Tribus engendrées par des applications* . . . 8
4. Ensembles négligeables, complétion . . . 9
4.1. *Ensembles négligeables* . . . 9
4.2. *Complétion* . . . 10
5. Séparabilité, atomes, séparation . . . 11
5.1. *Séparabilité* . . . 11
5.2. *Atomes* . . . 12
5.3. *Mesures atomiques* . . . 13
5.4. *Séparation* . . . 13
6. Séparabilité essentielle . . . 14
Exercices . . . 15
Appendice : quelques mots de topologie . . . 16
Références . . . 17