Jean-Claude Laleuf

Licence
Master

Bases mathématiques du calcul stochastique

ellipses

Chapitre 1

Espaces mesurables

On présente les définitions générales du calcul des probabilités et du calcul stochastique : tribus, espaces mesurables, espaces de probabilités, applications mesurables et processus stochastiques.

On établit le théorème de décomposition d'une mesure signée μ en la différence de deux mesures positives. On construit pour cela d'abord la décomposition de Hahn de l'espace Ω en deux parties mesurables disjointes N *négative* et P *positive* telles que pour toutes sous parties mesurables N' de N on a $\mu(N') \leq 0$ et pour toutes sous parties mesurables P' de P on a $\mu(P') \geq 0$. On obtient alors la décomposition de Jordan de $\mu = \mu^+ - \mu^-$ en posant $\mu^+(A) = \mu(A \cap P)$ et $\mu^-(A) = -\mu(A \cap N)$.

On démontre ensuite les théorèmes de classe monotone. D'abord le lemme des classes monotones permet l'extension d'une propriété satisfaite dans une classe $\mathcal{S}$ d'événements stables par intersection finie (π - système) à la tribu $\sigma(\mathcal{S})$ engendrée par $\mathcal{S}$.
Le premier théorème de classe monotone établit que pour un π - système $\mathcal{S}$ sur un ensemble non vide Ω et un espace vectoriel $\mathcal{H}$ d'applications de Ω dans $\mathbb{R}$ qui contient les constantes, les indicatrices des éléments de $\mathcal{S}$ et les limites des suites dans $\mathcal{H}$ croissantes et bornées, l'espace $\mathcal{H}$ contient toutes les applications $\sigma(\mathcal{S})$ mesurables et bornées de Ω dans $\mathbb{R}$.
Le deuxième théorème de classe monotone est une généralisation ou la classe des indicatrices d'un π - système $\mathcal{S}$ est remplaçée par une partie $\mathcal{G}$ de $\mathcal{H}$ stable par produit.
Le troisième théorème de classe monotone est adapté aux besoins de la théorie de l'intégration stochastique du chapitre 11. Dans ce résultat $\mathcal{H}$ n'est pas à priori un espace vectoriel mais il contient une algèbre $\mathcal{G}$.

1.1 Espaces mesurables et probabilités

Dans une expérience aléatoire on note en général Ω l'ensemble des résultats ou *réalisations* possibles de cette expérience. Un *événement* est un ensemble de réalisations, donc un sous ensemble de Ω. Une *tribu* $\mathcal{F}$ sur Ω est un ensemble d'événements pouvant se produire au cours d'une expérience aléatoire.

Une tribu $\mathcal{F}$ sur Ω est donc un sous ensemble de l'ensemble de tous les sous ensembles de Ω (ensemble des parties).

1.1.1 Définitions : tribus, algèbres, espaces mesurables

Un ensemble $\mathcal{F}$ de parties d'un ensemble non vide Ω est une *tribu* sur Ω si :

1. $\emptyset \in \mathcal{F}$
2. $A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$ (stabilité par complémentaire)
3. $\forall n,\ A_n \in \mathcal{F} \Longrightarrow \bigcup_n A_n \in \mathcal{F}$ (stabilité par union dénombrable)

Si dans (3) l'union est finie l'ensemble $\mathcal{F}$ est une *algèbre* de parties de Ω. Un couple $(\Omega, \mathcal{F})$ constitué d'un ensemble non vide Ω et d'une tribu $\mathcal{F}$ sur Ω est un *espace mesurable*. Une partie A d'un espace mesurable $(\Omega, \mathcal{F})$ est dite *mesurable* si $A \in \mathcal{F}$. On réservera dans la suite le terme *événement* dans un espace mesurable pour désigner les parties mesurables de Ω.

Remarque
Notons que les axiomes (1) et (2) impliquent $\Omega \in \mathcal{F}$ et que les axiomes (2) et (3) impliquent :

$$\forall n,\ A_n \in \mathcal{F} \Longrightarrow \bigcap_n A_n \in \mathcal{F} \text{ (stabilité par intersection dénombrable)}$$

Exemples
Dans un ensemble quelconque non vide Ω l'exemple le plus simple de tribu est l'ensemble $\mathcal{F} = \{\emptyset, \Omega\}$. C'est la plus petite tribu (au sens de l'inclusion) sur Ω. La plus grande est l'ensemble $\mathcal{P}_\Omega$ de tous les sous ensembles de Ω.

1.1.2 Définitions : tribu engendrée, tribu borélienne

Si Γ est un ensemble de parties de Ω on appelle *tribu engendrée* par Γ et on note $\sigma(\Gamma)$ la plus petite tribu (au sens de l'inclusion) sur Ω contenant Γ. Comme toutes intersections de tribus est une tribu il existe toujours une plus petite tribu contenant Γ c'est l'intersection de toutes les tribus sur Ω qui contiennent Γ.
Si E est un espace topologique on appelle *tribu borélienne* de E et on note $\mathcal{B}(E)$ la tribu sur E engendrée par les ouverts de E. On appelle *borélien* un élément d'une tribu borélienne.

Exemples
Sur $\mathbb{R}$ la tribu borélienne $\mathcal{B}(\mathbb{R})$ est aussi engendrée par les intervalles. Notons qu'il n'est pas simple en général de montrer qu'une partie de $\mathbb{R}$ n'est pas un borélien. On trouvera un exemple construit à l'aide de l'axiome du choix et de la mesure de Lebesgue au §1.2.1.
Si A est une partie non vide de Ω la tribu $\sigma(A)$ engendrée par le singleton $\{A\}$ est $\{\emptyset, A, A^c, \Omega\}$. Si (A_i) est une partition dénombrable de Ω la tribu $\sigma(A_i \mid i \in \mathbb{N})$ engendrée par cette partition est appelée *tribu atomique*. Elle est constituée de $\emptyset$ et des unions dénombrables d'ensembles A_i.

1.1.3 Définition : probabilité, espace de probabilité

Une probabilité est une application P d'une tribu $\mathcal{F}$ dans $\mathbb{R}^+$ telle que :

1. $P(\emptyset) = 0$,
2. $P(\Omega) = 1$,
3. $\forall (A_n) \subseteq \mathcal{F}, \forall i \neq j \in \mathbb{N}, A_i \cap A_j = \emptyset \Longrightarrow P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$.

La probabilité d'une union dénombrable d'événements disjoints est la somme de leurs probabilités. C'est la propriété de σ - additivité. Un *espace de probabilité* est un triplet $(\Omega, \mathcal{F}, P)$ où $(\Omega, \mathcal{F})$ est un espace mesurable et P une probabilité sur $\mathcal{F}$.

1.1.4 Proposition : fonction d'ensembles sur une algèbre

Soit $\mathcal{A}$ une algèbre de parties de Ω (§1.1.1) et P une application de $\mathcal{A}$ dans $\mathbb{R}^+$ qui satisfait (1) et (2) de §1.1.3 et qui est aussi *additive* dans $\mathcal{A}$ c'est-à-dire :

$$\forall A, B \in \mathcal{A},\ P(A \cup B) + P(A \cap B) = P(A) + P(B)$$

Dans ces conditions P est une application croissante dans $\mathcal{A}$, c'est-à-dire :

$$\forall A, B \in \mathcal{A},\ A \subseteq B \Longrightarrow P(A) \leq P(B)$$

et pour tout $(A_n) \subseteq \mathcal{A}$ et $A \in \mathcal{A}$ les assertions suivantes sont équivalentes :

1. $A_n \uparrow A \Longrightarrow P(A_n) \uparrow P(A)$,
2. $A_n \downarrow A \Longrightarrow P(A_n) \downarrow P(A)$,
3. $A_n \downarrow \emptyset \Longrightarrow P(A_n) \downarrow 0$,
4. $\forall i \neq j,\ A_i \cap A_j = \emptyset,\ \bigcup_n A_n = A \Longrightarrow P(A) = \sum_n P(A_n)$.

La propriété (4) de P est appelée σ - *additivité* dans $\mathcal{A}$.

Démonstration

L'application P additive dans $\mathcal{A}$ est une application croissante dans $\mathcal{A}$ car :

$$A \subseteq B \implies B = A \cup (B - A) \implies P(B) = P(A) + P(B - A) \geq P(A)$$

Les assertions (1) et (2) sont équivalentes par passage aux complémentaires car l'additivité de P et l'axiome (2) de §1.1.3 entraîne $P(A^c) = P(\Omega) - P(A) = 1 - P(A)$. L'assertion (3) est un cas particulier de (2). Inversement (3) entraîne (2) car :

$$A_n \downarrow A \implies A_n - A \downarrow \emptyset \implies \lim_n P(A_n) - P(A) = \lim_n P(A_n - A) = 0$$

Les assertions (1), (2) et (3) de §1.1.4 sont donc équivalentes. Montrons que (1) entraîne (4). Soit (A_n) une partition de A dans $\mathcal{A}$ et $B_n = \cup_{k \leq n} A_k$ il vient :

$$A = \bigcup_n B_n \implies P(A) = \lim_n \uparrow P(B_n) = \lim_n \uparrow \sum_{k \leq n} P(A_k) = \sum_n P(A_n)$$

Inversement montrons que (4) entraîne (1). Soit (A_n) croissante dans $\mathcal{A}$ telle que $\cup_n A_n = A$ posons $B_0 = A_0$ et $B_{n+1} = A_{n+1} - A_n$ il vient :

$$P(A) = \sum_n P(B_n) = \lim_n \uparrow \sum_{k \leq n} P(B_k) = \lim_n \uparrow P(A_n)$$

La σ - additivité de P dans $\mathcal{A}$ est donc équivalente à la continuité croissante de P dans $\mathcal{A}$, ce qui achève la démonstration de §1.1.4. $\square$

1.1.5 Définition : probabilité sur une algèbre

Soit $\mathcal{A}$ une algèbre de partie d'un ensemble non vide Ω on dit qu'une application P de $\mathcal{A}$ dans $\mathbb{R}^+$ est une *probabilité sur l'algèbre* $\mathcal{A}$ si elle satisfait (1) et (2) de §1.1.3 et est σ - additive dans $\mathcal{A}$ c'est-à-dire vérifie la propriété (4) de §1.1.4.

Remarques

1) Il résulte des résultats de §1.1.4 que puisque toute tribu est une algèbre toute probabilité sur une tribu est croissante et vérifie les assertions (1), (2), (3) de §1.1.4.

2) On montre (théorème de Carathéodory §3.4.4) que toute probabilité sur une algèbre $\mathcal{A}$ à un prolongement unique en une probabilité sur la tribu $\sigma(\mathcal{A})$ engendrée par $\mathcal{A}$.

1.1.6 Lemme de majoration et de Borel Cantelli

Dans un espace de probabilité $(\Omega, \mathcal{F}, P)$ pour toutes suites d'événements $(A_n) \subseteq \mathcal{F}$ on a $P(\cup_n A_n) \leq \sum_n P(A_n)$ et si $\sum_n P(A_n) < \infty$ alors $P(\limsup_n A_n) = 0$ avec $\limsup_n A_n = \cap_N \cup_{n \geq N} A_n$.

Démonstration

Posons $B_0 = A_0$ et pour $n \geq 1, B_n = A_n - \cup_{k \leq n-1} B_k$ on a $\cup_{k \leq n} B_k = \cup_{k \leq n} A_k$ et les B_k sont disjoints, on a donc :
$P(\cup_{k \leq n} A_k) = P(\cup_{k \leq n} B_k) = \sum_{k \leq n} P(B_k) \leq \sum_{k \leq n} P(A_k)$ d'où $P(\cup_n A_n) \leq \sum_n P(A_n)$.
Montrons le lemme de Borel Cantelli :

$$P(\limsup A_n) = \lim_m P(\cup_{n \geq m} A_n) \leq \lim_m \sum_{n \geq m} P(A_n)$$

Or $\sum_{n \geq m} P(A_n) \leq \sum_n P(A_n) < \infty$ donc $\lim_m \sum_{n \geq m} P(A_n) = 0$.

1.1.7 Définitions : parties négligeables, espaces complets

Dans un espace de probabilité $(\Omega, \mathcal{F}, P)$ on définit l'ensemble des *parties négligeables* $\mathcal{N}(\Omega, \mathcal{F}, P)$ par :

$$\mathcal{N}(\Omega, \mathcal{F}, P) = \{N \subseteq \Omega \mid \exists B \in \mathcal{F}, N \subseteq B, P(B) = 0\}$$

Si $A(\omega)$ est une assertion dépendante de $\omega \in \Omega$ on dira que A est vraie *presque sûrement* (PS) si elle est vraie pour tout ω sauf sur un ensemble négligeable de Ω. En particulier deux parties A, B de Ω sont égales presque sûrement si elles diffèrent d'une partie négligeable de Ω. On peut donc écrire : $B = A \cup N$ où $N \in \mathcal{N}(\Omega, \mathcal{F}, P)$.

Un espace de probabilité $(\Omega, \mathcal{F}, P)$ est dit *complet* si $\mathcal{N} = \mathcal{N}(\Omega, \mathcal{F}, P) \subseteq \mathcal{F}$. Tout espace de probabilité peut être complété en remplaçant $\mathcal{F}$ par $\sigma(\mathcal{F} \cup \mathcal{N})$ et en étendant P à $\sigma(\mathcal{F} \cup \mathcal{N})$ en posant $P = 0$ sur $\mathcal{N}$.

1.2 Mesure positive et mesure signée

1.2.1 Définition : mesure positive, mesure signée

Une *mesure positive* sur un espace mesurable $(\Omega, \mathcal{F})$ est une application μ de $\mathcal{F}$ dans $\mathbb{R}^+ \cup \{\infty\}$ telle que $\mu(\emptyset) = 0$ et qui soit σ - *additive*, c'est-à-dire :

$$\forall (A_n) \subseteq \mathcal{F}, \forall i \neq j, A_i \cap A_j = \emptyset \Longrightarrow \sum_n \mu(A_n) = \mu\left(\bigcup_n A_n\right) \text{ dans } \mathbb{R}^+ \cup \{\infty\}$$

Une *mesure* ou *mesure signée* sur un espace mesurable $(\Omega, \mathcal{F})$ est une application μ de $\mathcal{F}$ dans $\mathbb{R} \cup \{\infty\}$ telle que $\mu(\emptyset) = 0$ et qui est σ - additive, ce qui signifie que pour toute suite (A_n) d'événements disjoints la série $\sum_n \mu(A_n)$ est absolument convergente dans $\mathbb{R} \cup \{\infty\}$ et sa somme est $\mu(\bigcup_n A_n)$.

Remarque

Toute mesure positive μ est croissante dans $\mathcal{F}$ munie de l'inclusion puisque :

$$A, B \in \mathcal{F},\ A \subseteq B \Longrightarrow \mu(B) = \mu(A) + \mu(B - A) \text{ dans } \mathbb{R}^+ \cup \{\infty\}$$

donc $\mu(B) < \infty \Longrightarrow \mu(A) < \infty$ et $\mu(B) - \mu(A) = \mu(B - A) \geq 0$ dans $\mathbb{R}^+$.

Il en résulte qu'une mesure positive finie (à valeur dans $\mathbb{R}^+$) est bornée car :

$$\forall A \in \mathcal{F},\ \mu(A) \leq \mu(\Omega) \in \mathbb{R}^+$$

Exemple : la mesure de Lebesgue

On montre à l'aide du théorème de Carathéodory (§3.4.4) qu'il existe une unique probabilité P sur $[0, 1]$ muni de la tribu borélienne $\mathcal{B}([0, 1])$ telle que $P([0]) = 0$ et pour tout intervalle $]a, b] \subseteq [0, 1]$, $P(]a, b]) = b - a$ c'est la *probabilité uniforme*.

On étend cette probabilité en une mesure positive λ sur $\mathbb{R}$ en posant :

$$\forall B \in \mathcal{B}(\mathbb{R}),\ \lambda(B) = \sum_{k \in \mathbb{Z}} P(B \cap]k, k+1] - k)$$

Cette mesure est appelée *mesure de Lebesgue*. Elle est invariante par translation.

Application : Une partie non borélienne de $\mathbb{R}$

On va construire une partie A de $\mathbb{R}$ qui n'appartient pas à la tribu borélienne $\mathcal{B}(\mathbb{R})$. Définissons la relation d'équivalence $\mathcal{Q}$ dans $[-1, 1]$ par :

$$x \ \mathcal{Q} \ y \Longleftrightarrow x - y \in \mathbb{Q}$$

Soit $x_{\mathcal{Q}}$ la classe d'équivalence de $x \in [-1, 1]$ et f une application qui à toute classe $x_{\mathcal{Q}} \subseteq [-1, 1]$ associe un représentant de cette classe donc $f(x_{\mathcal{Q}}) \in x_{\mathcal{Q}}$ (une telle application existe d'après l'axiome du choix). Notons qu'en général $f(x_{\mathcal{Q}}) \neq x$ en effet si $\mathcal{C}$ est une classe et si $f(\mathcal{C}) = y$ alors pour tout $x \neq y$ de $\mathcal{C}$ on a $f(x_{\mathcal{Q}}) \neq x$. Posons :

$$A = \{f(x_{\mathcal{Q}}) \mid x \in [-1, 1]\} \text{ et } L = \bigcup \{A + r \mid r \in [-2, 2] \cap \mathbb{Q}\}$$

On va montrer par l'absurde que A n'appartient pas à $\mathcal{B}(\mathbb{R})$. Si $A \in \mathcal{B}(\mathbb{R})$ alors L aussi, ces deux événements ont donc une mesure de Lebesgue.
On établi d'abord $[-1,1] \subseteq L \subseteq [-3,3]$ donc :

$$2 \leq \lambda(L) \leq 6 \quad (1)$$

puis que les événements $A+r$ sont disjoints donc :

$$\lambda(L) = \sum_{r \in [-2,2] \cap \mathbb{Q}} \lambda(A+r) \quad (2)$$

Comme $\lambda(A+r) = \lambda(A)$ il résulte de (2) que $\lambda(A) = 0$ implique $\lambda(L) = 0$ et $\lambda(A) > 0$ implique $\lambda(L) = \infty$ ce qui contredit (1). Examinons les points (1) et (2) en détails :

1. $[-1,1] \subseteq L$ puisque $\forall x \in [-1,1], f(x_{\mathcal{Q}}) \in x_{\mathcal{Q}} \Longrightarrow x - f(x_{\mathcal{Q}}) = r$ donc $x \in A+r$ avec $r \in [-2,2] \cap \mathbb{Q}$. Par ailleurs $L \subseteq [-3,3]$ car $A \subseteq [-1,1]$ et $r \in [-2,2]$ donc $[-1,1] \subseteq L \subseteq [-3,3]$ et $2 \leq \lambda(L) \leq 6$
2. $\forall r, s \in \mathbb{Q}, r \neq s \Longrightarrow (A+r) \cap (A+s) = \emptyset$. Par l'absurde : $x = f(y_{\mathcal{Q}}) + r$ et $x = f(z_{\mathcal{Q}}) + s \Longrightarrow f(y_{\mathcal{Q}}) - f(z_{\mathcal{Q}}) = s - r \in \mathbb{Q} \Longrightarrow y_{\mathcal{Q}} = z_{\mathcal{Q}} \Longrightarrow r = s$ $\square$

On va montrer que toute mesure signée μ se décompose en la différence d'une mesure positive μ_+ et d'une mesures positives finie μ_-. On note dans la suite du paragraphe μ une mesure signée sur $(\Omega, \mathcal{F})$.

1.2.2 Définition : événement négatif, nul, positif

- Un événement N est *négatif* si pour tout événement $N' \subseteq N$, $\mu(N') \leq 0$
- Un événement O est *nul* si pour tout événement $O' \subseteq O$, $\mu(O') = 0$
- Un événement P est *positif* si pour tout événement $P' \subseteq P$, $\mu(P') \geq 0$

1.2.3 Théorème : décomposition de Hahn

Pour toute mesure signée μ sur $(\Omega, \mathcal{F})$ il existe $N, P \in \mathcal{F}$ tels que $N \cup P = \Omega$ et $N \cap P = \emptyset$ avec N négatif et P positif. Cette décomposition de Ω relative à μ est dite de *Hahn*. Elle est unique au sens où si (N', P') en est une autre alors $N \Delta N'$ et $P \Delta P'$ sont nuls. On établit un lemme avant de démontrer le théorème.

Lemme : existence d'un événement négatif
Si $D \in \mathcal{F}$ et $\mu(D) \leq 0$ il existe un événement négatif A inclus dans D tel que $\mu(A) \leq \mu(D)$.

Démonstration du lemme
On va construire une suite (B_n) d'événements disjoints inclus dans D et montrer que $A = D - \bigcup_n B_n$ est négatif et vérifie $\mu(A) \leq \mu(D)$.

On pose $t_0 = \sup\{\mu(B) \mid B \in \mathcal{F}, B \subseteq D\}$ on a $t_0 \geq 0$ car $\emptyset$ satisfait la condition. Il existe $B_0 \in \mathcal{F}$ tel que $B_0 \subseteq D$ avec $\mu(B_0) \geq \frac{t_0}{2} \wedge 1$ (il est nécéssaire de minorer la condition par 1 car on peut avoir $t_0 = \infty$) puis par récurrence pour $n \geq 1$, on pose :

$$t_n = \sup\{\mu(B) \mid B \in \mathcal{F}, B \subseteq D - \bigcup_{k \leq n-1} B_k\} \geq 0$$

($\emptyset$ vérifie la condition). Il existe B_n tel que

$$B_n \in \mathcal{F},\ B_n \subseteq D - \bigcup_{k\leq n-1} B_k \text{ avec } \mu(B_n) \geq \frac{t_n}{2} \wedge 1$$

Les B_n sont disjoints et $\cup_n B_n \subseteq D$. On peut avoir $t_n = 0$ à partir d'un certain rang, dans ce cas on prend $B_n = \emptyset$ à partir de ce rang. Posons :

$$A = D - \bigcup_n B_n$$

On a $\mu(A) = \mu(D) - \mu(\cup_n B_n) = \mu(D) - \sum_n \mu(B_n) \leq \mu(D) - \sum_n (\frac{t_n}{2} \wedge 1) \leq \mu(D)$
Montrons par l'absurde que A est négatif. S'il existe $B \subseteq A,\ \mu(B) > 0$ on a $\forall n \geq 1$,

$$B \subseteq D - \bigcup_{k\leq n-1} B_n \Longrightarrow \mu(B) \leq t_n$$

ce qui entraîne :

$$\mu(A) \leq \mu(D) - \sum_n \left(\frac{t_n}{2} \wedge 1\right) \leq \mu(D) - \sum_n \left(\frac{\mu(B)}{2} \wedge 1\right) = -\infty$$

qui est impossible car μ ne prend pas la valeur $-\infty$. $\square$

Démonstration du théorème de décomposition de Hahn §1.2.3
On va construire une suite (A_n) d'événements disjoints négatifs et montrer que $N = \bigcup_n A_n$ est négatif et que $P = \Omega - N$ est positif. On pose :

$$s_0 = \inf\{\mu(D) \mid D \in \mathcal{F}, D \subseteq \Omega\} \leq 0$$

Il existe :

1. $D_0 \subseteq \Omega$ tel que $\mu(D_0) \leq \dfrac{s_0}{2} \vee (-1)$
2. $A_0 \subseteq D_0$ négatif avec $\mu(A_0) \leq \mu(D_0)$ (lemme)

puis par récurence pour $n \geq 1$, on pose :

$$s_n = \inf\{\mu(D) \mid D \in \mathcal{F}, D \subseteq \Omega - \bigcup_{k\leq n-1} A_k\} \leq 0$$

Il existe :

1. $D_n \subseteq \Omega - \bigcup_{k\leq n-1} A_k$ tel que $\mu(D_n) \leq \frac{s_n}{2} \vee (-1)$
2. $A_n \subseteq D_n$ négatif avec $\mu(A_n) \leq \mu(D_n)$

Notons que que les A_n sont disjoints. Posons :

$$N = \bigcup_n A_n$$

D'abord N est négatif car $\forall B \subseteq N, \mu(B) = \sum_n \mu(B \cap A_n) \leq 0$ puisque A_n est négatif.
Ensuite $P = \Omega - N$ est positif car s'il existe $D \subseteq P,\ \mu(D) < 0$ on a pour tout $n \geq 1$,

$$\bigcup_{k\leq n-1} A_k \subseteq N \Longrightarrow \Omega - N \subseteq \Omega - \bigcup_{k\leq n-1} A_k \Longrightarrow D \subseteq \Omega - \bigcup_{k\leq n-1} A_k \Longrightarrow s_n \leq \mu(D) < 0$$

ce qui entraîne :

$$\mu(N)=\sum_n \mu(A_n)\le\sum_n\mu(D_n)\le\sum_n\left[\frac{s_n}{2}\vee(-1)\right]\le\sum_n\left[\frac{\mu(D)}{2}\vee(-1)\right]=-\infty$$

qui est impossible car μ ne prend pas la valeur $-\infty$.

Vérifions l'unicité de la décomposition de Hahn. Soit (P,N) et (P',N') deux décompositions de Hahn de Ω relative à μ. On a :

$$P=(P\cap N')\cup(P\cap P') \text{ et } P'=(P'\cap N)\cup(P'\cap P)$$

donc $P-P'=P\cap N'$ et $P'-P=P'\cap N$ ce qui entraîne :

$$P\Delta P'=(P-P')\cup(P'-P)=(P\cap N')\cup(P'\cap N)$$

De la même manière $N\Delta N'=(N\cap P')\cup(N'\cap P)$ donc $P\Delta P'=N\Delta N'$ et un événement à la fois positif et négatif est nul. □

1.2.4 Corollaire : décomposition de Jordan

Si (N,P) est une décomposition Hahn de Ω relative à μ alors $\mu=\mu_+-\mu_-$ où μ_+ est une mesure positive définie pour tout $A\in\mathcal{F}$ par $\mu_+(A)=\mu(A\cap P)$ et μ_- est une mesure positive finie définie pour tout $A\in\mathcal{F}$ par $\mu_-(A)=-\mu(A\cap N)$

Cette décomposition de μ relative à une décomposition Hahn de Ω relative à μ est dite de *Jordan*. Elle est unique au sens où si (N',P') est une autre décomposition de Hahn de Ω relative à μ alors $\mu'_+=\mu_+$ et $\mu'_-=\mu_-$

Démonstration : unicité de la décomposition de Jordan

Il vient en tenant compte du fait que $N'-N$ et $N-N'$ sont des événements nuls :

$$\begin{aligned}\mu'_-(A)=-\mu(A\cap N')=-\mu(A\cap N'\cap N)-\mu[A\cap(N'-N))\\=-\mu(A\cap N\cap N')-\mu[A\cap(N-N')]=-\mu(A\cap N)=\mu_-(A)\end{aligned}$$

De même $\mu'_+(A)=\mu_+(A)$ □

1.2.5 Proposition : minimalité de la décomposition de Jordan

Si une mesure positive μ_1 et une mesure positive finie μ_2 sont telles que $\mu=\mu_1-\mu_2$ alors $\mu_1\ge\mu_+$ et $\mu_2\ge\mu_-$ le couple (μ_+,μ_-) est donc minimal.

Démonstration

Remarquons d'abord que :

$$\mu_1\ge\mu_+ \text{ et } \mu_2\ge\mu_-\Longleftrightarrow\mu_1+\mu_2\ge\mu_++\mu_-$$

Le sens direct s'obtient par addition, montrons la réciproque. Si $\mu_1+\mu_2\ge\mu_++\mu_-$ comme $\mu=\mu_+-\mu_-=\mu_1-\mu_2$ il vient :

— $\mu_1\ge\mu_++\mu_--\mu_2=\mu_++\mu_--(\mu_1-\mu_++\mu_-)=2\mu_+-\mu_1\Longrightarrow\mu_1\ge\mu_+$

— $\mu_2\ge\mu_++\mu_--\mu_1=\mu_++\mu_--(\mu_2+\mu_+-\mu_-)=2\mu_--\mu_2\Longrightarrow\mu_2\ge\mu_-$

Pour montrer la minimalité du couple (μ_+, μ_-) il suffit donc d'établir la minimalité de la somme $\mu_+ + \mu_-$ c'est-à-dire $\mu_1 + \mu_2 \geq \mu_+ + \mu_-$ il vient :

$$\begin{aligned}\mu_1(A) + \mu_2(A) &= \mu_1(A \cap P) + \mu_1(A \cap N) + \mu_2(A \cap P) + \mu_2(A \cap N)\\ &\geq |\mu_1(A \cap P) - \mu_2(A \cap P)| + |\mu_1(A \cap N) - \mu_2(A \cap N)|\\ &= \mu(A \cap P) - \mu(A \cap N) = \mu_+(A) + \mu_-(A) \quad \square\end{aligned}$$

1.2.6 Définition et proposition : variation d'une mesure

On définit la *variation* $v(\mu)$ d'une mesure signé μ par

$$v(\mu) = \mu_+ + \mu_-$$

Elle vérifie pour tout $A \in \mathcal{F}$, $v(\mu)(A) = \sup \sum_n |\mu(A_n)|$ où le sup est étendu à toutes les partitions dénombrables mesurables de A et on a :

$$\mu_+ = \frac{v(\mu) + \mu}{2} \quad \text{et} \quad \mu_- = \frac{v(\mu) - \mu}{2}$$

Démonstration
Soit (A_n) une partitions mesurables de A, montrons $v(\mu)(A) \geq \sum_n |\mu(A_n)|$ il vient :

$$\begin{aligned}v(\mu)(A) &= \mu_+(A) + \mu_-(A) = \textstyle\sum_n \mu_+(A_n) + \sum_n \mu_-(A_n)\\ &= \textstyle\sum_n [\mu_+(A_n) + \mu_-(A_n)] \geq \sum_n |\mu_+(A_n) - \mu_-(A_n)| = \sum_n |\mu(A_n)|\end{aligned}$$

et on a l'égalité avec la partition $\{A \cap P, A \cap N\}$ puisque :

$$v(\mu)(A) = \mu_+(A) + \mu_-(A) = \mu(A \cap P) - \mu(A \cap N) = \mu(A \cap P) + |\mu(A \cap N)| \quad \square$$

1.2.7 Définition : mesure *Lebesgue-Stieltjes* et fonction de répartition

1) On appelle mesure de *Lebesgue-Stieltjes* une mesure signée λ définie sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ telle que $\forall x \in \mathbb{R},\ \lambda(]-\infty, x]) < \infty$

2) On appelle *fonction de répartition* d'une mesure de Lebesgue-Stieltjes λ l'application $F_\lambda : \mathbb{R} \to \mathbb{R}$ définie par $F_\lambda(x) = \lambda(]-\infty, x])$

3) On appelle mesure de Lebesgue-Stieltjes sur $(\mathbb{R}^+, \mathcal{B})$ où $\mathcal{B} = \mathcal{B}_{\mathbb{R}^+}$ une mesure signée λ définie sur $(\mathbb{R}^+, \mathcal{B})$ telle que $\forall x \in \mathbb{R}^+,\ \lambda([0, x]) < \infty$

4) On appelle *fonction de répartition* d'une mesure de *Lebesgue-Stieltjes* λ sur $(\mathbb{R}^+, \mathcal{B})$ l'application $F_\lambda : \mathbb{R}^+ \to \mathbb{R}$ définie par $F_\lambda(x) = \lambda([0, x])$

1.2.8 Proposition : fonctions de répartition

1) Si λ est une mesure positive de Lebesgue-Stieltjes sa fonction de répartition F_λ est une fonction *càdlàg* (c'est-à-dire continue à droite et ayant une limite à gauche en tout point) croissante et positive. Elle est bornée si et seulement si la mesure positive λ est finie.

2) Si une mesure signée λ est de Lebesgue-Stieltjes il en est de même des mesures λ_+ et λ_- de sa décomposition de Jordan. Si F_{λ_+} et F_{μ_-} sont les fonctions de répartition de λ_+ et λ_- on a $F_\lambda = F_{\lambda_+} - F_{\lambda_-}$ et F_{λ_-} est bornée. Si $F_{v(\lambda)}$ est la fonction de répartition de la variation $v(\lambda)$ de λ on a $F_{v(\lambda)} = F_{\lambda_+} + F_{\lambda_-}$ et

$$F_{\lambda_+} = \frac{F_{v(\lambda)} + F_\lambda}{2} \quad \text{et} \quad F_{\lambda_-} = \frac{F_{v(\lambda)} - F_\lambda}{2}$$

3) Si λ est une mesure de Lebesgue-Stieltjes on a :

$$\lambda(]a,b]) = F_\lambda(b) - F_\lambda(a) \quad \text{et} \quad \lambda([a]) = F_\lambda(a) - F_\lambda(a_-)$$

Notons que si λ est une mesure de Lebesgue-Stieltjes sur $(\mathbb{R}^+, \mathcal{B})$ on a $\lambda([0]) = F_\lambda(0)$.

4) Si $F : \mathbb{R} \to \mathbb{R}^+$ est une fonction *càdlàg*, croissante et positive elle est la fonction de répartition d'une unique mesure de Lebesgue-Stieltjes positive λ.

5) Si une fonction *càdlàg* $F : \mathbb{R} \to \mathbb{R}$ est la différence de deux fonction *càdlàg*, croissantes et positives $F = F_1 - F_2$ où F_2 est bornée elle est la fonction de répartition d'une mesure signée de *Lebesgue-Stieltjes* λ. Plus précisément si λ_1 et λ_2 sont les mesures positives de *Lebesgue-Stieltjes* de fonctions de répartition F_1 et F_2 on a la décomposition $\lambda = \lambda_1 - \lambda_2$ et la mesure λ_2 est finie.

6) Dans les notations du point (5) les assertions suivantes sont équivalentes :

1. le couple (F_1, F_2) est minimal au sens où $F = F_1 - F_2 = F_3 - F_4$ implique $F_1 \leq F_3$ et $F_2 \leq F_4$
2. le couple (λ_1, λ_2) est minimale au sens où $\lambda = \lambda_1 - \lambda_2 = \lambda_3 - \lambda_4$ implique $\lambda_1 \leq \lambda_3$ et $\lambda_2 \leq \lambda_4$ (voir §1.2.5)
3. $\lambda_1 = \lambda_+$ et $\lambda_2 = \lambda_-$

Démonstration
1) La fonction de répartition F_λ d'une mesure λ positive de *Lebesgue-Stieltjes* est positive, croissante et continue à droite puisque $\lim_{y\downarrow x} \,]-\infty, y] = \,]-\infty, x]$ entraîne $\lim_{y\downarrow x} F_\lambda(y) = F_\lambda(x)$. Elle est donc *càdlàg*.

2) Si λ est une mesure signée on a la décompositions $\lambda = \lambda_+ - \lambda_-$ où λ_- est finie donc $\lambda(]-\infty, x]) = \lambda_+(]-\infty, x]) - \lambda_-(]-\infty, x])$ entraîne $\lambda_+(]-\infty, x]) < \infty$ lorsque λ est de *Lebesgue-Stieltjes*, il en découle $F_\lambda = F_{\lambda_+} - F_{\lambda_-}$. Cette égalité entraîne le caractère *càdlàg* de F_λ en outre F_λ est bornée puisque la mesure λ_- est finie.
On a de même la décompositions $v(\lambda) = \lambda_+ + \lambda_-$ et donc $F_{v(\lambda)} = F_{\lambda_+} + F_{\lambda_-}$

3) L'égalité $\lambda([a]) = F_\lambda(a) - F_\lambda(a_-)$ découle de l'existence de limites à gauche de F_λ et de la continuité de λ.

4) Soit $F : \mathbb{R} \to \mathbb{R}^+$ une application *càdlàg* croissante et positive. Pour tous les intervalles de $\mathbb{R}$ de la forme $]a,b]$ on pose $\lambda(]a,b]) = F(b) - F(a)$. Soit $A < B$ dans $\mathbb{R}$, l'application λ peut être étendue par additivité à l'algèbre $\mathcal{A}_{]A,B]}$ des unions finies d'intervalles $]a,b]$ contenus dans l'intervalle fixe $]A,B]$ puis à la tribu borélienne $\mathcal{B}_{]A,B]}$ (voir le théorème de Carathéodory §3.4.4) et enfin à la tribu $\mathcal{B}_\mathbb{R}$ en posant pour tout $B \in \mathcal{B}_\mathbb{R}$:

$$\lambda(B) = \sum_{k\in\mathbb{Z}} \lambda(B \cap \,]k, k+1])$$

5) L'existence de λ résulte de celle de λ_1 et λ_2.

6) L'équivalence entre (2) et (3) a été établie au §1.2.5. L'équivalence avec (1) découle de la définition des fonctions de répartition. □

1.3 Applications mesurables

1.3.1 Définition : applications mesurables

Une application X d'un espace mesurable $(\Omega, \mathcal{F})$ dans un espace mesurable $(E, \mathcal{B})$ est *mesurable* si $X^{-1}(\mathcal{B}) \subseteq \mathcal{F}$ c'est-à-dire :

$$\forall B \in \mathcal{B},\ X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\} = (X \in B) \in \mathcal{F}$$

Dans le cas particulier ou E est la droite réelle fermée $E = \mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$ et ou $\mathcal{B}$ est sa tribu borélienne on dit que X est $\mathcal{F}$ - mesurable pour exprimer que X est mesurable de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{B})$. On appelle :

— variable aléatoire (VA) les applications $\mathcal{F}$ - mesurables.

— variable aléatoire réelle (VAR) les applications $\mathcal{F}$ - mesurables à valeurs réelles. On note $\mathcal{F}$ la classe des VAR.

— variable aléatoire positive ou temps aléatoire les applications $\mathcal{F}$ - mesurables à valeurs dans $\mathbb{R}^+ \cup \{\infty\}$ On note $\mathcal{F}^+$ la classe des VA positives. On montrera au corollaire §1.3.5 que $\mathcal{F}$ est une algèbre et que $\mathcal{F}^+$ un cône.

Exemple : fonction étagée

On appelle *fonction étagée* sur Ω une combinaison linéaire finie d'indicatrices de sous ensembles A_i de Ω. Une application X étagée sur Ω est donc de la forme :

$$X(\omega) = \sum_{i=1,n} x_i 1_{A_i}(\omega)$$

où x_i est un réel et $1_{A_i}(\omega)$ vaut 1 si $\omega \in A_i$ et 0 sinon. Si pour tout i l'ensemble A_i appartient à une tribu $\mathcal{F}$ de Ω on dit que X est étagée sur $\mathcal{F}$ elle est alors $\mathcal{F}$ - mesurable.

1.3.2 Définition : égalité presque sûre des variables aléatoires

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité. Deux VA X et Y sont *égales presque sûrement* si $(X \neq Y)$ est négligeable, on écrit $X = Y$ PS. On note $L^0 = L^0(\Omega, \mathcal{F}, P)$ l'espace vectoriel des *classes* pour l'égalité PS des VAR. On confondra souvent une VAR X et sa classe d'équivalence.

1.3.3 Proposition : critère de mesurabilité de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{B})$

Soit X une application de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{B})$. Si $\Gamma \subseteq \mathcal{B}$ est telle que $\sigma(\Gamma) = \mathcal{B}$ alors X est mesurable de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{B})$ si et seulement si $X^{-1}(\Gamma) \subseteq \mathcal{F}$.

Lemme de commutativité
Si X est une application de Ω dans E et Γ un ensemble de parties de E alors :

$$\sigma[X^{-1}(\Gamma)] = X^{-1}[\sigma(\Gamma)]$$

Démonstration
D'abord $\Gamma \subseteq \sigma(\Gamma)$ donc $X^{-1}(\Gamma) \subseteq X^{-1}[\sigma(\Gamma)]$ et comme $X^{-1}[\sigma(\Gamma)]$ est une tribu on obtient $\sigma[X^{-1}(\Gamma)] \subseteq X^{-1}[\sigma(\Gamma)]$. Inversement posons :

$$\mathcal{C} = \{B \in \sigma(\Gamma) \mid X^{-1}(B) \in \sigma[X^{-1}(\Gamma)]\}$$

On observe $\Gamma \subseteq \mathcal{C}$ donc $\sigma(\Gamma) \subseteq \mathcal{C}$ car $\mathcal{C}$ est une tribu. Il en résulte $\mathcal{C} = \sigma(\Gamma)$ donc $X^{-1}[\sigma(\Gamma)] \subseteq \sigma[X^{-1}(\Gamma)]$.
Démonstration de la proposition
Il faut montrer $X^{-1}(\Gamma) \subseteq \mathcal{F} \Longrightarrow X^{-1}(\mathcal{B}) \subseteq \mathcal{F}$ on a $X^{-1}(\Gamma) \subseteq \mathcal{F} \Longrightarrow \sigma[X^{-1}(\Gamma)] \subseteq \mathcal{F}$ car $\mathcal{F}$ est une tribu. Il vient alors en utilisant le lemme :

$$X^{-1}(\mathcal{B}) = X^{-1}[\sigma(\Gamma)] = \sigma[X^{-1}(\Gamma)] \subseteq \mathcal{F} \; \square$$

Corollaire
Une application X d'un espace mesurable $(\Omega, \mathcal{F})$ dans la droite réelle fermée munie de sa tribu borélienne est mesurable si et seulement si tous les ensembles de la forme :

$$(X \leq a) = X^{-1}([-\infty, a])$$

où $a \in \mathbb{R}$ appartiennent à $\mathcal{F}$ puisque les demi droites $[-\infty, a]$ engendrent $\mathcal{B}([-\infty, \infty])$. Il en est de même pour les ensembles de la forme $(X < a), (X \geq a), (X > a)$.
Remarque
Il résulte du corollaire que si X et Y sont des VA alors $(X \neq Y) \in \mathcal{F}$ car

$$(X \neq Y) = \cup_{q \in \mathbb{Q}} (X < q < Y) \cup (Y < q < X)$$

Par conséquent deux VA égales PS diffèrent sur un ensemble négligeable mesurable.

1.3.4 Proposition : mesurabilité des limites

Si $(X_n)_{n \in \mathbb{N}}$ est une suite d'applications $\mathcal{F}$ - mesurables, les applications suivantes : $\sup_n X_n$, $\inf_n X_n$, $\limsup_n X_n$, $\liminf_n X_n$ et $\lim_n X_n$ (si elle est définie) sont $\mathcal{F}$ - mesurables.
Démonstration
Posons $X = \sup_n X_n$ et $Y = \inf_n X_n$ on a

$$\forall a \in \mathbb{R}, (X \leq a) = \bigcap_n (X_n \leq a) \in \mathcal{F}, \; (Y \geq a) = \bigcap_n (X_n \geq a) \in \mathcal{F}$$

Les VA X et Y sont donc $\mathcal{F}$ - mesurables d'après le corollaire §1.3.3, il en est donc de même de $\limsup_n X_n$, $\liminf_n X_n$ et $\lim_n X_n$. $\square$

1.3.5 Proposition : approximation par les fonctions étagées

Une application X de Ω dans $\mathbb{R}^+ \cup \{\infty\}$ est $\mathcal{F}$ - mesurable si et seulement s'il existe une suite (X_n) croissante d'applications positives étagées sur $\mathcal{F}$ telle que $\lim_n \uparrow X_n = X$

Démonstration

S'il existe une suite (X_n) d'applications étagées sur $\mathcal{F}$ telle que $\lim_n \uparrow X_n = X$ l'application X est $\mathcal{F}$ - mesurables. Inversement, posons :

$$A_{n,k} = \left[\frac{k}{2^n} \leq X < \frac{k+1}{2^n}\right],\ X_n = \sum_{k=0,n2^n-1} \frac{k}{2^n} 1_{A_{n,k}}$$

L'application X_n est positive et étagée sur $\mathcal{F}$, la suite (X_n) est croissante et on vérifie que $\lim_n \uparrow X_n = X$ donc X est $\mathcal{F}$ - mesurables. $\square$

Corollaire

La classe $\mathcal{F}$ des VAR est une algèbre et la classe $\mathcal{F}^+$ des VA positives est un cône.

Démonstration

Le corollaire résulte du fait que la somme et le produit de fonctions étagées sont des fonctions étagées.

1.3.6 Définition : tribu engendrée par une famille d'applications

Soit $\mathcal{C}$ une famille d'applications de Ω dans $(E, \mathcal{B})$. On appelle tribu engendrée par $\mathcal{C}$ et on note $\sigma(\mathcal{C})$ la plus petite tribu sur Ω qui rend mesurables toutes les applications $X \in \mathcal{C}$, on a donc :

$$\forall X \in \mathcal{C}, \forall B \in \mathcal{B},\ (X \in B) \in \sigma(\mathcal{C})$$

Exemples

1) La tribu $\sigma(X)$ engendrée par une seule application X est égale à $X^{-1}(\mathcal{B})$, en effet $\sigma(X)$ contient nécessairement $X^{-1}(\mathcal{B})$ et l'image réciproque d'une tribu est une tribu.

2) Si E est un espace métrique la tribu borélienne $\mathcal{B}_E$ coïncide avec la *tribu de Baire* $\sigma(\mathcal{C}_E)$ engendrée par les applications continues de E dans $\mathbb{R}$ (espace $\mathcal{C}_E$). En effet si G est un ouvert de $\mathbb{R}$ et $f \in \mathcal{C}_E$ alors $f^{-1}(G)$ est un ouvert de E donc $\sigma(\mathcal{C}_E) \subseteq \mathcal{B}_E$ (ce qui est vrai pour tout espace topologique E). Inversement si F est un fermé de l'espace métrique E alors $F = d_F^{-1}(0)$ où $d_F(x) = \inf\{d(x,y) \mid y \in F\}$ est continue donc $F \in \sigma(\mathcal{C}_E)$ par conséquent $\mathcal{B}_E \subseteq \sigma(\mathcal{C}_E)$ et finalement $\mathcal{B}_E = \sigma(\mathcal{C}_E)$.

1.3.7 Proposition : factorisation des applications mesurables

Soient X une application mesurable de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{B})$ et Y une application mesurable de $(\Omega, \mathcal{F})$ dans $\mathbb{R}$. L'application Y est $\sigma(X)$ - mesurable si et seulement s'il existe une application g mesurable de $(E, \mathcal{B})$ dans $\mathbb{R}$ telle que $Y = g(X)$.

Démonstration

Si Y est de la forme $Y = g(X)$ elle est mesurable de $[\Omega, \sigma(X)]$ dans $\mathbb{R}$ puisque X est mesurable de $[\Omega, \sigma(X)]$ dans $(E, \mathcal{B})$, que g est mesurable de $(E, \mathcal{B})$ dans $\mathbb{R}$ et que pour tout borélien B on a $Y^{-1}(B) = X^{-1}[g^{-1}(B)] \in \sigma(X)$.

Inversement construisons g mesurable de $(E,\mathcal{B})$ dans $\mathbb{R}$ telle que $Y = g(X)$.
Supposons d'abord que Y est étagée sur $\sigma(X)$. Il existe une partition (A_i) de Ω telle que $Y = \sum_i y_i 1_{A_i}$ avec les y_i distincts. Comme $\sigma(X) = X^{-1}(\mathcal{B})$ pour tout A_i il existe $B_i \in \mathcal{B}$ tel que $A_i = X^{-1}(B_i)$ posons alors $g = \sum_i y_i 1_{B_i}$ on a $Y = g(X)$ puisque :

$$\forall\omega, \exists i,\ \omega \in A_i \Longrightarrow X(\omega) \in B_i \Longrightarrow g[X(\omega)] = y_i = Y(\omega)$$

Si Y est $\sigma(X)$ - mesurable positive on a $Y = \lim_n \uparrow Y_n$ avec Y_n étagée sur $\sigma(X)$ donc pour tout n, il existe g_n mesurable de $(E,\mathcal{B})$ dans $\mathbb{R}$ telle que $Y_n = g_n(X)$ et la suite (g_n) est croissante sur $X(\Omega)$ (qui n'appartient pas nécéssairement à $\mathcal{B}$). Notons :

$$E_0 = \{x \in E \mid \forall n, g_n(x) \leq g_{n+1}(x)\} = \cap_{n\in\mathbb{N}} \cup_{r\in\mathbb{Q}} (g_{n+1} \geq r, g_n \leq r)$$

la partie de E où (g_n) est croissante on a $X(\Omega) \subseteq E_0$ et $E_0 \in \mathcal{B}$ posons alors :

$$g(x) = \lim_n \uparrow g_n(x) \text{ sur } E_0 \text{ et } g(x) = 0 \text{ sur } E_0^c$$

Comme pour tout $\omega \in \Omega$, $\lim_n \uparrow g_n[X(\omega)] = \lim_n \uparrow Y_n(\omega) = Y(\omega)$ on a $Y = g(X)$ et l'application g est mesurable de $(E,\mathcal{B})$ dans $\mathbb{R}$.

Si Y est de signe quelconque on a $Y = Y^+ - Y^-$ avec $Y^+ = \max(0,Y)$ et
$Y^- = -\min(0,Y)$ d'ou $Y = g^+(X) - g^-(X) = g(X)$ avec $g = g^+ - g^-$ $\square$

1.3.8 Définitions et proposition : tribu produit et mesurabilité

Soit $(\Omega_1,\mathcal{F}_1)$ et $(\Omega_2,\mathcal{F}_2)$ des espaces mesurables, la *tribu produit* $\mathcal{F}_1 \otimes \mathcal{F}_2$ sur $\Omega_1 \times \Omega_2$ est définie par :

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(A_1 \times A_2 \mid A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2)$$

1) Si $A \subseteq \Omega$ on définit les *coupes* $A(\omega_1) = \{\omega_2 \in \Omega_2 \mid (\omega_1,\omega_2) \in A\}$ et de même $A(\omega_2)$. Notons que si $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ alors $A(\omega_1) \in \mathcal{F}_2$ et $A(\omega_2) \in \mathcal{F}_1$. (Les coupes d'une partie mesurable sont mesurables mais que la réciproque est fausse).

2) Si $X \in \mathcal{F}_1 \otimes \mathcal{F}_2$ les *applications partielles* $X(\omega_1, .)$ et $X(., \omega_2)$ sont respectivement $\mathcal{F}_2$ et $\mathcal{F}_1$ mesurables. (Les applications partielles d'une application mesurable sont mesurables mais que la réciproque est fausse).

3) Si X est de la forme $X(\omega_1,\omega_2) = X_1(\omega_1)X_2(\omega_2)$ alors $X \in \mathcal{F}_1 \otimes \mathcal{F}_2$ si et seulement si $X_1 \in \mathcal{F}_1$ et $X_2 \in \mathcal{F}_2$

Démonstration
La classe $\mathcal{A}(\omega_1) = \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2 \mid A(\omega_1) \in \mathcal{F}_2\}$ est une tribu puisque $\mathcal{F}_2$ en est une et elle contient les rectangles $A_1 \times A_2$ où $A_1 \in \mathcal{F}_1$ et $A_2 \in \mathcal{F}_2$ car
$A_1 \times A_2(\omega_1) = A_2$ ou $\emptyset$. Elle contient donc $\mathcal{F}_1 \otimes \mathcal{F}_2$. Le deuxième point découle de $X(\omega_1)^{-1}(B) = X^{-1}(B)(\omega_1) \in \mathcal{F}_2$ Le troisième point découle de $1_{A_1}1_{A_2} = 1_{A_1 \times A_2}$ et du fait que le premier point entraîne $A_1 \times A_2 \in \mathcal{F}_1 \otimes \mathcal{F}_2 \Longleftrightarrow A_1 \in \mathcal{F}_1$ et $A_2 \in \mathcal{F}_2$ $\square$

Exemple : une partie non mesurable de $\Omega_1 \times \Omega_2$ de coupes mesurables.
Soit $\Omega_1 = \Omega_2 = \mathbb{R}^+$ et $\mathcal{F}$ la tribu sur $\mathbb{R}^+$ engendrée par les singletons de $\mathbb{R}^+$. Les coupes de la diagonale $\Delta = \{(u,u) \mid u \in \mathbb{R}^+\}$ sont des singletons de $\mathbb{R}^+$ ils sont donc

mesurables mais Δ n'appartiennent pas à $\mathcal{F}^2$. En revanche si $\mathcal{F} = \mathcal{B}(\mathbb{R})$ alors Δ est mesurable car :

$$\Delta = \bigcap_n \bigcup_k [k2^{-n}, (k+1)2^{-n}[\times [k2^{-n}, (k+1)2^{-n}[$$

Remarque : mesurabilité des projections
On définit les projecteurs π_1 et π_2 de $\Omega_1 \times \Omega_2$ sur Ω_1 et Ω_2 par $\pi_1(\omega_1, \omega_2) = \omega_1$ et $\pi_2(\omega_1, \omega_2) = \omega_2$. Si $\mathcal{F}_1$ et $\mathcal{F}_2$ sont des tribus sur Ω_1 et Ω_2 les projections $\pi_1(A)$ et $\pi_2(A)$ de $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ n'appartiennent pas en général à $\mathcal{F}_1$ et $\mathcal{F}_2$. Toutefois la construction d'ensembles mesurables de projections non mesurables n'est pas simple et utilise le plus souvent des éléments de théorie des ensembles analytiques.

Rappelons que si X est un espace métrique séparables complet (MSC ou espace polonais) une partie $A \subseteq X$ est dite *analytique* s'il existe un espace MSC Y et une application continue $f : Y \to X$ telle que $A = f(Y)$. On montre alors que $A \subseteq X$ est analytique si et seulement s'il existe un espace MSC Y et une partie borélienne B de $X \times Y$ telle que $A = \pi_X(B)$. Comme il existe des ensembles analytiques non boréliens ces derniers constituent des exemples de projections non mesurables d'ensembles mesurables.

Sans utiliser la théorie des ensembles analytiques on établira le théorème de capacité de Choquet §3.3.2 et en corollaire le théorème de projection mesurable §3.6.3 de Meyer qui donne des conditions dans lesquelles les projections sont mesurables.

1.4 Processus stochastiques

1.4.1 Définitions : processus stochastiques

1) Soit T un ensemble non vide, $(\Omega, \mathcal{F})$ et $(E, \mathcal{B})$ des espaces mesurables, un processus X est une application de $T \times \Omega$ dans $(E, \mathcal{B})$ telle que pour tout $t \in T$ l'application partielle $\omega \to X_t(\omega) = X(t, \omega)$ est mesurable de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{B})$. Si $(\Omega, \mathcal{F})$ est muni d'une probabilité P on dira qu'un processus défini sur $T \times \Omega$ est un *processus stochastique*.

2) On dira qu'un processus est à *temps discret* si $T = \mathbb{N}$ et à *temps continu* si $T = \mathbb{R}^+$. Un processus X est *réel* si $E = \mathbb{R}$ et $\mathcal{B} = \mathcal{B}(\mathbb{R})$. Un processus réel à temps continu est donc une famille $X = (X_t)_{t \in \mathbb{R}^+}$ de VAR. On appellera dans la suite *processus* un processus réel à temps continu. Dans certains cas on complétera un processus X par une *variable aléatoire terminale* X_∞ on aura alors $T = [0, \infty]$. Ce sera notamment le cas lorsque X_t à une limite lorsque $t \to \infty$ au sens de la convergence simple.

3) Un processus est *càglàd* si ses *trajectoires* c'est-à-dire les applications partielles $t \to X(t, \omega)$ sont continues à gauche (*càg*) en tout $t > 0$ et ont une limite à droite en tout $t \geq 0$. Un processus est *càdlàg* si ses trajectoires sont continues à droite (*càd*) en tout $t \geq 0$ et ont une limite à gauche en tout $t > 0$. Pour un processus X *càdlàg* on définit le *processus des limites à gauche* X_- par $X_-(t) = X(t-)$ pour $t > 0$ et $X_-(0) = X_{0-} = 0$. On définit également pour un processus X *càdlàg* le *processus de sauts* ΔX par $\Delta X_t = X_t - X_{t-}$ pour $t > 0$ et $\Delta X_0 = \Delta X_\infty = 0$ (pas de sauts en zéro, ni à l'infini).

4) Un *ensemble aléatoire* A c'est-à-dire une partie de $\mathbb{R}^+ \times \Omega$ est *évanescent* si sa projection $\pi(A)$ sur Ω est négligeable (voir au §1.1.7 la définition des négligeables), on a donc :

$$\pi(A) = \{\omega \mid \exists t,\ (t,\omega) \in A\} \in \mathcal{N}(\Omega, \mathcal{F}, P)$$

5) Deux processus X et Y sont *indistinguables* si $(X \neq Y) = \{(t,\omega) \mid X_t(\omega) \neq Y_t(\omega)\}$ est évanescent, on note $X = Y$ EP (à un évanescent près).

6) Un processus X est *évanescent* s'il est indistinguable du processus nul.

7) Un processus Y est une *modification* d'un processus X si $\forall t, P(X_t = Y_t) = 1$ on a donc dans ce cas $\forall t, X_t = Y_t$ PS (voir §1.3.2). Notons que si Y est une modification de X alors X est une modification de Y et que si X et Y sont indistinguables alors Y est une modification de X.

Remarque

Un ensemble évanescent A n'est pas nécessairement mesurable au sens ou $A \in \mathcal{B} \otimes \mathcal{F}$. Bien entendu on peut compléter la tribu $\mathcal{B} \otimes \mathcal{F}$ en la remplaçant par $\sigma(\mathcal{B} \otimes \mathcal{F}, \mathcal{E})$ ou $\mathcal{E}$ est l'ensemble des évanescents.

Exemple : processus modifiés et processus indistingables

Si $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}[0,1], \lambda)$ où λ est la mesure de Lebesgue alors le processus X défini sur $T \times \Omega$ où $T = [0,1]$ par $X_t(\omega) = 0$ pour $t \neq \omega$ et $X_t(\omega) = 1$ pour $t = \omega$ c'est-à-dire $X = 1_\Delta$ ou Δ est la diagonale de $[0,1]^2$ est une modification du processus nul mais n'est pas évanescent. On a en effet $\pi(X \neq 0) = \pi(\Delta) = [0,1]$

1.4.2 Proposition : modification, indistinguabilité et continuité

Soit X et Y des processus continus à droite ou à gauche. Si X est une modification de Y alors X est indistinguables de Y. Il en résulte qu'une modification *càdlàg* Y d'un processus X est unique à un évanescent près.

Démonstration

Dire que X est une modification de Y signifie $\forall t \in \mathbb{R}^+, P(X_t = Y_t) = 1$ et dans ce cas $P(\forall s \in \mathbb{Q}^+, X_s = Y_s) = P(\cap_{s \in \mathbb{Q}^+} X_s = Y_s) = 1$ car

$$P(\cup_{s \in \mathbb{Q}^+} X_s \neq Y_s) \leq \sum_{s \in \mathbb{Q}^+} P(X_s \neq Y_s) = 0$$

Posons $A = (\forall t \in \mathbb{R}^+, X_t = Y_t)$ et $B = (\forall s \in \mathbb{Q}^+, X_s = Y_s)$. On a montré $P(B) = 1$. La continuité à droite (ou à gauche) de X et Y entraîne $B \subseteq A$ donc $P(A) = 1$ ce qui signifie que X est indistinguable de Y. $\square$

1.4.3 Définitions : filtration et complétude

1) Une *filtration* sur un espace mesurable $(\Omega, \mathcal{F})$ est une suite $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ croissante (au sens de l'inclusion) de sous tribus de $\mathcal{F}$. Un espace mesurable muni d'une filtration est appelé *espace filtré*.

2) Si $[\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}^+}]$ est un espace filtré on définit deux autres filtrations $(\mathcal{F}_{t+})_{t \in \mathbb{R}^+}$ et $(\mathcal{F}_{t-})_{t \in \mathbb{R}^+}$ par $\mathcal{F}_{t+} = \cap_{s>t} \mathcal{F}_s$ et pour $t > 0, \mathcal{F}_{t-} = \sigma(\cup_{s<t} \mathcal{F}_s)$ et $\mathcal{F}_{0-} = \mathcal{F}_0$

3) Un processus $X = (X_t)_{t\in\mathbb{R}^+}$ défini sur un espace filtré $[\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}^+}]$ est dit *adapté* si pour tout $t \in \mathbb{R}^+$, la variable X_t est mesurable de $(\Omega, \mathcal{F}_t)$ dans $\mathbb{R}$. Si X est complété par une variable terminale X_∞ on supposera alors que X_∞ est mesurable de $(\Omega, \mathcal{F}_\infty)$ dans $\mathbb{R} \cup \{-\infty, \infty\}$ avec $\mathcal{F}_\infty = \sigma(\cup_t \mathcal{F}_t)$

4) Une filtration $(\mathcal{F}_t)_{t\in\mathbb{R}^+}$ dans $(\Omega, \mathcal{F}, P)$ est *complète* si $\mathcal{N}(\Omega, \mathcal{F}, P) \subseteq \mathcal{F}_0$ (voir §1.1.7) Un espace de probabilité filtré $[\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}^+}, P]$ est *complet* si $(\mathcal{F}_t)_{t\in\mathbb{R}^+}$ est complète. Dans ce cas l'espace $(\Omega, \mathcal{F}, P)$ et tous les espaces $(\Omega, \mathcal{F}_t, P)$ sont aussi complets.

1.4.4 Proposition : mesurabilité des limites à droite et à gauche

Soit X un processus adapté à une filtration $(\mathcal{F}_t)_{t\in\mathbb{R}^+}$. Si X a une limite à droite X_{t+} en $t \geq 0$ alors $X_{t+} \in \mathcal{F}_{t+}$ Si X a une limite à gauche X_{t-} en $t > 0$ alors $X_{t-} \in \mathcal{F}_{t-}$.

Démonstration
Si X à une limite à droite X_{t+} en $t \geq 0$, pour toute suite (t_n) décroissante vers t telle que $\forall n,\ t_n > t$ posons $Y_N = \sup_{n\geq N} X_{t_n}$ on a :

$$X_{t+} = \lim_n X_{t_n} = \limsup_n X_{t_n} = \inf_N Y_N \quad \text{et} \quad \forall x \in \mathbb{R}, (X_{t+} \geq x) = \cap_N (Y_N \geq x)$$

Il en résulte $\forall m, (X_{t+} \geq x) = \cap_{N\geq m}(Y_N \geq x) \in \mathcal{F}_{t_m}$ ce qui entraîne $(X_{t+} \geq x) \in \mathcal{F}_{t+}$ donc $X_{t+} \in \mathcal{F}_{t+}$.
Si X à une limite à gauche X_{t-} en $t > 0$, pour toute suite (t_n) croissante vers t, telle que $\forall n,\ t_n < t$ on a $X_{t-} = \lim_n X_{t_n}$ et $X_{t_n} \in \mathcal{F}_{t_n} \subseteq \cup_n \mathcal{F}_{t_n} \subseteq \mathcal{F}_{t-}$ donc $X_{t-} \in \mathcal{F}_{t-}$.
□

1.4.5 Définitions : processus mesurables

Un processus X défini un espace mesurable $(\Omega, \mathcal{F})$ est *mesurable* si l'application $(t, \omega) \to X(t, \omega)$ est mesurable de $(\mathbb{R}^+ \times \Omega, \mathcal{B} \otimes \mathcal{F})$ ou $\mathcal{B} = \mathcal{B}(\mathbb{R}^+)$ dans $[\mathbb{R}, \mathcal{B}(\mathbb{R})]$ On note $\mathcal{B} \otimes \mathcal{F}$ l'algèbre des processus mesurables et $b(\mathcal{B} \otimes \mathcal{F})$ l'algèbre des processus mesurables bornés.

1.4.6 Définitions : processus déterministes, processus intemporels

Un processus X est *déterministe* s'il existe une application $f : \mathbb{R}^+ \to \mathbb{R}$ telle que $\forall (t, \omega) \in \mathbb{R}^+ \times \Omega, X(t, \omega) = f(t)$ (une seule trajectoire indépendante de ω). Un processus X est indépendant du temps ou *intemporel* s'il est de la forme $X(t, \omega) = Z(\omega)$ où Z est une VAR.

Remarques
Si f est une application quelconque de $\mathbb{R}^+$ dans $\mathbb{R}$ et si X est défini par $X(t, \omega) = f(t)$ alors X est un processus adapté à toute filtration $(\mathcal{F}_t)$ sur $(\Omega, \mathcal{F})$ puisque $X_t^{-1}(B) = \emptyset$ si $f(t) \notin B$ et $X_t^{-1}(B) = \Omega$ si $f(t) \in B$
Un processus intemporel X de la forme $X(t, \omega) = Z(\omega)$ est adapté si et seulement si $Z \in \mathcal{F}_0$

1.4.7 Proposition : processus déterministes et mesurabilité

Soit X un processus déterministe de la forme $X(t,\omega) = f(t)$ les assertions suivantes sont équivalentes, on note $\mathcal{B} = \mathcal{B}(\mathbb{R}^+)$:

1. X est mesurable
2. f est mesurable de $(\mathbb{R}^+, \mathcal{B})$ dans $[\mathbb{R}, \mathcal{B}(\mathbb{R})]$
3. X est mesurable de $(\mathbb{R}^+ \times \Omega, \mathcal{B} \otimes \{\emptyset, \Omega\})$ dans $[\mathbb{R}, \mathcal{B}(\mathbb{R})]$

Démonstration
Si un processus déterministe X de la forme $X(t,\omega) = f(t)$ est mesurable alors l'application partielle f est mesurable (§1.3.8). Comme $\forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) = f^{-1}(B) \times \Omega$ le processus X est mesurable de $(\mathbb{R}^+ \times \Omega, \mathcal{B} \otimes \{\emptyset, \Omega\})$ dans $[\mathbb{R}, \mathcal{B}(\mathbb{R})]$ et inversement dans ce cas comme $\{\emptyset, \Omega\} \subseteq \mathcal{F}$ le processus X est mesurable. $\square$

1.5 Théorèmes de classes monotones

Le lemme des classes monotones permet l'extension d'une propriété satisfaite dans une classe $\mathcal{S}$ d'événements stable par intersection finie (π - système) à la tribu $\sigma(\mathcal{S})$ engendrée par $\mathcal{S}$.

1.5.1 Définition : π et δ système

Une classe $\mathcal{S}$ de parties d'un ensemble non vide est appelée un π - *système* si elle est stable par intersection finie. Une classe $\mathcal{M}$ de parties de Ω est une *classe monotone* ou δ - *système* si :

1. $\Omega \in \mathcal{M}$,
2. $A, B \in \mathcal{M}, A \subseteq B \Longrightarrow B - A \in \mathcal{M}$
3. $(A_n) \subseteq \mathcal{M}, (A_n) \uparrow \Longrightarrow \bigcup_n A_n \in \mathcal{M}$

Remarques
Si $\mathcal{M}$ est une classe monotone alors $A \in \mathcal{M} \Longrightarrow A^c = \Omega - A \in \mathcal{M}$ donc $\Omega \in \mathcal{M}$ et $(A_n) \subseteq \mathcal{M}, (A_n) \downarrow \Longrightarrow \cap_n A_n \in \mathcal{M}$
Si un δ - système $\mathcal{M}$ est aussi un π - système alors $\mathcal{M}$ est une tribu. Enfin toute intersection de classes monotones est une classe monotone.

1.5.2 Lemme des classes monotones

Si une classe monotone $\mathcal{M}$ contient une classe $\mathcal{S}$ stable par intersection finie alors $\mathcal{M}$ contient $\sigma(\mathcal{S})$. Plus précisément si $\mu(\mathcal{S})$ est la plus petite classe monotone contenant $\mathcal{S}$ alors $\mu(\mathcal{S}) = \sigma(\mathcal{S})$.

Démonstration
On veut montrer $\mu(\mathcal{S}) = \sigma(\mathcal{S})$. Comme $\mu(\mathcal{S}) \subseteq \sigma(\mathcal{S})$ il suffit de montrer que $\mu(\mathcal{S})$ est une tribu et pour cela de vérifier que $\mu(\mathcal{S})$ est stable par intersection. On pose :

$$\mathcal{M}_1 = \{B \in \mu(\mathcal{S}) \mid \forall A \in \mathcal{S}, B \cap A \in \mu(\mathcal{S})\}$$

On vérifie que $\mathcal{M}_1$ est une classe monotone contenant $\mathcal{S}$ donc $\mu(\mathcal{S}) \subseteq \mathcal{M}_1$. On pose alors :

$$\mathcal{M}_2 = \{B \in \mu(\mathcal{S}) \mid \forall A \in \mu(\mathcal{S}), B \cap A \in \mu(\mathcal{S})\}$$

On vérifie que $\mathcal{M}_2$ est une classe monotone contenant $\mathcal{S}$ donc $\mu(\mathcal{S}) \subseteq \mathcal{M}_2$ et par conséquent $\mu(\mathcal{S}) = \mathcal{M}_2$ ce qui signifie que $\mu(\mathcal{S})$ est stable par intersection. $\square$

Le premier théorème des classes monotones ci-dessous est la traduction fonctionnelle du lemme §1.5.2. La classe monotone $\mathcal{M}$ devient un espace vectoriel $\mathcal{H}$ d'applications de Ω dans $\mathbb{R}$.

1.5.3 Premier théorème des classes monotones (TCM1)

Soit Ω un ensemble non vide, $\mathcal{S}$ un π - système sur Ω et $\mathcal{H}$ un espace vectoriel d'applications de Ω dans $\mathbb{R}$ qui contient les constantes et les indicatrices des éléments de $\mathcal{S}$. Dans ces conditions :

1. Si $\mathcal{H}$ contient les limites des suites dans $\mathcal{H}$ croissantes positives et convergentes alors $\mathcal{H}$ contient les applications $\sigma(\mathcal{S})$ mesurables de Ω dans $\mathbb{R}$.
2. Si $\mathcal{H}$ contient les limites des suites dans $\mathcal{H}$ croissantes positives et bornées alors $\mathcal{H}$ contient les applications $\sigma(\mathcal{S})$ mesurables et bornées de Ω dans $\mathbb{R}$.

Remarques
Si (f_n) est une suite croissante d'applications de Ω dans $\mathbb{R}$ il existe toujours une application f de Ω dans $\mathbb{R} \cup \{\infty\}$ telle que $f = \lim_n f_n$ au sens de la convergence simple. On dira que (f_n) est convergente si la limite f est finie c'est-à-dire à valeur dans $\mathbb{R}$. Si (f_n) est croissante et bornée par $K \in \mathbb{R}$ alors (f_n) est convergente et sa limite f est aussi bornée par K.

Démonstration du TCM1
La classe $\mathcal{M} = \{A \subseteq \Omega \mid 1_A \in \mathcal{H}\}$ est une classe monotone contenant le π - système $\mathcal{S}$ donc $\mathcal{M}$ contient $\sigma(\mathcal{S})$. Par conséquent $\mathcal{H}$ contient les indicatrices des éléments de $\sigma(\mathcal{S})$ donc aussi les fonctions étagées sur $\sigma(\mathcal{S})$ puisque $\mathcal{H}$ est un espace vectoriel. Le TCM1 s'obtient alors par limite croissantes et décomposition en parties positives et négatives des applications de Ω dans $\mathbb{R}$. $\square$

Le deuxième théorème des classes monotones ci-dessous est une généralisation du TCM1. La classe des indicatrices d'un π - système $\mathcal{S}$ est ici une partie $\mathcal{G}$ de $\mathcal{H}$ stable par produit. Cette généralisation nécessite une démonstration plus complexe qui utilise le théorème de Weierstrass. Notons que l'espace $\mathcal{H}$ est ici constitué d'applications bornées.

1.5.4 Deuxième théorème des classes monotones (TCM2)

Soit Ω un ensemble non vide et $\mathcal{H}$ un espace vectoriel d'applications bornées de Ω dans $\mathbb{R}$ qui contient les constantes et les limites des suites de $\mathcal{H}$ croissantes et bornées. Si $\mathcal{G}$ est une partie de $\mathcal{H}$ stable par produit alors $\mathcal{H}$ contient les applications $\mathcal{G}$ - mesurables et bornées de Ω dans $\mathbb{R}$.

Démonstration
On montre d'abord (lemme 1) que $\mathcal{H}$ est stable pour la convergence uniforme. On montre ensuite (lemme 2) que si $\phi(x_1, \dots x_d)$ est une fonction continue de $\mathbb{R}^d$ dans $\mathbb{R}$ et si $g_1, \dots g_d$ est une famille d'applications de $\mathcal{G}$ alors $\phi(g_1, \dots g_d)$ est une application de $\mathcal{H}$. On montre enfin (lemme 3) que $1(g_1 > a_1, \dots g_d > a_d) \in \mathcal{H}$
La classe $\mathcal{S}$ des événements de la forme $(g_1 > a_1, \dots g_d > a_d)$ est stable par intersection et $1(g_1 > a_1, \dots g_d > a_d) \in \mathcal{H}$. Il en resulte d'après TCM1 que l'espace $\mathcal{H}$ contient les applications bornées et mesurables pour la tribu $\sigma(\mathcal{S}) = \sigma(\mathcal{G})$. $\square$

Lemme 1
Montrons que $\mathcal{H}$ contient les limites des suites de $\mathcal{H}$ uniformément convergentes.

Démonstration
Soit $(f_n) \subseteq \mathcal{H}$ uniformément convergente de limite f. Posons $a_n = ||f_{n+1} - f_n||$ ou $||f|| = \sup |f(\omega)|$ on peut supposer (en prenant une sous suite) que la série $\sum_n a_n$ est convergente. Soit $r_n = \sum_{k \geq n} a_n$le reste de cette série et $g_n = f_n - r_n$. La suite (g_n) est incluse dans $\mathcal{H}$, croissante et bornée, sa limite f appartient donc à $\mathcal{H}$. $\square$

Lemme 2
Si $\phi(x_1, \dots x_d)$ une fonction continue de $\mathbb{R}^d$ dans $\mathbb{R}$ et si $g_1, \dots g_d$ est une famille d'applications de $\mathcal{G}$ alors $\phi(g_1, \dots g_d)$ est une application de $\mathcal{H}$.

Démonstration
D'après le théorème de Weierstrass la fonction continue $\phi(x_1, \dots x_d)$ est la limite uniforme dans toute boule centrée fermée d'une suite $[P_n(x_1, \dots x_d)]$ de polynomes sur $\mathbb{R}^d$. Comme les applications g_i sont bornées $\phi(g_1, \dots g_d)$ est la limite uniforme de la suite $[P_n(g_1, \dots g_d)]$ qui est incluse dans $\mathcal{H}$ elle appartient donc à $\mathcal{H}$. $\square$

Lemme 3
Si $g_1, \dots g_d$ est une famille d'applications de $\mathcal{G}$ alors $1(g_1 > a_1, \dots g_d > a_d) \in \mathcal{H}$.

Démonstration
On pose : $\phi_n(x_1, \dots x_d) = \prod_{i=1}^d (n(x_i - a_i)^+ \wedge 1)$ on a :

$$\lim_n \uparrow \phi_n(x_1, \dots x_d) = 1(x_1 > a_1 \dots x_d > a_d)$$

L'application ϕ_n est continue. La suite (ϕ_n) est croissante, bornée donc
$\lim_n \uparrow \phi_n(g_1, \dots g_d) = 1(g_1 > a_1 \dots g_d > a_d) \in \mathcal{H}$ $\square$

Le corollaire relaxe l'hypothèse $\mathcal{H}$ borné et distingue la stabilité de $\mathcal{H}$ par limite croissantes et bornées ou croissantes et convergentes.

1.5.5 Corollaire du TCM2

Soit Ω un ensemble non vide et $\mathcal{H}$ un espace vectoriel d'applications (pas nécéssairement bornées) de Ω dans $\mathbb{R}$ qui contient les constantes et une partie $\mathcal{G}$ stable par produit.

1. Si $\mathcal{H}$ contient les limites des suites de $\mathcal{H}$ croissantes et bornées alors $\mathcal{H}$ contient les applications $\sigma(b\mathcal{G})$ mesurables et bornées de Ω dans $\mathbb{R}$.
2. Si $\mathcal{H}$ contient les limites des suites de $\mathcal{H}$ croissantes et convergente alors $\mathcal{H}$ contient les applications $\sigma(b\mathcal{G})$ mesurables de Ω dans $\mathbb{R}$.

Rappelons qu'on note $b\mathcal{G}$ la partie de $\mathcal{G}$ constituée des applications bornées.

Démonstration

Dans le premier cas $b\mathcal{H}$ contient les limites de toute suite croissante et bornée de $b\mathcal{H}$ donc d'après le TCM2 $b\mathcal{H}$ contient $b\sigma(b\mathcal{G})$.

Dans le deuxième cas et en application du premier $b\mathcal{H}$ contient $b\sigma(b\mathcal{G})$ mais alors $\mathcal{H}$ contient $\sigma(b\mathcal{G})$ car si $f \in \sigma(b\mathcal{G})$ est positif $f \wedge n \in b\sigma(b\mathcal{G}) \subseteq b\mathcal{H}$ et donc $f \in \mathcal{H}$ et sinon on décompose f en partie positive et négative. $\square$

Remarque

Si pour tout $g \in \mathcal{G}$ il existe une suite $(g_n) \subseteq b\mathcal{G}$ telle que $g_n \to g$ on a $\sigma(b\mathcal{G}) = \sigma(\mathcal{G})$. C'est le cas courant d'emploi du TCM2 mais ce n'est pas le cas général par exemple si $\mathcal{G}$ est l'espace des polynômes sur $\mathbb{R}$ alors $\sigma(\mathcal{G}) = \mathcal{B}(\mathbb{R})$ mais $b\mathcal{G}$ est réduit aux polynômes constants donc $\sigma(b\mathcal{G}) = \{\emptyset, \Omega\}$.

Exemple : égalité des probabilités et égalité des moments

Examinons comment s'applique le corollaire du TCM2 à l'étude de l'égalité des probabilités sur $\mathbb{R}$ de mêmes moments.

Soient P et Q des probabilités sur $\mathbb{R}$ de même moments c'est-à-dire :

$$\forall k \geq 1, \int_{\mathbb{R}} x^k P(dx) = \int_{\mathbb{R}} x^k Q(dx)$$

On voudrait montrer $P = Q$. Notons $\mathcal{H}$ l'espace des fonctions h mesurables sur $\mathbb{R}$, intégrables pour P et Q et vérifiant

$$\int_{\mathbb{R}} h(x)P(dx) = \int_{\mathbb{R}} h(x)Q(dx)$$

Pour montrer $P = Q$ il faudrait que $\mathcal{H}$ contienne les fonctions mesurables bornées sur $\mathbb{R}$. La classe $\mathcal{H}$ est un espace vectoriel contenant les constantes contenant les limites des suites croissantes bornées dans $\mathcal{H}$ et contenant la classe $\mathcal{G}$ des polynômes. Le corollaire du TCM2 montre donc $b\sigma(b\mathcal{G}) \subseteq \mathcal{H}$ mais $b\mathcal{G}$ est réduit aux polynômes constants donc $\sigma(b\mathcal{G}) = \{\emptyset, \Omega\}$ ce qui ne permet pas de conclure et heureusement car il existe des probabilités distinctes sur $\mathbb{R}$ de même moment.

Toutefois le résultat est vrai sur le segment $[a, b]$ où les polynomes sont bornés car alors $b\sigma(b\mathcal{G}) = b\sigma(\mathcal{G}) = b\mathcal{B}([a, b])$ donc $b\mathcal{B}([a, b]) \subseteq \mathcal{H}$.

Le troisième théorème des classes monotones ci-dessous est adapté aux besoins de la théorie de l'intégration stochastique du chapitre 11. Dans ce résultat $\mathcal{H}$ n'est pas à priori un espace vectoriel mais il contient une algèbre $\mathcal{G}$.

1.5.6 Troisième théorème des classes monotones (TCM3)

Soit $(\Omega, \mathcal{F})$ un espace mesurable et $\mathcal{H}$ une classe d'applications mesurables bornées de Ω dans $\mathbb{R}$ fermée pour la convergence monotone bornée et la convergence uniforme. S'il existe une algèbre $\mathcal{G} \subseteq \mathcal{H}$ d'applications de $\mathcal{H}$ telle que $\sigma(\mathcal{G}) = \mathcal{F}$ et si $\mathcal{G}$ contient une suite croissante de limite 1 alors $\mathcal{H}$ contient toutes les applications mesurables bornées de Ω dans $\mathbb{R}$, soit $\mathcal{H} = b\mathcal{F}$.

Démonstration
Soit $\mathcal{K}$ la plus petite classe contenant $\mathcal{G}$ et fermée pour la convergence monotonne bornée (CMB) et la convergence uniforme (CU). On a $\mathcal{G} \subseteq \mathcal{K} \subseteq \mathcal{H} \subseteq b\mathcal{F}$. On montre (lemme 1 ci dessous) que $\mathcal{K}$ est une algèbre unitaire puis on pose :

$$\Gamma = \{A \in \mathcal{F} \mid 1_A \in \mathcal{K}\} \text{ et } \mathcal{A} = \{(g \leq a) \mid g \in \mathcal{G}, a \in \mathbb{R}\}$$

On vérifie :

1. $\Gamma \subseteq \mathcal{F}$ est une tribu (car $\mathcal{K}$ est une algèbre unitaire fermée pour la CMB)
2. $b\Gamma \subseteq \mathcal{K}$ (car $\mathcal{K}$ contient les fonctions étagées sur Γ et est fermée pour la CMB)
3. $\sigma(\mathcal{A}) = \sigma(\mathcal{G})$ (par définition de $\mathcal{A}$)
4. $\sigma(\mathcal{A}) = \mathcal{F}$ (d'après 3 car par hypothèse $\sigma(\mathcal{G}) = \mathcal{F}$))

et on établit (lemme 2 ci dessous) que $\mathcal{A} \subseteq \Gamma$ on a alors

$$\mathcal{F} =_4 \sigma(\mathcal{A}) \subseteq \sigma(\Gamma) =_1 \Gamma \subseteq \mathcal{F} \text{ donc } b\mathcal{F} = b\Gamma \subseteq_2 \mathcal{K} \subseteq \mathcal{H} \subseteq b\mathcal{F}$$

D'où $\mathcal{H} = b\mathcal{F}$ $\square$

Lemme 1
$\mathcal{K}$ est une algèbre unitaire.

Démonstration
Montrons d'abord que $\mathcal{K}$ est un espace vectoriel, posons pour tout $f \in b\mathcal{F}$:

$$\mathcal{K}_0(f) = \{k \in \mathcal{K} \mid \forall a, b \in \mathbb{R}, af + bk \in \mathcal{K}\}$$

Il vient :

1. $g \in \mathcal{G} \Longrightarrow \mathcal{K} = \mathcal{K}_0(g)$
   car $\mathcal{G} \subseteq \mathcal{K}_0(g) \subseteq \mathcal{K}$ et $\mathcal{K}_0(f)$ est fermée pour la CMB et la CU.
2. $k \in \mathcal{K} \Longrightarrow \mathcal{K} = \mathcal{K}_0(k)$
   car $\mathcal{G} \subseteq \mathcal{K}_0(k) \subseteq \mathcal{K}$ et $\mathcal{K}_0(k)$ est fermée pour la CMB et la CU.
3. $k_1, k_2 \in \mathcal{K} \Longrightarrow k_2 \in \mathcal{K}_0(k_1) \Longrightarrow \forall a, b \in \mathbb{R}, ak_1 + b_2 \in \mathcal{K}$

(3) signifie que $\mathcal{K}$ est un espace vectoriel.
Montrons que $\mathcal{K}$ est une algèbre unitaire, posons pour tout $f \in b\mathcal{F}$:

$$\mathcal{K}_1(f) = \{k \in \mathcal{K} \mid kf \in \mathcal{K}\}$$

Il vient comme précédamment :

1. $g \in \mathcal{G} \Longrightarrow \mathcal{K} = \mathcal{K}_1(g)$
2. $k \in \mathcal{K} \Longrightarrow \mathcal{K} = \mathcal{K}_1(k)$
3. $k_1, k_2 \in \mathcal{K} \Longrightarrow k_2 \in \mathcal{K}_1(k_1) \Longrightarrow k_1 k_2 \in \mathcal{K}$

(3) signifie que $\mathcal{K}$ est une algèbre. Comme 1 est limite croissante d'une suite de $\mathcal{G} \subseteq \mathcal{K}$ $\mathcal{K}$ est unitaire. $\square$

Lemme 2
$\mathcal{A} \subseteq \Gamma$ c'est-à-dire $1_A \in \mathcal{K}$ où $A = (g \leq a)$ et $g \in \mathcal{G}$

Démonstration
Posons $\phi(x) = 1$ pour $x \leq a$, $\phi(x) = 0$ pour $x \geq a + 1$ et $\phi(x) = 1 - x + a$ pour $a < x < a+1$. On a $\phi(g)^n \to 1_A$ Soit c tel que $|g| \leq c$, il existe une suite de polynomes (p_n) qui converge uniformément vers ϕ sur $[-c, c]$ donc $\phi(g)^n \in \mathcal{K}$ et $1_A \in \mathcal{K}$. $\square$

Chapitre 2

Espérance

On définit l'espérance et l'espérance conditionnelle des variables aléatoires en appliquant une méthode générale d'extension d'opérateurs décrite au premier paragraphe. Dans cette méthode dérivée de la construction de l'intégrale de Daniel on suppose défini au départ un opérateur (application linéaire) p croissant sur un sous espace vectoriel *réticulé* $\mathcal{E}$ de l'algèbre $\mathcal{F}$ des applications mesurables de $(\Omega, \mathcal{F})$ dans $\mathbb{R}$ et à valeurs dans $\mathcal{F}$. On suppose que l'espace $\mathcal{E}$ satisfait la condition (1) et l'opérateur p la condition (2) de la définition §2.1.1 ci-dessous. On étend alors cet opérateur au cône $\mathcal{F}^+$ des applications mesurables positives puis par décomposition en parties positive et négative à la classe des applications mesurables X pour lesquelles la différence $p(X^+) - p(X^-)$ est définie dans $\mathbb{R}^+ \cup \{-\infty, \infty\}$. On montre alors pour l'application p des résultats équivalents au théorème de convergence monotone, aux inégalités de Fatou et au théorème de convergence dominée.

2.1 Extension d'opérateurs de variables aléatoires

Rappels et notations
Soit $(\Omega, \mathcal{F})$ un espace mesurable, rappelons (voir §1.3.1) qu'une variable aléatoire (VA) est une application mesurable de $(\Omega, \mathcal{F})$ dans $\mathbb{R} \cup \{-\infty, \infty\}$ muni de sa tribu borélienne, qu'une variable aléatoire positive est une VA à valeur dans $\mathbb{R}^+ \cup \{\infty\}$ et qu'une variable aléatoire réelle (VAR) est une VA à valeur dans $\mathbb{R}$. On note aussi $\mathcal{F}$ l'algèbre des VAR et $\mathcal{F}^+$ le cône des VA positives.

2.1.1 Définition : opérateur p de $\mathcal{E}$ dans $\mathcal{F}$

Soit $\mathcal{E}$ un sous espace vectoriel *réticulé* de l'algèbre $\mathcal{F}$ c'est-à-dire un sous espace de $\mathcal{F}$ stable par min et max : $\forall X, Y \in \mathcal{E},\ X \wedge Y = \min(X, Y) \in \mathcal{E},\ X \vee Y = \max(X, Y) \in \mathcal{E}$
Soit p un opérateur *croissant* de $\mathcal{E}$ dans $\mathcal{F}$ c'est-à-dire une application linéaire de $\mathcal{E}$ dans $\mathcal{F}$ telle que : $\forall X, Y \in \mathcal{E},\ X \leq Y \Longrightarrow p(X) \leq p(Y)$
Hypothèses sur le sous espace $\mathcal{E}$ et l'opérateur p :

1. Pour toute VA positive X il existe une suite croissante (X_n) dans $\mathcal{E}$ de limite X dans $\mathcal{F}^+$ c'est-à-dire $\lim_n \uparrow X_n(\omega) = X(\omega)$ dans $\mathbb{R}^+ \cup \{\infty\}$ pour tout $\omega \in \Omega$.

2. Pour toute suite (X_n) dans $\mathcal{E}$ décroissante vers 0 la suite $[p(X_n)]$ est décroissante vers 0.

Remarque : convergence monotone dans $\mathcal{E}$
D'après l'hypothèse (2) plus généralement pour toutes suites (X_n) dans $\mathcal{E}$ monotones de limite X dans $\mathcal{E}$ la suite $[p(X_n)]$ est monotone de limite $p(X)$.

Exemples
1) L'espérance $E(.)$ est définie initialement au §2.3.1 sur l'espace réticulé $\mathcal{E}$ des VAR étagées sur $\mathcal{F}$. L'espace $\mathcal{E}$ vérifie (1) de §2.1.1 et l'espérance $E(.)$ est un opérateur croissant de $\mathcal{E}$ dans l'espace isomorphe à $\mathbb{R}$ des VAR constantes. L'espérance vérifie (2) de §2.1.1. Elle est étendue selon le procédé ci-dessous au cône $\mathcal{F}^+$ des VA positives puis à la classe des VA X quasi intégrables pour lesquelles la différence $E(X^+) - E(X^-)$ où $X^+ = X \vee 0$ et $X^- = -(X \wedge 0)$ est définie dans $\mathbb{R}^+ \cup \{-\infty, \infty\}$ et en particulier à l'espace des VAR intégrables pour lesquelles $E(X^+) - E(X^-)$ est définie dans $\mathbb{R}$ (voir §2.3)

2) Si $\mathcal{G}$ est une sous tribu de $\mathcal{F}$ l'espérance conditionnelle $E(. \mid \mathcal{G})$ est définie initialement au §2.5.1 sur l'espace réticulé $L^2(\mathcal{F})$ des VAR de carré intégrable. L'espace $L^2(\mathcal{F})$ vérifie (1) de §2.1.1 et l'espérance conditionnelle est un opérateur croissant de $L^2(\mathcal{F})$ dans l'espace $L^2(\mathcal{G}) \subseteq L^2(\mathcal{F})$. Elle vérifie (2) de §2.1.1. On confondra une VAR X et sa classe d'équivalence (voir §1.3.2). L'espérance conditionnelle est étendue selon le procédé ci-dessous au cône $\mathcal{F}^+$ des VA positives puis à la classe des VA X ayant une espérance conditionnelle c'est-à-dire telles que $E(X^+ \mid \mathcal{G}) - E(X^- \mid \mathcal{G})$ soit définie dans $\mathbb{R}^+ \cup \{-\infty, \infty\}$ (voir §2.6).

3) Les projecteurs optionnels et prévisibles ${}^o(.)$ et ${}^p(.)$ sont définis initialement au §9.1.3 sur l'espace réticulé $b(\mathcal{B} \otimes \mathcal{F})$ où $\mathcal{B} = \mathcal{B}(\mathbb{R}^+)$ des processus mesurables bornés. L'espace $b(\mathcal{B} \otimes \mathcal{F})$ vérifie (1) de §2.1.1. Ce sont des opérateurs croissants de $b(\mathcal{B} \otimes \mathcal{F})$ dans $b\mathcal{O}$ et $b\mathcal{P}$ (espace des processus bornés, optionnels ou prévisibles) qui vérifient (2) de §2.1.1. Ils sont étendus selon le procédé ci-dessous au cône $(\mathcal{B} \otimes \mathcal{F})^+$ puis aux classes des processus X mesurables de $(\mathbb{R}^+ \times \Omega, \mathcal{B} \otimes \mathcal{F})$ dans $\mathbb{R} \cup \{-\infty, \infty\}$ ayant une projections optionnelle ou prévisible c'est-à-dire tels que ${}^oX^+ - {}^oX^-$ ou ${}^pX^+ - {}^pX^-$ soient définis dans $\mathbb{R} \cup \{-\infty, \infty\}$ (voir §9.2). $\square$

2.1.2 Lemme : comparaison des suites croissantes dans $\mathcal{E}$

Si (X_n) et (Y_n) sont des suites croissantes dans $\mathcal{E}$ alors

$$\lim_n \uparrow X_n \leq \lim_n \uparrow Y_n \Longrightarrow \lim_n \uparrow p(X_n) \leq \lim_n \uparrow p(Y_n)$$

En particulier $\lim_n \uparrow X_n = \lim_n \uparrow Y_n \Longrightarrow \lim_n \uparrow p(X_n) = \lim_n \uparrow p(Y_n)$

Démonstration

1. $X_n \leq \lim_m \uparrow X_m \leq \lim_m \uparrow Y_m$ (hypothèses)
2. $\lim_m \uparrow (X_n \wedge Y_m) = X_n \wedge \lim_m \uparrow Y_m = X_n$
$\Longrightarrow \lim_m \uparrow p(X_n \wedge Y_m) = p(X_n)$ (convergence monotone dans $\mathcal{E}$)
3. $p(X_n \wedge Y_m) \leq p(Y_m) \Longrightarrow \lim_m \uparrow p(X_n \wedge Y_m) \leq \lim_m \uparrow p(Y_m)$
4. $p(X_n) \leq \lim_m \uparrow p(Y_m) \Longrightarrow \lim_n \uparrow p(X_n) \leq \lim_m \uparrow p(Y_m)$ $\square$

2.1.3 Définition et proposition : opérateur p de $\mathcal{F}^+$ dans $\mathcal{F}^+$

Soit $(\Omega, \mathcal{F})$ un espace mesurable, $\mathcal{E}$ un sous espace réticulé de l'espace $\mathcal{F}$ des VAR et p un opérateur croissant de $\mathcal{E}$ dans $\mathcal{F}$. On suppose que $\mathcal{E}$ et p vérifient (1) et (2) de §2.1.1. D'après (1) de §2.1.1 pour tout $X \in \mathcal{F}^+$ (cône des VA positives) il existe une suite $(X_n) \subseteq \mathcal{E}$ telle que $X_n \uparrow X$ on pose alors : $p(X) = \lim_n \uparrow p(X_n)$

Cette définition est indépendante de la suite (X_n) d'après le lemme §2.1.2 ci-dessus. On définit ainsi une première extension de l'opérateur p de $\mathcal{E}$ à $\mathcal{F}^+$. Cette extension est croissante, linéaire dans le cône $\mathcal{F}^+$, c'est-à-dire :

$$\forall X, Y \in \mathcal{F}^+, \ \forall a, b \in \mathbb{R}^+, \ p(aX + bY) = ap(X) + bp(X)$$

et vérifie le théorème de convergence croissante dans $\mathcal{F}^+$:

$$\forall (X_n) \subseteq \mathcal{F}^+, X \in \mathcal{F}^+, \ X_n \uparrow X \Longrightarrow p(X_n) \uparrow p(X)$$

Démonstration

Montrons la convergence croissante dans $\mathcal{F}^+$. Soit $(X_n) \subseteq \mathcal{F}^+$ une suite croissante de VA positives. Pour tout X_n il existe une suite croissante $(X_{n,m})$ dans $\mathcal{E}$ telle que $\lim_m \uparrow X_{n,m} = X_n$. Posons $Y_m = \max_{n \leq m} X_{n,m}$. On vérifie que $\lim_m \uparrow Y_m = X$ on en déduit :

1. $Y_m \leq X_m \Longrightarrow p(X) = \lim_m \uparrow p(Y_m) \leq \lim_m \uparrow p(X_m)$
2. $X_n \leq X \Longrightarrow p(X_n) \leq p(X) \Longrightarrow \lim_n \uparrow p(X_n) \leq p(X)$

Finalement $\lim_n \uparrow p(X_n) = p(X)$ $\square$

2.1.4 Lemme : propriétés de X^+ et X^-

Pour toutes VAR X, Y on note $X^+ = X \vee 0$ et $X^- = -(X \wedge 0)$ on a :

1. $X^- = (-X)^+$ et $X^+ = (-X)^-$
2. $(X + Y)^+ \leq X^+ + Y^+$ et $(X + Y)^- \leq X^- + Y^-$
3. $X \leq Y \Longleftrightarrow X^+ \leq Y^+$ et $X^- \geq Y^-$
4. $(X \wedge Y)^+ \leq (X \vee Y)^+ \leq X^+ + Y^+$ et $(X \vee Y)^- \leq (X \wedge Y)^- \leq X^- + Y^-$
5. $|X| = X^+ + X^-$

Démonstration

Premier point :

— $X \geq 0 \Longrightarrow X^- = 0$ et $-X \leq 0 \Longrightarrow (-X)^+ = 0 \Longrightarrow X^- = (-X)^+$
— $X \leq 0 \Longrightarrow X^- = -X \geq 0 \Longrightarrow (-X)^+ = -X \Longrightarrow X^- = (-X)^+$
— $X^- = (-X)^+ \Longleftrightarrow (-X)^- = X^+$

Deuxième point :

— $X \geq 0, Y \geq 0 \Longrightarrow (X+Y)^+ = X+Y = X^+ + Y^+$ et $(X+Y)^- = X^- = Y^- = 0$
— $X \leq 0, Y \leq 0 \Longrightarrow (X+Y)^+ = X^+ = Y^+ = 0$ et $(X+Y)^- = X+Y = X^- + Y^-$
— $X \geq 0, Y \leq 0, X + Y \geq 0 \Longrightarrow (X + Y)^+ \leq X^+ \leq X^+ + Y^+$ et $(X + Y)^- = 0 \leq X^- + Y^-$
— $X \geq 0, Y \leq 0, X + Y \leq 0 \Longrightarrow (X + Y)^+ = 0 \leq X^+ + Y^+$ et $(X + Y)^- \leq Y^- \leq X^- + Y^-$

Troisième point :

— $X \geq 0, Y \geq 0 \Longrightarrow X^+ = X, Y^+ = Y, X^- = Y^- = 0 \Longrightarrow (3)$
— $X \leq 0, Y \leq 0 \Longrightarrow X^+ = Y^+ = 0, X^- = X, Y^- = Y \Longrightarrow (3)$
— $X \geq 0, Y \leq 0 \Longrightarrow (X \leq Y \Longleftrightarrow X = Y = 0 \Longleftrightarrow X^+ \leq Y^+, X^- \geq Y^-)$ (3)
— $X \leq 0, Y \geq 0 \Longrightarrow X \leq Y, X^+ \leq Y^+, X^- \geq Y^- \Longrightarrow (3)$

Quatrième point, succesivement :

— $X \wedge Y \leq X \vee Y \Longrightarrow (X \wedge Y)^+ \leq (X \vee Y)^+$
— $(X \vee Y)^+ \leq X^+ + Y^+$ on distingue $X \geq Y$ et $Y \geq X$
— $(X \wedge Y)^+ \leq (X \vee Y)^+ \leq X^+ + Y^+$ c'est le point 4.1
— $X \vee Y \geq X \wedge Y \Longrightarrow -(X \vee Y) \leq -(X \wedge Y)$
— $(X \vee Y)^- = [-(X \vee Y)]^+ \leq [-(X \wedge Y)]^+ = (X \wedge Y)^-$
— $(X \wedge Y)^- \leq X^- + Y^-$ on distingue $X \leq Y$ et $Y \leq X$
— $(X \vee Y)^- \leq (X \wedge Y)^- \leq X^- + Y^-$ c'est le point 4.2

Cinquième point : on distingue $X \geq 0$ et $X \leq 0$ $\square$

2.1.5 Définition et proposition : variables aléatoires opérables

Si X est une VA c'est-à-dire une application X mesurable de $(\Omega, \mathcal{F})$ dans $\mathbb{R} \cup \{-\infty, \infty\}$ on décompose X en parties positive et négative X^+ et X^-. On dit que X est *opérable* par p si les deux applications $p(X^+)$ et $p(X^-)$ de $\mathcal{F}^+$ ne sont pas simultanément infinies et dans ce cas l'opérateur p est défini pour X par : $p(X) = p(X^+) - p(X^-)$ qui est à valeur dans $\mathbb{R} \cup \{-\infty, \infty\}$.
Si X est une VA opérable par p et si les deux applications $p(X^+)$ et $p(X^-)$ sont finies on dit que X est opérable par p dans $\mathbb{R}$ dans ce cas $p(X) \in \mathcal{F}$.
Si X et Y sont des VAR opérables par p dans $\mathbb{R}$ alors $X + Y, X \vee Y, X \wedge Y$ et $|X|$ le sont aussi d'après les majorations (2) et (4) et l'égalité (5) du lemme §2.1.4.
On note $\mathcal{F}_p$ le sous espace vectoriel réticulé de $\mathcal{F}$ constitué des VAR opérables par p dans $\mathbb{R}$.

2.1.6 Définition et proposition : négligeables pour p dans $\mathcal{F}$

Si $A \in \mathcal{F}$ alors $1_A \in \mathcal{F}^+$ et on dit que A est *négligeable pour* p si $p(1_A) = 0$. Avec la convention $\infty 0 = 0$ on a :

1. Si $A \in \mathcal{F}$ alors $p(\infty 1_A)$ est soit nul soit infini selon que A est négligeable pour p ou non.
2. Si X est une VA opérable par p et si A est négligeable pour p alors $p(X1_A) = 0$ et $p(X1_{A^c}) = p(X)$
3. Si X et Y sont des VA égales à un négligeable pour p près alors X est opérable par p si et seulement si Y l'est et dans ce cas $p(X) = p(Y)$
4. Si X est une VA opérable par p dans $\mathbb{R}$ alors $(|X| = \infty)$ est négligeable pour p
5. Une VA X est opérable par p dans $\mathbb{R}$ si et seulement si elle est égale à un négligeable pour p près à une VAR Y opérable par p dans $\mathbb{R}$.

Démonstration

1. Comme $\infty 1_A = \lim_n \uparrow n 1_A$ donc $p(\infty 1_A) = \lim_n \uparrow p(n 1_A) = \lim_n \uparrow np(1_A)$ on a $p(1_A) = 0 \Longrightarrow p(\infty 1_A) = 0$ et $p(1_A) > 0 \Longrightarrow p(\infty 1_A) = \infty$
2. Pour $X \in \mathcal{F}^+$ et A négligeable on a $p(X 1_A) \leq p(\infty 1_A) = 0$ ce qui implique $p(X) = p(X 1_A) + p(X 1_{A^c}) = p(X 1_{A^c})$. Pour une VA X opérable par p on décompose X en parties positive et négative.
3. Le troisième point découle du second en posant $A = (X \neq Y)$
4. Pour $X \in \mathcal{F}^+$ comme $p(X) = p(X 1_{X<\infty}) + p(X 1_{X=\infty})$ l'hypothèse $p(X) < \infty$ implique $p(X 1_{X=\infty}) < \infty$ donc $p(X 1_{X=\infty}) = p(\infty 1_{X=\infty}) = 0$ et par conséquent $(X = \infty)$ est négligeable pour p. Pour une VA X opérable par p on décompose X en parties positive et négative.
5. Si une VA X est opérables par p dans $\mathbb{R}$ alors $(|X| = \infty)$ est négligeable pour p et $P(X) = p(X 1_{|X|<\infty})$ donc $Y = X 1_{|X|<\infty}$ est une VAR opérable par p dans $\mathbb{R}$ égale à X à un négligeable pour p près. Inversement si X et Y sont égales à un négligeable pour p près alors $p(X) = p(Y)$ donc si Y est opérable par p dans $\mathbb{R}$ il en est de même de X.

2.1.7 Théorème : croissance de p et linéarité dans $\mathcal{F}_p$

L'application p est croissante sur la classe des VA opérables par p et sa restriction au sous espace vectoriel $\mathcal{F}_p$ de $\mathcal{F}$ constitué des VAR opérables par p dans $\mathbb{R}$ est un opérateur croissant de $\mathcal{F}_p$ dans $\mathcal{F}$. Notons que $P(X^+) \geq P(X)^+$ et $P(X^-) \geq P(X)^-$

Démonstration
La croissance de p sur la classe des VA opérables par p résulte de la croissance de p dans $\mathcal{F}^+$ et de $X \leq Y \Longleftrightarrow X^+ \leq Y^+, X^- \geq Y^-$ (lemme §2.1.4-3)
Montrons la linéarité de p dans $\mathcal{F}_p$, pour toutes VAR $X, Y \in \mathcal{F}_p$ il vient :
$X + Y = (X + Y)^+ - (X + Y)^-$ et $X + Y = (X^+ - X^-) + (Y^+ - Y^-)$
$\Longrightarrow\ (X + Y)^+ + X^- + Y^- = (X + Y)^- + X^+ + Y^+$
$\Longrightarrow\ p[(X + Y)^+] + p(X^-) + p(Y^-) = p[(X + Y)^-] + p(X^+) + p(Y^+)$
$\Longrightarrow\ p(X + Y) = p[(X + Y)^+] - p[(X + Y)^-] = p(X^+) + p(Y^+) - p(X^-) + -p(Y^-)$
$\Longrightarrow\ p(X + Y) = p(X) + p(Y)$
Enfin on vérifie par distinction de cas $\forall a \in \mathbb{R},\ p(aX) = ap(X)$

2.1.8 Théorème de convergence monotone (TCM)

1) Si (X_n) est une suite croissante de VA qui ne part pas de trop bas, ce qui signifie $p(X_0^-) < \infty$ alors les VA X_n et la VA limite $X = \lim_n \uparrow X_n$ sont opérables par p et $p(X) = \lim_n \uparrow p(X_n)$ est définie de $(\Omega, \mathcal{F})$ dans $\mathbb{R} \cup \{\infty\}$.

2) Si (X_n) est une suite décroissante de VA qui ne part pas de trop haut, ce qui signifie $p(X_0^+) < \infty$ alors les VA X_n et la VA limite $X = \lim_n \downarrow X_n$ sont opérables par p et $p(X) = \lim_n \downarrow p(X_n)$ est définie de $(\Omega, \mathcal{F})$ dans $\mathbb{R} \cup \{-\infty\}$.

Lemme
Si (Y_n) est une suite décroissante de VAR positives telle que $p(Y_0) < \infty$ alors la limite $Y = \lim_n \downarrow Y_n$ vérifie $p(Y) = \lim_n \downarrow p(Y_n)$ dans $\mathbb{R}$.

Démonstration du lemme
On pose $Z_n = Y_0 - Y_n$ on a $Z_n \geq 0, Z_n \uparrow Z = Y_0 - Y$ et $p(Z) = \lim_n \uparrow p(Z_n)$ d'après le théorème de convergence croissante dans $\mathcal{F}^+$, il vient alors :

$$\begin{aligned} p(Y) = p(Y_0 - Z) = p(Y_0) - p(Z) = p(Y_0) - \lim_n \uparrow p(Z_n) = \lim_n \downarrow [p(Y_0) - p(Z_n)] \\ = \lim_n \downarrow p(Y_0 - Z_n) = \lim_n \downarrow p(Y_n) \ \square \end{aligned}$$

Démonstration du théorème
On décompose $X_n = X_n^+ - X_n^-$ la croissance de (X_n) entraîne la croissance de (X_n^+) et la décroissance de (X_n^-) d'après le lemme §2.1.4-3.
Pour (X_n^+) on applique le théorème de convergence croissante dans $\mathcal{F}^+$ on obtient

$$p(X^+) = \lim_n \uparrow p(X_n^+) \text{ dans } \mathbb{R}^+ \cup \{\infty\}$$

Pour (X_n^-) on pose $A = (X_0^- = \infty)$ qui est négligeable pour p puisque $p(X_0^-) < \infty$ puis $Y_n = X_n^- 1_{A^c} < \infty$ et $Y = X^- 1_{A^c} < \infty$ où $X^- = \lim_n \downarrow X_n^-$. On a alors $p(Y_n) = p(X_n^-), p(Y) = p(X^-)$ et $p(Y_0) = p(X_0^-) < \infty$ et d'après le lemme ci-dessus $p(Y) = \lim_n \downarrow p(Y_n)$ dans $\mathbb{R}$ donc

$$p(X^-) = \lim_n \downarrow p(X_n^-) \text{ dans } \mathbb{R}$$

Il en résulte $p(X) = \lim_n \uparrow p(X_n)$ dans $\mathbb{R} \cup \{\infty\}$. On étudie de même $X = \lim_n \downarrow X_n$.
$\square$

2.1.9 Proposition : inégalités de Fatou

Si (X_n) est une suite de VA et s'il existe une VA Y telle que $p(Y^-) < \infty$ et $\forall n, Y \leq X_n$ alors les VA X_n et la VA $\liminf_n X_n$ sont opérables par p et

$$p(\liminf_n X_n) \leq \liminf_n p(X_n)$$

Symétriquement s'il existe une VA Z telle que $p(Z^+) < \infty$ et $\forall n, X_n \leq Z$ alors les VA X_n et la VA $\limsup_n X_n$ sont opérables par p et

$$p(\limsup_n X_n) \geq \limsup_n p(X_n)$$

Démonstration
D'abord les VA X_n sont opérables par p puisque $Y \leq X_n \Longrightarrow X_n^- \leq Y^-$ et $p(Y^-) < \infty$. Posons $U_m = \inf_{n \geq m} X_n$, la suite (U_m) est croissante et

$$\forall n, X_n \geq Y \Longrightarrow U_0 \geq Y \Longrightarrow U_0^- \leq Y^- \Longrightarrow p(U_0^-) \leq p(Y^-) < \infty$$

Les VA U_m et la VA limite $U = \lim_m \uparrow U_m$ sont donc opérables par p et on a :

$$\begin{aligned} \forall n \geq m, U_m \leq X_n \Longrightarrow p(U_m) \leq p(X_n) \Longrightarrow p(U_m) \leq \inf\nolimits_{n \geq m} p(X_n) \\ \Longrightarrow \lim\nolimits_m \uparrow p(\inf\nolimits_{n \geq m} X_n) \leq \lim\nolimits_m \uparrow \inf\nolimits_{n \geq m} p(X_n) \end{aligned}$$

Le TCM entraîne alors :

$$p(\liminf_n X_n) = p(\lim_m \uparrow \inf_{n \geq m} X_n \leq \lim_m \uparrow \inf_{n \geq m} p(X_n) = \liminf_n p(X_n) \ \square$$

2.1.10 Théorème de convergence dominée (TCD)

Si (X_n) est une suite de VA telle que $\lim_n X_n = X$ et s'il existe une VA positive Z opérable par p dans $\mathbb{R}$ c'est-à-dire vérifiant $p(Z) < \infty$ qui *domine* la suite (X_n) c'est-à-dire $\forall n, |X_n| \leq Z$ alors les VA X_n et la limite X sont opérables par p dans $\mathbb{R}$ et $p(X) = \lim_n p(X_n)$

Démonstration
On a $\forall n, Y = -Z \leq X_n \leq Z$ avec $p(Y^-) < \infty$ et $p(Z^+) < \infty$ les inégalités de Fatou s'appliquent :

$$p(\liminf_n X_n) \leq \liminf_n p(X_n) \leq \limsup_n p(X_n) \leq p(\limsup_n X_n)$$

Par conséquent $X = \lim_n X_n = \liminf_n X_n = \limsup_n X_n)$
entraîne $p(X) = \lim_n p(X_n)$ $\square$

2.1.11 Proposition : inégalité de Jensen

Si ϕ est une application convexe de $\mathbb{R}$ dans $\mathbb{R}$ telle que $\forall X \in \mathcal{F}_p, \phi(X) \in \mathcal{F}_p$ on a *l'inégalité de Jensen* :

$$\forall X \in \mathcal{F}_p,\ \phi[p(X)] \leq p[\phi(X)]$$

Démonstration
Pour $\phi(x) = ax + b$ et pour tout $X \in \mathcal{F}_p$ on a

$$\phi[p(X)] = ap(X) + b = p(aX + b) = p[\phi(X)]$$

Comme toute fonction convexe est enveloppe supérieure d'une famille dénombrable de fonction affine ϕ_n on a $\phi = \sup_n \phi_n$ donc $\phi_n(X) \leq \phi(X)$. Il vient alors :

$$\forall n, \phi_n[p(X)] = p[\phi_n(x)] \leq p[\phi(X)] \Longrightarrow \phi[p(X)] = \sup_n \phi_n[p(X)) \leq p[\phi(X)] \quad \square$$

2.2 Projecteurs dans un espace vectoriel

2.2.1 Définition : projecteur dans un espace vectoriel

Soit $\mathcal{E}$ un espace vectoriel, on appelle *projecteur* dans $\mathcal{E}$ une application linéaire p de $\mathcal{E}$ dans lui-même (endomorphisme d'espace vectoriel) *idempotente* c'est-à-dire :

$$\forall x \in \mathcal{E},\ p \circ p(x) = p(x)$$

Si $\mathcal{E}'$ est un sous espace de $\mathcal{E}$ on dit que p est projecteur de $\mathcal{E}$ sur $\mathcal{E}'$ si $p(\mathcal{E}) = \mathcal{E}'$.

2.2.2 Proposition : idempotence et invariance de $p(\mathcal{E})$

L'idempotence pour une application p d'un espace vectoriel $\mathcal{E}$ dans lui-même est équivalente à l'invariance de l'espace image $p(\mathcal{E})$ au sens ou ses éléments sont invariants par p, autrement dit la restriction de p à $p(\mathcal{E})$ est l'identité.

De plus si p est un endomorphisme dans $\mathcal{E}$ et si $\mathcal{E}'$ est un sous espace de $\mathcal{E}$ alors p est un projecteur de $\mathcal{E}$ sur $\mathcal{E}'$ si et seulement si :

1. $p(\mathcal{E}) \subseteq \mathcal{E}'$ et
2. $\forall y \in \mathcal{E}', \; p(y) = y \;$ ($\mathcal{E}'$ est invariant par p)

Démonstration
Pour tout $y \in p(\mathcal{E})$ il existe $x \in \mathcal{E}$ tel que $y = p(x)$ donc si p est idempotent on a $p(y) = p[p(x)] = p(x) = y$ et $p(\mathcal{E})$ est invariant. Inversement si $\forall y \in p(\mathcal{E}), p(y) = y$ alors $\forall x \in \mathcal{E}, p[p(x)] = p(x)$ car $p(x) \in p(\mathcal{E})$ donc p est idempotent.
Les assertions (1) et (2) implique l'invariance de $p(\mathcal{E})$ donc l'idempotence de p. Si p est un endomorphisme p est donc un projecteur. Comme (2) implique $\mathcal{E}' \subseteq p(\mathcal{E})$ les assertions (1) et (2) implique $p(\mathcal{E}) = \mathcal{E}'$ il s'agit donc d'un projecteur sur $\mathcal{E}'$.
Inversement si p est un projecteur de $\mathcal{E}$ sur $\mathcal{E}'$ on a $p(\mathcal{E}) = \mathcal{E}'$ et $p(\mathcal{E})$ est invariant puisque p est idempotent, on a donc (1) et (2).

Notations
Dans les propositions suivantes $\mathcal{E}$ est un sous espace vectoriel réticulé de l'algèbre $\mathcal{F}$ des VAR définies sur un espace mesurable $(\Omega, \mathcal{F})$ et p un opérateur croissant de $\mathcal{E}$ dans $\mathcal{F}$. On suppose que $\mathcal{E}$ et p vérifient les hypothèses (1) et (2) du §2.1.1 et on étend p selon §2.1 au cône $\mathcal{F}^+$ des VA positives et à l'espace $\mathcal{F}_p$ des VAR opérables par p dans $\mathbb{R}$. Cette extension est encore notée p.

2.2.3 Proposition : extension de projecteurs

Si l'opérateur p est un projecteur dans $\mathcal{E}$ (endomorphisme idempotent dans $\mathcal{E}$) alors son extension à $\mathcal{F}_p$ est un projecteur dans $\mathcal{F}_p$.

Démonstration
Soit $X \in \mathcal{F}_p$ donc $p(X) \in \mathcal{F}$ il faut montrer $p(X) \in \mathcal{F}_p$ et $p \circ p(X) = p(X)$.
Pour $X \geq 0$ on a $X \in \mathcal{F}^+$ et il existe une suite (X_n) dans $\mathcal{E}$ telle que $X_n \uparrow X$ donc $p(X_n) \uparrow p(X)$ et on a :

$$p \circ p(X) = p[\lim_n \uparrow p(X_n)] = \lim_n \uparrow p \circ p(X_n)] = \lim_n \uparrow p(X_n) = p(X)$$

Comme $p(X) \in \mathcal{F}$ et $p[p(X)] = p(X)$ on a donc $p(X) \in \mathcal{F}_p$.
Pour $X \in \mathcal{F}_p$ de signe quelconque on décompose $X = X^+ - X^-$. $\square$

2.2.4 Proposition : espace image de l'extension d'un projecteur

Soit $\mathcal{E}' \subseteq \mathcal{E}$ un sous espace réticulé de $\mathcal{E}$ et $\mathcal{F}' \subseteq \mathcal{F}_p$ un sous espace réticulé de $\mathcal{F}_p$. Si l'opérateur p est un projecteur de $\mathcal{E}$ sur $\mathcal{E}'$ et si les VAR positives de $\mathcal{F}'$ sont les VAR positives de $\mathcal{F}_p$ limites croissantes de VAR de $\mathcal{E}'$ alors l'extension de p à $\mathcal{F}_p$ est un projecteur de $\mathcal{F}_p$ sur $\mathcal{F}'$.

Démonstration
On sait d'après la proposition §2.2.3 que l'extension de p est un projecteur dans $\mathcal{F}_p$. Supposons que les VAR positives de $\mathcal{F}'$ sont les VAR positives de $\mathcal{F}_p$ limites croissantes

de VAR de $\mathcal{E}'$ et montrons $p(\mathcal{F}_p) = \mathcal{F}'$. Il faut établir (voir §2.2.2) $p(\mathcal{F}_p) \subseteq \mathcal{F}'$ et $\forall Y \in \mathcal{F}', p(Y) = Y$.
D'abord $p(\mathcal{F}_p) \subseteq \mathcal{F}'$. Soit $Y \in p(\mathcal{F}_p)$ on a $Y \in \mathcal{F}_p$ et $Y = p(X)$ avec $X \in \mathcal{F}_p$ On suppose $X \geq 0$ il existe dans ce cas une suite $(X_n) \subseteq \mathcal{E}$ telle que $X_n \uparrow X$ donc $p(X_n) \uparrow p(X)$ et $p(X_n) \in \mathcal{E}'$ puisque p est un projecteur de $\mathcal{E}$ sur $\mathcal{E}'$. Posons $Y_n = p(X_n)$ on a $Y_n \uparrow Y$ et $Y_n \in \mathcal{E}'$ ce qui entraîne $Y \in \mathcal{F}'$ d'après l'hypothèse. Pour X de signe quelconque on décompose $X = X^+ - X^-$ on a $Y = p(X^+) - p(X^-)$ et d'après le raisonnement précédent $p(X^+) \in \mathcal{F}'$ et $p(X^-) \in \mathcal{F}'$ donc $Y \in \mathcal{F}'$.
Ensuite $\forall Y \in \mathcal{F}', p(Y) = Y$. Pour $Y \in \mathcal{F}', Y \geq 0$ il existe d'après l'hypothèse une suite $(Y_n) \subseteq \mathcal{E}'$ telle que $Y_n \uparrow Y$ donc $p(Y_n) \uparrow p(Y)$. Comme $\forall n, p(Y_n) = Y_n$ puisque p est un projecteur sur $\mathcal{E}'$ on obtient $p(Y) = Y$. Pour Y de signe quelconque on décompose $Y = Y^+ - Y^-$. $\square$

2.3 Espérance mathématique

2.3.1 Définition : espérance mathématique

Si X est une variable aléatoire (VA) définie sur un espace de probabilité $(\Omega, \mathcal{F}, P)$ on appelle *espérance* de X et on note $E(X)$ l'intégrale de Lebesgue (voir [3]) de X (lorsqu'elle existe) relativement à la probabilité P.

$$E(X) = \int_\Omega X(\omega)P(d\omega) = \int_\Omega XdP$$

On va construire ci-dessous l'intégrale de Lebesgue d'une VA relativement à une probabilité P (c'est-à-dire une mesure positive finie). On définit d'abord explicitement l'espérance E sur l'espace réticulé $\mathcal{E}$ des VAR étagées sur la tribu $\mathcal{F}$ et on montre qu'il s'agit d'un opérateur croissant à valeur dans l'espace des VAR constantes. On étend ensuite cet opérateur au cône $\mathcal{F}^+$ des VA positives puis à l'espace $\mathcal{F}_E$ des VAR intégrables en appliquant la méthode décrite au §2.1.

Définition initiale : espérance d'une VAR étagée sur la tribu $\mathcal{F}$
L'espace $\mathcal{E}$ des VAR étagées sur la tribu $\mathcal{F}$ est un espace réticulé contenant les constantes réelles. Pour X étagée sur $\mathcal{F}$ de la forme $X = \sum_{i=1,n} x_i 1_{A_i}$ on pose $E(X) = \sum_{i=1,n} x_i P(A_i)$. Cette définition est indépendante de la décomposition de X sur $\mathcal{F}$. L'application $X \to E(X)$ est un projecteur croissant de $\mathcal{E}$ sur $\mathbb{R}$ et elle vérifie la propriété de convergence décroissante vers 0 (lemme ci-dessous).

Lemme : convergence décroissante vers 0
Si $X_n \downarrow 0$ dans $\mathcal{E}$ alors $E(X_n) \downarrow 0$. C'est l'hypothèse (2) du §2.1.1.

Démonstration
Pour tout $\epsilon \geq 0$, $E(X_n) = E(X_n 1_{X_n \leq \epsilon}) + E(X_n 1_{X_n > \epsilon}) \leq \epsilon + \sup(X_0) P(X_n > \epsilon)$ et $P(X_n > \epsilon) \downarrow 0$ car $(X_n > \epsilon) \downarrow \emptyset$ puisque $X_n \downarrow 0$

Première extension : espérance d'une VA positive
Pour toute VA positive X il existe (§1.3.5) une suite croissante (X_n) de VAR étagées sur $\mathcal{F}$ telle que $\lim_n \uparrow X_n = X$. C'est l'hypothèse (1) du §2.1.1.
On pose $E(X) = \lim_n \uparrow E(X_n)$. Cette définition est indépendante de la suite (X_n) de

fonctions étagées telle que $\lim_n \uparrow X_n = X$. L'application $X \to E(X)$ est croissante, linéaire et elle vérifie le théorème de convergence croissante dans le cône $\mathcal{F}^+$ des VA positives (§2.1.3).

Deuxième extension : espérance d'une VA quasi intégrable
Pour toute VA X on décompose X en parties positive X^+ et négative X^- et on dit que X est *quasi intégrable* pour la probabilité P (opérable par E au §2.1.5) si $E(X^+)$ et $E(X^-)$ ne sont pas simultanément infinies, on pose alors $E(X) = E(X^+) - E(X^-)$ dans $\mathbb{R} \cup \{-\infty, \infty\}$. On dit que X est *intégrable* pour la probabilité P (opérable par E dans $\mathbb{R}$ au §2.1.5) lorsque $E(X^+)$ et $E(X^-)$ sont toutes les deux finies, on a alors $E(X) = E(X^+) - E(X^-)$ dans $\mathbb{R}$.

Remarque
Comme $\forall A \in \mathcal{F}, E(1_A) = P(A)$ les négligeables pour E coincident avec les parties négligeables mesurables de $(\Omega, \mathcal{F}, P)$ c'est-à-dire les éléments de $\mathcal{N}(\Omega, \mathcal{F}, P) \cap \mathcal{F}$.

Lemme : égalité PS et espérance
Si X et Y sont deux VA intégrables et $A \in \mathcal{F}$ on note $E(X, A) = E(X1_A) = \int_A XdP$ et on a :

1. $\forall A \in \mathcal{F}, E(X, A) \geq 0 \Longleftrightarrow X \geq 0 \ PS$
2. $\forall A \in \mathcal{F}, E(X, A) = 0 \Longleftrightarrow X = 0 \ PS$
3. $\forall A \in \mathcal{F}, E(X, A) = E(Y, A) \Longleftrightarrow X = Y \ PS$

Démonstration
Montrons $\forall A \in \mathcal{F}, E(X, A) \geq 0 \Longrightarrow X \geq 0 \ PS$. Soit $A_n = (X \leq -\frac{1}{n})$ on a $A_n \in \mathcal{F}$ donc $E(X, A_n) \geq 0$ or $X \leq -\frac{1}{n}$ sur A_n donc $P(A_n) = 0$ et $P(X < 0) = P(\cup_n A_n) = 0$ c'est-à-dire $X \geq 0$ PS. □

2.3.2 Définition et proposition : espace $\mathcal{F}_E$ et espace L^1

On note $\mathcal{F}_E$ l'espace vectoriel des VAR intégrables et $L^1 = L^1(\Omega, \mathcal{F}, P)$ l'espace vectoriel des classes pour l'égalité PS des VAR intégrables. Notons que deux VAR d'une même classe ont même intégrale. L'espace quotient L^1 est un espace vectoriel et $||X||_1 = E(|X|)$ définit une norme sur l'espace L^1 qui muni de cette norme est un espace de Banach.

Démonstration
Montrons que l'application $X \to ||X||_1 = E(|X|)$ est une norme sur L^1.

1. $||X||_1 = 0 \Longleftrightarrow X = 0 \ PS$ d'après le lemme ci-dessus.
2. $||\lambda X||_1 = E(|\lambda X|) = E(|\lambda||X|) = |\lambda|E(|X|) = |\lambda|||X||_1$
3. $||X + Y||_1 = E(|X + Y|) \leq E(|X| + |Y|) = E(|X|) + E(|Y|) = ||X||_1 + ||Y||_1$

La complétude de L^1 sera établie au §2.4.4 □

Remarque
Si $X_n \to X$ dans L^1 comme $|E(X_n) - E(X)| = |E(X_n - X)| \leq E(|X_n - X|)$ on a $E(X_n) \to E(X)$

Le théorème ci-dessous rappelle les propriétés générales de l'extension d'un opérateur p de variables aléatoires (§2.1) dans le cas de l'espérance E qui est un projecteur de l'espace $\mathcal{E}$ des VAR étagées sur la tribu $\mathcal{F}$ sur l'espace des VAR constantes.

2.3.3 Théorème : propriétés de l'espérance des VA

1. L'espérance $X \to E(X)$ est un projecteur croissant de l'espace $\mathcal{F}_E$ des VAR intégrables sur l'espace des VAR constantes. Elle vérifie :
2. Le théorème de convergence monotone (TCM) :
   — $E(X_0^-) < \infty, X_n \uparrow X \Longrightarrow E(X_n) \uparrow E(X)$
   — $E(X_0^+) < \infty, X_n \downarrow X \Longrightarrow E(X_n) \downarrow E(X)$
3. Les inégalités de Fatou :
   — $\forall n, Y \leq X_n, E(Y^-) < \infty \Longrightarrow E(\liminf_n X_n) \leq \liminf_n E(X_n)$
   — $\forall n, X_n \leq Z, E(Z^+) < \infty \Longrightarrow E(\limsup_n X_n) \geq \limsup_n E(X_n)$
4. Le théorème de convergence dominée (TCD) :
$$\lim_n X_n = X, \forall n, |X_n| \leq Z, E(Z) < \infty \Longrightarrow E(X) = \lim_n E(X_n)$$
5. L'inégalité de Jensen :
$$\phi : \mathbb{R} \to \mathbb{R} \text{ convexe}, \phi(X) \in \mathcal{F}_E \Longrightarrow \phi[E(X)] \leq E[\phi(X)]$$

Démonstration
Le procédé d'extension d'opérateur du §2.1 utilisé ici pour étendre $X \to E(X)$ de $\mathcal{E}$ à $\mathcal{F}_E$ permet d'obtenir un opérateur croissant. Comme l'application $E(.)$ est un projecteur de $\mathcal{E}$ sur $\mathcal{E}' = \mathbb{R}$ et que les limites croissantes finies de VAR constantes positives sont aussi des VAR constantes positives l'extension à $\mathcal{F}_E$ de l'espérance est un projecteur de $\mathcal{F}_E$ sur $\mathbb{R}$ (voir §2.2.4). L'extension de $E(.)$ de $\mathcal{E}$ à la classe des VA quasi intégrables satisfait le TCM (2) les inégalités de Fatou (3) le TCD (4) et l'inégalité de Jensen (5) d'après le §2.1.

Remarque I
Une conséquence immédiate du théorème de convergence monotone est que la convergence dans L^1 de (X_n) vers X entraîne la convergence PS vers X d'une sous suite (X_{n_k}). Montrons le pour $X = 0$ prenons (n_k) telle que $\sum_k E(|X_{n_k}|) < \infty$ on a $E(\sum_k |X_{n_k}|) < \infty$ donc $\sum_k |X_{n_k}| < \infty$ et il en résulte $X_{n_k} \to 0$ (terme général d'une série convergente).

Remarque II
Une conséquence du théorème de convergence dominée est que si $X_n \to X$ PS et si $\forall n, |X_n| \leq Y \in L^1$ alors $X_n \to X$ dans L^1 car $|X_n - X| \leq Y + |X| \in L^1$ donc $|X_n - X| \to 0$ entraîne $E(|X_n - X|) \to 0$

Exemples de convergence
On dispose dans L^1 de deux modes de convergence : la convergence simple sauf éventuellement sur un ensemble négligeable et la convergence dans L^1 au sens de la norme. On donne ci dessous dans l'espace $([0,1], \mathcal{B}([0,1]), \lambda)$ où λ est la mesure de Lebesgue deux exemples de suites qui convergent PS ou dans L^1 mais pas les deux.

1) D'abord une suite qui converge PS mais qui ne converge pas dans L^1.
Soit $X_n = n1_{[0,\frac{1}{n}]}$ pour $n \geq 1$ on a $\lim_n X_n = 0$ PS (sauf en 0) et $\forall n, E(X_n) = 1$. Si (X_n) converge dans L^1 vers une VAR X il existe une sous suite qui converge PS vers X donc $X = 0$ et $E(X_n) \to E(X) = 0$ ce qui contredit $\forall n, E(X_n) = 1$.

2) Ensuite une suite qui converge dans L^1 mais qui ne converge pas PS.
Soit $X_1 = 1_{[0,\frac{1}{2}]}$, $X_2 = 1_{[\frac{1}{2},1]}$, $X_3 = 1_{[0,\frac{1}{3}]}$, $X_4 = 1_{[\frac{1}{3},\frac{2}{3}]}$, $X_5 = 1_{[\frac{2}{3},1]}$, $X_6 = 1_{[0,\frac{1}{4}]}, \ldots$

Il s'agit d'une sorte de *bosse circulante*. On a $\lim_n X_n = 0$ dans L^1 car $E(X_n) \downarrow 0$ mais pour tout $u \in [0,1]$, la suite $[X_n(u)]$ n'a pas de limite.

2.3.4 Définition et proposition : convergence en probabilité

Soit (X_n) et X des VAR défines sur $(\Omega, \mathcal{F}, P)$ on définit la *convergence en probabilité* de X_n vers X par :

$$\forall \epsilon > 0, \ P(|X_n - X| > \epsilon) \to 0$$

L'application $(X,Y) \to d(X,Y) = E(|X-Y| \wedge 1)$ est une distance sur l'espace $L^0(\Omega, \mathcal{F}, P)$ des classes de VAR définies sur $(\Omega, \mathcal{F}, P)$ et $X_n \to X$ en probabilité si et seulement si $d(X_n, X) \to 0$. Notons que la convergence dans L^1 et la convergence PS implique chacune la convergence en probabilité et que si $X_n \to X$ en probabilité il existe une sous suite (X_{n_k}) telle que $X_{n_k} \to X \ PS$.

Démonstration

Montrons que $d(X,Y))$ est une distance sur l'espace $L^0(\Omega, \mathcal{F}, P)$. On a (trivialement) d(X, Y) = d(Y, X). Supposons $d(X,Y) = 0$ et montrons $X = Y$ PS, il vient en notant $E(X, A) = E(X 1_A) = \int_A X dP$:

$$E(|X-Y| \wedge 1) = E(|X-Y|, |X-Y| \leq 1) + P(|X-Y| > 1) = 0$$
$$\Longrightarrow E(|X-Y|, |X-Y| \leq 1) = 0 \text{ et } P(|X-Y| > 1) = 0$$
$$\Longrightarrow E(|X-Y|) = E(|X-Y|, |X-Y| \leq 1) + E(|X-Y|, |X-Y| > 1) = 0$$
$$\Longrightarrow X = Y \text{ PS}$$

L'inégalité triangulaire découle de :

$$E(|X-Z| \wedge 1) \leq E[(|X-Y| + |Y-Z|) \wedge 1] \leq E(|X-Y| \wedge 1 + |Y-Z| \wedge 1)$$
$$= E(|X-Y| \wedge 1) + E(|Y-Z| \wedge 1)$$

Montrons que $X_n \to 0$ en probabilité équivaut à $d(X_n, 0) = E(|X_n| \wedge 1) \to 0$. Pour tout $0 < \epsilon < 1$ on a :

$$E(|X_n| \wedge 1) = E(|X_n|, \ |X_n| \leq \epsilon) + E(|X_n|, \ \epsilon < |X_n| \leq 1) + E(1, |X_n| > 1)$$
$$\leq \epsilon + P(\epsilon < |X_n|) + P(|X_n| > 1)$$

Donc $P(|X_n| > \eta) \to 0$ entraîne $E(|X_n| \wedge 1) \to 0$.

Inversement $E(|X_n| \wedge 1) \geq \epsilon P(\epsilon < |X_n| \leq 1) + P(|X_n| > 1)$ donc $E(|X_n| \wedge 1) \to 0$ entraîne $P(\epsilon < |X_n| \leq 1) \to 0$ et $P(|X_n| > 1) \to 0$ d'où $P(|X_n| > \epsilon) \to 0$.

La convergence dans L^1 implique la convergence en probabilité car $X_n \to 0$ dans L^1 s'écrit $E(|X_n|) \to 0$ ce qui entraîne $E(|X_n| \wedge 1) \to 0$ et la convergence PS implique aussi convergence en probabilité car $X_n \to 0$ PS entraîne $E(|X_n| \wedge 1) \to 0$ (TCD).

Enfin si $X_n \to X$ en probabilité il existe une sous suite (X_{n_k}) telle que $X_{n_k} \to X$ PS. Pour $X = 0$ il suffit de choisir (n_k) telle que $\sum_k E(|X_{n_k}| \wedge 1) < \infty$ ce qui entraîne $\sum_k |X_{n_k}| \wedge 1) < \infty$ PS et il en résulte $X_{n_k} \to 0$ PS (remarque I de §2.3.3). □

Remarque : convergence uniforme en probabilité

Comme pour tout $t \leq \theta$ on a $d(X_t^{(n)}, X_t) = d(X_t^{(n)} - X_t, 0) \leq d(Y_\theta^{(n)}, 0)$. La convergence en probabilité $X_t^{(n)} \to X_t$ uniforme en $t \in [0, \theta]$ implique la convergence de la distance $d(X_t^{(n)}, X_t)$ vers 0 uniforme en $t \in [0, \theta]$ mais la réciproque est fausse.

Exemple

Soit $\Omega = [0,1]$ et $\theta = 1$ si $X_t(\omega) = 1_{[t]}(\omega)$ on a $E_\lambda(X_t) = 0$ et $\sup_{t \leq 1} X_t = 1$ donc $E_\lambda(\sup_{t \leq 1} X_t) = 1$

2.3.5 Proposition : un résultat de convergence en probabilité

Soit (X_n) une suite de VAR et X une VAR. Si $X_n \to X$ en probabilité et si $\forall \epsilon > 0, \exists K \in \mathbb{R}^+, \forall n,\ P(|X_n| \geq K) \leq \epsilon$ alors $P(|X| \geq K) \leq \epsilon$.

Démonstration

Pour tout $\eta > 0$ et $K \in \mathbb{R}^+$ tel que $\forall n, P(|X_n| \geq K) \leq \epsilon$ il vient :

1. $|X - X_n| < \eta, |X_n| < K \Longrightarrow |X| \leq |X - X_n| + |X_n| < K + \eta$
2. $(|X| \geq K + \eta) \subseteq (|X - X_n| \geq \eta)\ \cup\ (|X_n| \geq K)$
3. $P(|X| \geq K + \eta) \leq P(|X - X_n| \geq \eta) + P(|X_n| \geq K)$
4. $P(|X| \geq K + \eta) \leq \epsilon$ car $P(|X - X_n| \geq \eta) \to 0$

Il en résulte $P(|X| \geq K) = P[\cup_n (|X| \geq K + 2^{-n}] = \sup_n P(|X| \geq K + 2^{-n}) \leq \epsilon$ $\square$

2.3.6 Définition : indépendance

1) Une famille $(A_i)_{i=1,n}$ d'événements (mesurables) est *indépendante* si

$$P\left(\bigcap_{i=1,n} A_i\right) = \prod_{i=1,n} P(A_i)$$

2) Une famille $(\mathcal{G}_i)_{i=1,n}$ de sous tribus est indépendante si $\forall A_i \in \mathcal{G}_i$ la famille $(A_i)_{i=1,n}$ est indépendante.
3) Une famille $(X_i)_{i=1,n}$ de VA est indépendante si la famille de sous tribus $[\sigma(X_i)]_{i=1,n}$ est indépendante.
Ces définitions s'étendent aux familles de cardinalité quelconque en disant qu'une famille est indépendante si toute sous famille finie est indépendante.

Remarque

Si X et Y sont indépendantes, alors X^+ et Y^+, X^+ et Y^-, X^- et Y^+, X^- et Y^- le sont aussi. Plus généralement si f et g sont mesurables de $\mathbb{R}$ dans $\mathbb{R}$, et si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ le sont aussi.

2.3.7 Proposition : produit de VA indépendantes

Si X et Y sont des VA indépendantes, intégrables et telles que XY soit intégrable, alors $E(XY) = E(X)E(Y)$.

Démonstration

On montre le résultat pour X et Y étagées, on approche ensuite X et Y positives par des suites croissantes de fonctions étagées et on utilise le TCM. On décompose enfin X et Y en parties positives et négatives.

Un contre exemple

Des événements A_i où $i = 1, n$ peuvent être indépendants deux à deux sans que la famille $(A_i)_{i=1,n}$ soit indépendante.
Par exemple, si X et Y sont des VAR indépendantes de *Bernouilli* de paramètre $\frac{1}{2}$ et si on pose $A = (X = 0), B = (Y = 0)$ et $C = (X = 0, Y = 0) \cup (X = 1, Y = 1)$ on a : $P(A) = P(B) = P(C) = \frac{1}{2}$ et $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$, les événements A, B, C sont indépendants deux à deux mais $P(A \cap B \cap C) = \frac{1}{4}$ et $P(A)P(B)P(C) = \frac{1}{8}$

2.4 Famille uniformément intégrable

On définit ci-dessous les familles de VAR uniformément intégrables et on va montrer que dans une famille de VAR uniformément intégrables la convergence en probabilité est équivalente à la convergence dans L^1, c'est-à-dire au sens de la norme L^1.

Pour toute VAR X définie et intégrable sur un espace de probabilité $(\Omega, \mathcal{F}, P)$ et tout événement $A \in \mathcal{F}$ on pose $E(X, A) = E(X 1_A) = \int_A X dP$

2.4.1 Définition : famille de VAR uniformément intégrables

Une famille $\mathcal{H}$ de VAR intégrables est *uniformément intégrable* (UI) si pour $M \in \mathbb{R}^+$:

$$\lim_{M \to \infty} \sup_{X \in \mathcal{H}} E(|X|, |X| > M) = 0$$

Ce qui revient à dire $E(|X|, |X| > M) \to 0$ lorsque $M \to \infty$ uniformément en $X \in \mathcal{H}$.

Exemple
Une famille $\mathcal{H}$ de VAR intégrables dominée par une variable intégrable Y est UI. En effet si $\forall X \in \mathcal{H},\ |X| \leq Y$ alors $E(|X|, |X| > M) \leq E(Y, Y > M)$ et comme $E(Y, Y > M) \to 0$ lorsque $M \to \infty$ car $E(Y, Y \leq M) \uparrow E(Y)$ lorsque $M \to \infty$ donc $E(Y, Y > M) \downarrow 0$ on a $E(|X|, |X| > M) \to 0$. En particulier une famille réduite à une ou à un nombre fini de variables intégrables est UI.

Un contre exemple : une suite non UI
Les variables $X_n = n 1_{[0, \frac{1}{n}]}$ sont définies et intégrables sur $\mathbb{R}$ mais ne constituent pas une famille UI, en effet $\forall M, n > M \Longrightarrow E(|X_n|, |X_n| > M) = 1$.

2.4.2 Définition et proposition : famille de VAR équi-intégrables

Une famille $\mathcal{H}$ de VAR intégrables est *équi-intégrable* (EI) si pour $A \in \mathcal{F}$:

$$\lim_{P(A) \to 0} \sup_{X \in \mathcal{H}} E(|X|, A) = 0$$

Une famille $\mathcal{H}$ de VAR est UI si et seulement si elle est EI et bornée dans L^1

Démonstration
Une famille $\mathcal{H}$ de VAR UI est toujours bornée dans L^1 car pour $\epsilon > 0$ il existe $M \in \mathbb{R}^+$ tel que $\forall X \in \mathcal{H}, E(|X|) = E(|X|, |X| \leq M) + E(|X|, |X| > M) \leq M + \epsilon$ et est EI car $E(|X|, A) = E(|X|, A \cap |X| \leq M) + E(|X|, A \cap |X| > M) \leq MP(A) + \epsilon$
Inversement, si $\mathcal{H}$ est EI et bornée dans L^1 alors

1. $\forall X \in \mathcal{H}, P(A) \leq \eta \Longrightarrow E(|X|, A) \leq \epsilon$
2. $\forall M \in \mathbb{R}^+,\ MP(|X| > M) \leq E(|X|, |X| > M) \leq E(|X|) \leq K$

Donc $M \geq \frac{K}{\eta} \Longrightarrow \frac{K}{M} \leq \eta \Longrightarrow P(|X| > M) \leq \eta \Longrightarrow E(|X|, |X| > M) \leq \epsilon.$ $\square$

2.4.3 Théorème : convergence en probabilité et convergence dans L^1

Soit (X_n) une suite de VAR intégrables et X une VAR, les assertions suivantes sont équivalentes :

1. (X_n) est UI et $X_n \to X$ en probabilité
2. X est intégrable et $X_n \to X$ dans L^1

Démonstration
Si $X_n \to X$ en probabilité il existe une suite extraite (X_{n_k}) telle que $X_{n_k} \to X$ PS donc $|X_{n_k}| \to |X|$ PS et si (X_n) est UI donc bornée dans L^1 :

$$E(|X|) = E(\liminf_k |X_{n_k}|) \leq \liminf_k E(|X_{n_k}|) \leq \sup_k E(|X_{n_k}|) < \infty$$

X est donc intégrable. Par ailleurs

$$\begin{aligned} ||X_n - X||_1 &= E(|X_n - X|, |X_n - X| \leq \epsilon) + E(|X_n - X|, |X_n - X| > \epsilon) \\ &\leq \epsilon + E(|X_n|, |X_n - X| > \epsilon) + E(|X|, |X_n - X| > \epsilon) \end{aligned}$$

Si $X_n \to X$ en probabilité $P(|X_n - X| > \epsilon) \to 0$ il en résulte $E(|X_n|, |X_n - X| > \epsilon) \to 0$ car (X_n) est UI donc EI et $E(|X|, |X_n - X| > \epsilon) \to 0$ car X est intégrable, finalement $||X_n - X||_1 \to 0$. On a donc établit (1) implique (2).

Montrons (2) implique (1). Si $X_n \to X$ dans L^1 alors $X_n \to X$ en probabilité. La suite (X_n) est bornée dans L^1 car $E(|X_n|) \leq E(|X|) + E(|X_n - X|)$ et $E(|X_n - X|) \to 0$ elle est aussi EI car $E(|X_n|, A) \leq E(|X|, A) + E(|X_n - X|)$ elle est donc UI. $\square$

2.4.4 Proposition : complétude de L^1

Si (X_n) est une suite de VAR intégrables de Cauchy pour la norme L^1 alors (X_n) est UI et il existe une VAR intégrable X telle que $X_n \to X$ dans L^1.

Démonstration
La suite (X_n) de Cauchy pour la norme L^1 est bornée dans L^1. Elle est EI car $\forall n \geq N, E(|X_n|, A) \leq E(|X_N|, A) + E(|X_N - X_n|, A) \leq E(|X_N|, A) + \epsilon$ elle est donc UI. Comme il existe une sous suite (X_{n_k}) telle que (X_{n_k}) soit PS de Cauchy dans $\mathbb{R}$ (voir §2.3.4) il existe une VAR X tels que $X_{n_k} \to X$ PS donc $X_{n_k} \to X$ en probabilité (§2.3.4) par conséquent puisque (X_n) est UI la VAR X est intégrable et $X_{n_k} \to X$ dans L^1 (§2.4.3). Il en résulte $X_n \to X$ dans L^1 car $||X - X_n||_1 \leq ||X - X_{n_k}||_1 + ||X_{n_k} - X_n||_1$ et (X_n) est de Cauchy dans L^1. $\square$

2.5 Espérance conditionnelle dans $L^2(\mathcal{F})$

Rappels : projecteurs, espace de Hilbert
Un projecteur dans un espace vectoriel (voir §2.2.1) est une application linéaire f d'un espace vectoriel $\mathcal{E}$ dans lui-même (c'est-à-dire un endomorphisme) telle que $f \circ f = f$ (propriété d'idempotance). Si $\mathcal{E}'$ est un sous espace de $\mathcal{E}$, alors f est un projecteur de $\mathcal{E}$ sur $\mathcal{E}'$ si $f(\mathcal{E}) = \mathcal{E}'$.

Si $\mathcal{H}$ est un espace de Hilbert et si $\mathcal{G}$ est un sous espace fermé de $\mathcal{H}$ alors pour tout $x \in \mathcal{H}$ il existe un unique $y \in \mathcal{G}$ tel que $x - y$ soit othogonal à $\mathcal{G}$ c'est la *projection orthogonale* de x sur $\mathcal{G}$.

On note $L^2(\Omega, \mathcal{F}, P)$ ou plus simplement $L^2(\mathcal{F})$ l'espace vectoriel des classes (pour l'égalité PS) de VAR de carré intégrable définies sur $(\Omega, \mathcal{F}, P)$. Comme on confond souvent une VAR et sa classe d'équivalence on confondra aussi souvent $L^2(\mathcal{F})$ et l'espace vectoriel des VAR de carré intégrable. En définissant dans $L^2(\mathcal{F})$ le produit scalaire $< X, Y >= E(XY)$ on muni $L^2(\mathcal{F})$ d'une structure d'espace de Hilbert. On note $X \perp Y$ pour $< X, Y >= 0$. Si $\mathcal{G}$ est une sous tribu de $\mathcal{F}$, $L^2(\mathcal{G})$ est un sous espace vectoriel fermé de $L^2(\mathcal{F})$. Pour tout élément X de $L^2(\mathcal{F})$ il existe donc un unique élément Y dans $L^2(\mathcal{G})$ telle que $X - Y \perp L^2(\mathcal{G})$ ce qui signifie $\forall Z \in L^2(\mathcal{G}), X - Y \perp Z$. Notons que Y est le vecteur de $L^2(\mathcal{G})$ dont la distance à X (au sens de la norme L^2) est minimale. L'application $X \to Y$ ainsi définie est le *projecteur orthogonal* de $L^2(\mathcal{F})$ sur $L^2(\mathcal{G})$.

2.5.1 Définition : espérance conditionnelle dans $L^2(\mathcal{F})$

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et $\mathcal{G}$ une sous-tribu de $\mathcal{F}$, on appelle *espérance conditionnelle* relativement à $\mathcal{G}$ et on note $E(.|\ \mathcal{G})$ le projecteur orthogonal de $L^2(\mathcal{F})$ sur $L^2(\mathcal{G})$. L'espérance conditionnelle définie sur $L^2(\mathcal{F})$ est donc linéaire et l'espace image $L^2(\mathcal{G})$ est invariant en particulier les classes de constantes sont invariantes.
Si X est une VAR de carré intégrable on appelle espérance conditionnelle de X relativement à $\mathcal{G}$ et on note $E(X|\ \mathcal{G})$ un élément quelconque de la classe projetée de la classe de X. Si Y est une VAR de la classe projetée de X on écrira donc $Y = E(X|\ \mathcal{G})$ PS autrement dit l'espérance conditionnelle d'une VAR X de carré intégrable est définie à un négligeable près.

2.5.2 Proposition : caractérisation de l'espérance conditionnelle

Soit $\mathcal{G}$ une sous tribu de $\mathcal{F}$. Pour toute VAR X de carré intégrable l'espérance conditionnelle $Y = E(X|\ \mathcal{G})$ de X est caractérisée à un négligeable près par la conjonction de $Y \in L^2(\mathcal{G})$ et de $\forall Z \in L^2(\mathcal{G}), X - Y \perp Z$. Cette dernière assertion étant équivalente à chacune des trois assertions :

1. $\forall Z \in L^2(\mathcal{G}), E(XZ) = E(YZ)$
2. $\forall Z \in b\mathcal{G}, E(XZ) = E(YZ)$
3. $\forall A \in \mathcal{G}, E(X, A) = E(Y, A)$

Remarques
Il résulte de §2.5.2 *l'identité remarquable des espérances conditionnelles* :

$$\forall X \in L^2(\mathcal{F}), \forall Z \in L^2(\mathcal{G}),\ E(XZ) = E[E(X|\ \mathcal{G})Z]$$

En particulier $E(X) = E[E(X|\ \mathcal{G})]$. Notons que si X est positif on a pour tout $A \in \mathcal{G}$ $E[E(X|\ \mathcal{G}), A] = E(X, A) \geq 0$ donc $E(X|\ \mathcal{G}) \geq 0$. L'espérances conditionnelle est positive comme elle est linéaire elle est croissante.

2.5.3 Définition : espérance conditionnée par une VAR

Si $X \in L^2(\mathcal{F})$ et si Y est une VAR on pose $E(X|Y) = E[X|\, \sigma(Y)]$

Remarque : factorisation de l'espérance conditionnelle

Tout représentant de la classe $E(X \mid Y)$ est $\sigma(Y)$ mesurable. Il existe donc (proposition §1.3.7) une application f mesurable de $\mathbb{R}$ dans $\mathbb{R}$ telle que $E(X \mid Y) = f(Y)\ PS$ L'espérance de X conditionnée par Y et factorisée par Y notée $E(X \mid Y = y) = f(y)$ est PS la valeur prise par $E(X \mid Y)(\omega)$ lorsque $Y(\omega) = y$.

Exemple : conditionnement par une tribu atomique

Soit $X \in L^2(\Omega, \mathcal{F}, P)$ et $\mathcal{G}$ une sous tribu atomique (voir §1.1.2) de $\mathcal{F}$ engendrée par une partition mesurable finie (A_i) de Ω. L'application $\omega \to E(X|\, \mathcal{G})(\omega)$ est mesurable par rapport à $\mathcal{G}$ elle est donc constante sur chaque A_i, on a donc :

$$\forall \omega \in A_i, \int_{A_i} X(\omega) P(d\omega) = \int_{A_i} E(X \mid \mathcal{G})(\omega) P(d\omega) = E(X \mid \mathcal{G})(\omega) P(A_i)$$

l'expression de $E(X|\, \mathcal{G})$ en découle :

$$E(X \mid \mathcal{G}) = \sum_{P(A_i) \neq 0} \frac{E(X, A_i)}{P(A_i)} 1_{A_i}$$

Si $A = \cup_{P(A_i) \neq 0} A_i$ on a $P(A^c) = 0$. L'espérance condidionnelle $E(X \mid \mathcal{G})$ est arbitraire sur l'événement A^c de probabilité nulle, ici $E(X \mid \mathcal{G})(\omega) = 0$

En particulier pour $A, B \in \mathcal{F}$ où $0 < P(A) < 1$ on obtient :

$$\begin{aligned} E(1_B|1_A) = E[1_B|\sigma(A, A^c)] &= \tfrac{P(A \cap B)}{P(A)} 1_A + \tfrac{P(A^c \cap B)}{P(A^c)} 1_{A^c} \\ &= P(B|A) 1_A + P(B|A^c) 1_{A^c} \end{aligned}$$

2.5.4 Théorème : propriétés de l'espérance conditionnelle dans $L^2(\mathcal{F})$

1. L'espérance conditionnelle $X \to E(X \mid \mathcal{G})$ est un projecteur croissant de $L^2(\mathcal{F})$ sur $L^2(\mathcal{G})$, elle vérifie :
2. le théorème de convergence monotone dans $L^2(\mathcal{F})$: si (X_n) est une suite monotone dans $L^2(\mathcal{F})$ convergente vers $X \in L^2(\mathcal{F})$ alors $\lim_n X_n = X$
3. le théorème de projections successives : si $\mathcal{H}$ est une sous tribu de $\mathcal{G}$ alors pour tout $X \in L^2(\mathcal{F})$, $E(X \mid \mathcal{H}) = E[E(X \mid \mathcal{G}) \mid \mathcal{H}]$ notée $E(X \mid \mathcal{G} \mid \mathcal{H})$
4. le théorème du produit : $\forall X \in L^2(\mathcal{F}), \forall Y \in b\mathcal{G},\ E(XY \mid \mathcal{G}) = Y E(X \mid \mathcal{G})$
5. le théorème d'indépendance : si $X \in L^2(\mathcal{F})$ est indépendante de $\mathcal{G}$ alors $E(X \mid \mathcal{G}) = E(X)$.

Démonstration

(1) L'application $X \to E(X|\, \mathcal{G})$ est par définition un projecteur de $L^2(\mathcal{F})$ sur $L^2(\mathcal{G})$ l'espace image $L^2(\mathcal{G})$ est invariant et en particulier les constantes $C \in \mathbb{R}$ sont invariantes. Elle est positive (voir la remarque du §2.5.2) donc croissante puisqu'elle est linéaire.

(2) Théorème de convergence monotone : si (X_n) est une suite de $L^2(\mathcal{F})$ décroissante vers 0 on a $E[\lim_n \downarrow E(X_n|\ \mathcal{G})] = \lim_n \downarrow E[E(X_n|\ \mathcal{G})] = \lim_n \downarrow E(X_n) = 0$ et $\lim_n \downarrow E(X_n|\ \mathcal{G}) \geq 0$ donc $\lim_n \downarrow E(X_n|\ \mathcal{G}) = 0$. Il en résulte que si (X_n) est une suite de $L^2(\mathcal{F})$ croissante ou décroissante vers X appartenant à $L^2(\mathcal{F})$ on a $\lim_n E(X_n|\ \mathcal{G}) = E(X|\ \mathcal{G})$

(3) Théorème des projections successives : la caractérisation des projections s'écrit : $\forall Z \in b\mathcal{G}, E(XZ) = E[E(X|\ \mathcal{G})Z]$ et $\forall Z \in b\mathcal{H}, E[E(X|\ \mathcal{G})Z] = E[E(X|\ \mathcal{G} \mid \mathcal{H})Z]$. Par hypothèse $b\mathcal{H} \subseteq b\mathcal{G}$ donc $\forall Z \in b\mathcal{H}, E(XZ) = E[E(X|\ \mathcal{G} \mid \mathcal{H})Z]$ ce qui signifie $E(X|\ \mathcal{H}) = E(X|\ \mathcal{G} \mid \mathcal{H})$.

(4) Théorème du produit : par hypothèse $Y \in b\mathcal{G}$ donc $\forall Z \in b\mathcal{G}, YZ \in b\mathcal{G}$ et $E(XYZ) = E[E(X|\ \mathcal{G})YZ]$ ce qui signifie $E(XY|\ \mathcal{G}) = E(X|\ \mathcal{G})Y$

(5) Théorème d'indépendance : Si $X \in L^2(\mathcal{F})$ est indépendante de $\mathcal{G}$ alors pour tout $Z \in b\mathcal{G}$, X est indépendante de Z donc $E(XZ) = E(X)E(Z)$ et comme $E(X)$ est une constante on a $E[E(X)Z] = E(X)E(Z)$ ce qui signifie $E(X|\ \mathcal{G}) = E(X)$. $\square$

2.6 Espérance conditionnelle des variables aléatoires

L'espérance conditionnelle $E(.|\ \mathcal{G})$ est définie au §2.5 sur l'espace réticulé $L^2(\mathcal{F})$ des VAR de carré intégrable. On l'étend selon §2.1 au cône $\mathcal{F}^+$ des VA positives puis à la classe des VA ayant une espérance conditionnelle. On note :

$$\mathcal{F}_{\mathcal{G}} = \{X \in \mathcal{F} \mid E(X^+|\ \mathcal{F}) < \infty \text{ et } E(X^-|\ \mathcal{G}) < \infty\}$$

C'est l'espace vectoriel des VAR qui ont une espérance conditionnelle $E(.|\ \mathcal{G})$ à valeurs réelles.

2.6.1 Théorème : propriétés de l'espérance conditionnelle des VA

1. L'espérance conditionnelle $X \to E(.|\ \mathcal{G})$ est un projecteur croissant de $\mathcal{F}_{\mathcal{G}}$ sur l'espace $\mathcal{G}$ des VAR G mesurables. Elle vérifie :
2. Le théorème de convergence monotone (TCM) :
   — $E(X_0^-|\ \mathcal{G}) < \infty, X_n \uparrow X \Longrightarrow E(X_n|\ \mathcal{G}) \uparrow E(X|\ \mathcal{G})$
   — $E(X_0^+|\ \mathcal{G}) < \infty, X_n \downarrow X \Longrightarrow E(X_n|\ \mathcal{G}) \downarrow E(X|\ \mathcal{G})$
3. Les inégalités de Fatou :
   — $\forall n, Y \leq X_n, E(Y^-|\ \mathcal{G}) < \infty \Longrightarrow E(\liminf_n X_n|\ \mathcal{G}) \leq \liminf_n E(X_n|\ \mathcal{G})$
   — $\forall n, X_n \leq Z, E(Z^+|\ \mathcal{G}) < \infty \Longrightarrow E(\limsup_n X_n|\ \mathcal{G}) \geq \limsup_n E(X_n|\ \mathcal{G})$
4. Le théorème de convergence dominée (TCD) :
$$\lim X_n = X, \forall n, |X_n| \leq Z, E(Z|\ \mathcal{G}) < \infty \Longrightarrow \lim_n E(X_n|\ \mathcal{G}) = E(X|\ \mathcal{G})$$
5. L'inégalité de Jensen :
$$\phi : \mathbb{R} \to \mathbb{R} \text{ convexe, } \phi(X) \in \mathcal{F}_{\mathcal{G}} \Longrightarrow \phi[E(X|\ \mathcal{G})] \leq E[\phi(X)|\ \mathcal{G})$$
6. Le théorème de projections successives :
   Si $\mathcal{H}$ est une sous tribu de $\mathcal{G}$ et $X \in \mathcal{F}_{\mathcal{G}} \cap \mathcal{F}_{\mathcal{H}}$ alors $E(X|\ \mathcal{H}) = E(X|\ \mathcal{G}|\ \mathcal{H})$

7. Le théorème du produit :
   On a $E(XY|\mathcal{G}) = YE(X|\mathcal{G})$ lorsque $X \in \mathcal{F}^+$ et $Y \in \mathcal{G}^+$ où lorsque $X \in \mathcal{F}_\mathcal{G}$, $Y \in \mathcal{G}$ et $XY \in \mathcal{F}_\mathcal{G}$.
8. Le théorème d'indépendance : toute VA X indépendante de $\mathcal{G}$ et quasi intégrable a une espérance conditionnelle égale à $E(X)$.

Remarque
Une conséquence de l'inégalité de Jensen est que l'application : $X \to E(X|\mathcal{G})$ est une contraction de $L^2(\mathcal{F})$ sur $L^2(\mathcal{G})$ car en prenant $\phi(x) = x^2$ on obtient :

$$||E(X|\mathcal{G})||_2 \leq ||X||_2$$

Démonstration
Le procédé d'extension d'opérateur du §2.1 utilisé ici pour étendre $X \to E(X|\mathcal{G})$ de $L^2(\mathcal{F})$ à $\mathcal{F}_\mathcal{G}$ permet d'obtenir un opérateur croissant. Comme l'application $E(.|\mathcal{G})$ est un projecteur de $L^2(\mathcal{F})$ sur $L^2(\mathcal{G})$ et que les VAR $\mathcal{G}$ mesurables positives sont les limites croissantes finies d'éléments de $L^2(\mathcal{G})$ l'extension à $\mathcal{F}_\mathcal{G}$ de l'espérance conditionnelle est un projecteur de $\mathcal{F}_\mathcal{G}$ sur $\mathcal{G}$ (§2.2.4). L'extension de $E(.|\mathcal{G})$ de $L^2(\mathcal{F})$ à la classe des VA ayant une espérance conditionnelle satisfait le TCM (2) les inégalités de Fatou (3) le TCD (4) et l'inégalité de Jensen (5) (voir §2.1). Les résultats (6), (7), (8) sont spécifiques de l'espérance conditionnelle.

Vérifions le théorème des projections successives (6). Si $X \in \mathcal{F}_\mathcal{G} \cap \mathcal{F}_\mathcal{H}$ est une VA positive on a $E(X \wedge n|\mathcal{H}) = E(X \wedge n|\mathcal{G}|\mathcal{H})$ d'où le résultat par limite croissante. Pour X de signe quelconque on décompose $X = X^+ - X^-$.

Vérifions le théorème du produit (7). Si $X \in \mathcal{F}^+$ et $X \in \mathcal{G}^+$ sont des VA positives alors $E[(X\wedge n)(Y\wedge n)|\mathcal{G}] = (Y\wedge n)E(X\wedge n|\mathcal{G})$ d'où le résultat dans le premier cas par limite croissante. Pour $X \in \mathcal{F}_\mathcal{G}$ et $Y \in \mathcal{G}$ telles que $XY \in \mathcal{F}_\mathcal{G}$ on pose $X_n = X1_{|X|\leq n}$ et $Y_n = Y1_{|Y|\leq n}$ on a $E[(X_nY_n)|\mathcal{G}) = Y_nE(X_n|\mathcal{G})$ d'où le résultat dans le deuxième cas par le TCD.

Vérifions enfin le théorème d'indépendance (8). Si X est une VA positive indépendante de $\mathcal{G}$ on a $E(X \wedge n|\mathcal{G}) = E(X \wedge n)$ d'ou le résultat le par limite croissante. Pour X de signe quelconque on décompose $X = X^+ - X^-$.

2.6.2 Exemple : espérance conditionnelle d'une VAR non intégrable

Soit $\Omega = [0,1], \mathcal{F} = \mathcal{B}([0,1]), P = \lambda$ (mesure de Lebesgue), posons $X(u) = u^{-2}$ pour $u \in]0,1]$ et $X(0) = 0$. On a $\int_{[0,1]} XdP = \infty$ la VAR X n'est pas intégrable mais on va montrer qu'elle a une espérance conditionnelle finie pour une sous tribu $\mathcal{G}$ de $\mathcal{F}$ et une espérance conditionnelle infinie pour une sous tribu $\mathcal{H}$ de $\mathcal{G}$. Soit $\mathcal{G} = \sigma(A_i \mid i \in \mathbb{N})$ la tribu sur Ω engendrée par la suite d'intervalles $A_0 = [0]$ puis pour $i \geq 1$, $A_i =]\frac{1}{i+1}, \frac{1}{i}]$. On a $P(A_0) = 0$, $P(A_i) = \frac{1}{i(i+1)}$ et $\int_{A_i} XdP = \frac{1}{i(i+1)}$ d'où :

$$E(X|\mathcal{G}) = \sum_{i\geq 1} \frac{\int_{A_i} XdP}{P(A_i)} 1_{A_i} = \sum_{i\geq 1} 1_{A_i} < \infty$$

donc $X \in \mathcal{F}_\mathcal{G}$. En revanche si $\mathcal{H} = \sigma(A_i \mid 1 \leq i \leq m+1)$ est la sous tribu de $\mathcal{G}$ sur Ω engendrée par la suite finie d'intervalles $A_i =]\frac{1}{i+1}, \frac{1}{i}]$ où $1 \leq i \leq m$ et $A_{m+1} = [0, \frac{1}{m+1}]$. Comme pour $n \geq (m+1)^2$ on a $\int_0^{\frac{1}{m+1}} (u^{-2} \wedge n) du \geq n^{\frac{1}{2}}$ il vient :

$$E(X \wedge n \mid \mathcal{H}) = \sum_{i=1}^{m+1} \frac{\int_{A_i} (X \wedge n) dP}{P(A_i)} 1_{A_i} \geq \frac{\int_{A_{m+1}} (X \wedge n) dP}{P(A_{m+1})} 1_{A_{m+1}} \geq (m+1) n^{\frac{1}{2}} 1_{A_{m+1}}$$

donc $\lim_n \uparrow E(X \wedge n \mid \mathcal{H})(u) = \infty$ pour $u \in A_{m+1}$ et $X \notin \mathcal{F}_\mathcal{H}$ $\square$

2.6.3 Proposition : espérance conditionnelle dans $L^1(\mathcal{F})$

Pour toute sous tribu $\mathcal{G}$ de $\mathcal{F}$ toute VAR $X \in L^1(\mathcal{F})$ à une espérance conditionnelle $E(X|\ \mathcal{G})$ dans $\mathbb{R}$ soit $L^1(\mathcal{F}) \subseteq \mathcal{F}_\mathcal{G}$. En outre l'espérance conditionnelle $Y = E(X|\ \mathcal{G})$ est caractérisée à un négligeable près par $Y \in L^1(\mathcal{G})$ et $\forall Z \in b\mathcal{G}, E(XZ) = E(YZ)$.

Démonstration

Si $X \in L^1(\mathcal{F})$ est positif alors $E(X|\ \mathcal{G}) = \lim_n \uparrow E(X \wedge n \mid \mathcal{G})$ et puisque $X \wedge n \in L^2(\mathcal{F})$ on a $E[E(X \wedge n \mid \mathcal{G})] = E(X \wedge n) \leq E(X)$ donc $E(X|\ \mathcal{G}) < \infty$ PS et X a une espérance conditionnelle dans $\mathbb{R}$ intégrable. Pour $X \in L^1(\mathcal{F})$ de signe quelconque on décompose $X = X^+ - X^-$. L'espérance conditionnelle $Y = E(X|\ \mathcal{G}) \in L^1(\mathcal{G})$ pour $X \in L^1(\mathcal{F})$ vérifie $\forall Z \in b\mathcal{G}, E(XZ) = E(YZ)$ et cette égalité est caractéristique puisque pour tout $Y_1, Y_2 \in L^1(\mathcal{F})$ on a $\forall A \in \mathcal{G},\ E(X, A) = E(Y_1, A) = E(Y_2, A) \Longrightarrow Y_1 = Y_2$ PS $\square$

2.6.4 Proposition : famille $[E(X|\ \mathcal{G})]_\mathcal{G}$

Si X est intégrable, alors la famille $[E(X|\ \mathcal{G})]_\mathcal{G}$ ou $\mathcal{G}$ est une sous tribu de $\mathcal{F}$ est uniformément intégrable (UI).

Démonstration

Posons $A_M = [|E(X|\ \mathcal{G})| > M]$ on a $A_M \in \mathcal{G}$ et

$$E[|E(X|\ \mathcal{G})|, A_M] \leq E[E(|X| \mid \mathcal{G}), A_M] = E(|X|, A_M)$$

Par ailleurs :

$$MP(A_M) \leq E[E(X|\ \mathcal{G})|, A_M] \leq E[E(X|\ \mathcal{G})] \leq E(|X|)$$

donc $M \to \infty \Longrightarrow P(A_M) \to 0 \Longrightarrow E(|X|, A_M) \to 0 \Longrightarrow E[|E(X|\ \mathcal{G})|, A_M] \to 0$ $\square$

Chapitre 3

Capacité

On démontre de manière élémentaire le théorème de capacité de Choquet qui nous permettra d'établir les théorèmes de projection mesurable et de section de Meyer. L'étape essentielle est un résultat ensembliste sur les ensembles pavés : le théorème d'approximation par dedans (§3.2.2) que l'on établit à l'aide des lissages où rabotages de Sierpinski [5].

On définit d'abord un *pavage* sur un ensemble non vide E comme une famille $\mathcal{E}$ de parties de E contenant E et $\emptyset$ et stable par union et intersection finies. On appelle alors *pavé* un élément de $\mathcal{E}$, *mozaique* un pavage stable par union et par intersection dénombrables et on note $m(\mathcal{E})$ la mozaique engendrée par $\mathcal{E}$ c'est-à-dire la plus petite mozaique contenant $\mathcal{E}$. On note aussi $\mathcal{C}$ une famille de parties de E *ascendante* dans E c'est-à-dire telle que :

$$\forall A \in \mathcal{C},\ \forall B \subseteq E,\ A \subseteq B \Longrightarrow B \in \mathcal{C} \quad \text{et} \quad \forall (A_n) \subseteq \mathcal{C},\ \cup_n A_n \in \mathcal{C} \Longrightarrow \exists k, A_k \in \mathcal{C}$$

Dans ces notations pour tout $A \in \mathcal{C} \cap m(\mathcal{E})$ il existe une suite (B_n) décroissante dans $\mathcal{C} \cap m(\mathcal{E})$ telle que $\cap_n B_n \subseteq A$ c'est le théorème d'approximation par dedans.

Une *capacité* relative au pavage $\mathcal{E}$ est une application I croissante de l'ensemble des parties $\mathcal{P}_E$ de E dans $\mathbb{R} \cup \{-\infty, \infty\}$ continue pour les suites croissante dans $\mathcal{P}_E$ et pour les suites décroissante dans $\mathcal{E}$. Une partie $A \subseteq E$ est *capacitable* pour I si

$$I(A) = \sup\{I(\cap_n B_n) \mid B_n \in \mathcal{E},\ \cap_n B_n \subseteq A\}$$

Le théorème de Choquet s'énonce simplement :

Toute partie de la mosaique $m(\mathcal{E})$ est capacitable pour toute capacité relative à $\mathcal{E}$

Le théorème de section mesurable de Meyer qui en est une conséquence s'écrit :

> Si A est un ensemble aléatoire mesurable c'est-à-dire une partie de l'ensemble produit $\mathbb{R}^+ \times \Omega$ appartenant à la tribu $\mathcal{B} \otimes \mathcal{F}$ il existe une variable aléatoire positive T dont le graphe $[T]$ est inclu dans A et qui *sectionne* A presque sûrement au sens où la mesure (probabilité) des projections sur Ω de A et de $[T]$ sont égales.

Dans le cas ou Ω est muni d'une filtration et si l'ensemble A est optionnel ou prévisible le temps T peut être un temps d'arrêt (temps optionnel) ou un temps prévisible (théorèmes de section optionnel où prévisible §6.4.3 et §6.4.4). Ces théorèmes sont notamment utilisés pour montrer l'unicité des projections optionnelle et prévisible (§9.1.2 et 9.1.3).

3.1 Espace pavé

3.1.1 Définitions : pavage et mosaique

Un *pavage* sur un ensemble non vide E est une famille $\mathcal{E}$ de parties de E contenant E et $\emptyset$ et stable par union finie (notée $\cup_f$) et par intersection finie (notée $\cap_f$). On appelle *pavé* un élément de $\mathcal{E}$ et *espace pavé* le couple $(E, \mathcal{E})$.
Si $(E, \mathcal{E})$ et $(F, \mathcal{F})$ sont des espaces pavés on appelle *pavage produit* le pavage $E \otimes F$ sur $E \times F$ constitué des unions finis d'ensembles de la forme $A \times B$ où $A \in \mathcal{E}$ et $B \in \mathcal{F}$.
On appelle *mosaïque* sur un ensemble non vide E une famille de parties de E contenant E et $\emptyset$ et stable par union dénombrable (notée $\cup_\delta$) et par intersection dénombrable (notée $\cap_\delta$), c'est donc un pavage. Si $\mathcal{E}$ est un pavage sur E on note $m(\mathcal{E})$ la plus petite mosaïque contenant $\mathcal{E}$. C'est la mosaïque sur E engendrée par le pavage $\mathcal{E}$.
On dit qu'un pavage $\mathcal{E}$ est *monotone* s'il est stable par limite séquentielle monotone c'est-à-dire par union séquentielle croissante (notée $\cup_c$) et par intersection séquentielle décroissante (notée $\cap_d$).

3.1.2 Proposition : mosaïque et pavage monotone

Une famille $\mathcal{E}$ de parties de E est une mosaïque si et seulement si c'est un pavage monotone.

Démonstration
Une mosaïque est trivialement un pavage monotone. Inversement si $\mathcal{E}$ est un pavage monotone et (A_k) une suite quelconque dans E alors (B_n) définie par $B_n = \cup_{k \leq n} A_k$ est une suite croissante dans $\mathcal{E}$ et on a $\cup_k A_k = \cup_n B_n \in \mathcal{E}$. De même (C_n) définie par $C_n = \cap_{k \leq n} A_k$ est une suite décroissante dans $\mathcal{E}$ et on a $\cap_k A_k = \cap_n C_n \in \mathcal{E}$. $\square$

3.1.3 Proposition : mosaïque engendrée

Si $\mathcal{E}$ est un pavage sur E alors la mosaïque $m(\mathcal{E})$ engendrée par $\mathcal{E}$ est la plus petite classe de partie de E stable par $\cup_c$ et $\cap_d$ et contenant $\mathcal{E}$.

Démonstration
Notons $\mathcal{M}$ la plus petite classe de partie de E stable par $\cup_c$ et $\cap_d$ et contenant $\mathcal{E}$. Trivialement $\mathcal{M} \subseteq m(\mathcal{E})$. Pour établir l'égalité il suffit de montrer que $\mathcal{M}$ est une mosaïque ou ce qui est équivalent un pavage monotone et comme $\mathcal{M}$ est stable par $\cup_c$ et $\cap_d$ il suffit de vérifier que $\mathcal{M}$ est un pavage. Pour tout $B \subseteq E$ posons :

$$\Gamma(B) = \{A \subseteq E \mid A \cup B \in \mathcal{M}, \text{ et } A \cap B \in \mathcal{M}\}$$

Il vient :

1. $B \in \mathcal{E} \Longrightarrow \mathcal{E} \subseteq \Gamma(B)$ car $\mathcal{E}$ est stable par $\cup_f$ et $\cap_f$
2. $B \in \mathcal{E} \Longrightarrow \mathcal{M} \subseteq \Gamma(B)$ car (1) et $\Gamma(B)$ est stable par $\cup_c$ et $\cap_d$
3. $B \in \Gamma(A) \Longleftrightarrow A \in \Gamma(B)$
4. $A \in \mathcal{M}, B \in \mathcal{E} \Longrightarrow A \in \Gamma(B) \Longrightarrow B \in \Gamma(A)$
5. $A \in \mathcal{M} \Longrightarrow \mathcal{E} \subseteq \Gamma(A)$
6. $A \in \mathcal{M} \Longrightarrow \mathcal{M} \subseteq \Gamma(A)$

Ce qui signifie que $\mathcal{M}$ est stable par $\cup_f$ et $\cap_f$ donc $\mathcal{M}$ est un pavage. $\square$

3.1.4 Définition et proposition : pavage compact

Un pavage $\mathcal{K}$ sur un ensemble non vide E est dit *compact* si une des assertions équivalentes suivantes est satisfaite :

1. Si $(K_n) \subseteq \mathcal{K}$ est une suite décroissante de pavés non vides alors la limite $\cap_n K_n$ est non vide
2. Si $(K_n) \subseteq \mathcal{K}$ est une suite de pavés telle que pour tout entier N, $\cap_{n\leq N} K_n \neq \emptyset$ alors $\cap_n K_n \neq \emptyset$
3. Si $(K_n) \subseteq \mathcal{K}$ est une suite de pavés telle que $\cap_n K_n = \emptyset$ alors il existe un entier N telle que $\cap_{n\leq N} K_n = \emptyset$

Si E est un espace topologique séparé la classe $\mathcal{K}$ des parties compactes de E complétée par E lui-même constitue un pavage compact. On convient que $\emptyset$ est compact.

Démonstration
L'assertion (1) est un cas particulier de (2) et (1) implique (2).
En posant $C_N = \cap_{n\leq N} K_n$ on obtient une suite décroissante de pavés non vides donc $\cap_N C_N = \cap_n K_n \neq \emptyset$. Enfin (3) est la contraposée de (2).
Montrons le dernier point, soit (K_n) une suite décroissante de pavés non vides de E. Si la suite (K_n) est constante égale à E on a $\cap_n K_n = E \neq \emptyset$. Sinon on peut supposer $K_0 \neq E$. Posons alors pour $n \geq 1, G_n = K_0 - K_n$ le pavé K_n est un fermé de l'espace compact K_0 et G_n est un ouvert de K_0 donc par définition des sous espaces compacts :

$$\cup_{n\geq 1} G_n = K_0 \Longrightarrow \exists N \geq 1,\ \cup_{1\leq n\leq N} G_n = K_0$$

Ce qui équivaut à :

$$\forall N \geq 1, \cap_{1\leq n\leq N} K_n = K_N \neq \emptyset \Longrightarrow \cap_{n\geq 1} K_n \neq \emptyset$$

C'est la définition (1) d'un pavage compact. $\square$

3.1.5 Définition et proposition : ensemble $\mathcal{K}_\delta$ et limites décroissantes

Pour tout ensemble de parties $\mathcal{K}$ d'un ensemble non vide E on note $\mathcal{K}_\delta$ le plus petit ensemble de parties de E stable par intersections dénombrables et contenant $\mathcal{K}$ et

$\mathcal{K}_{\lim\downarrow}$ l'ensemble des limites des suites décroissantes d'élèments de $\mathcal{K}$. Lorsque $\mathcal{K}$ est stable par intersection finies on a l'égalité :

$$\mathcal{K}_\delta = \mathcal{K}_{\lim\downarrow}$$

En particulier lorsque $\mathcal{K}$ est un pavage l'ensemble $\mathcal{K}_\delta$ est constitué des intersections dénombrables d'éléments de $\mathcal{K}$.

Démonstration

On a $\mathcal{K} \subseteq \mathcal{K}_{\lim\downarrow} \subseteq \mathcal{K}_\delta$ donc si $\mathcal{K}_{\lim\downarrow}$ est stable par $\cap_\delta$ on aura $\mathcal{K}_{\lim\downarrow} = \mathcal{K}_\delta$

Vérifions que $\mathcal{K}_{\lim\downarrow}$ est stable par $\cap_\delta$. Soit (A_n) une suite quelconque dans $\mathcal{K}_{\lim\downarrow}$ on peut écrire $A_n = \cap_k A_{n,k}$ où $(A_{n,k})_{k\in\mathbb{N}}$ est une suite décroissante dans $\mathcal{K}$ donc si on pose $B_m = \cap_{n\leq m,k\leq m} A_{n,k}$ on obtient $\cap_n A_n = \cap_m B_m$ et (B_m) est une suite décroissante dans $\mathcal{K}$ ce qui entraîne $\cap_n A_n \in \mathcal{K}_{\lim\downarrow}$. $\square$

3.1.6 Proposition : pavage compact $\mathcal{K}_\delta$

Si $\mathcal{K}$ est un pavage compact sur un ensemble non vide E alors $\mathcal{K}_\delta$ est aussi un pavage compact.

Démonstration

L'ensemble $\mathcal{K}_\delta$ est constitué des intersections dénombrables de pavés de $\mathcal{K}$ il est donc stable par union finie puisque si $K_1 = \cap_i K_{1,i}$ et $K_2 = \cap_j K_{2,j}$ où $K_{1,i}, K_{1,i} \in \mathcal{K}$ alors $K_1 \cup K_2 = \cap_{i,j}(K_{1,i} \cup K_{i,j})$. Comme il est trivialement stable par intersection finie il s'agit bien d'un pavage. Reste à montrer que toute suite (K_n) décroissante de pavés non vides de $\mathcal{K}_\delta$ a une intersection non vide, il vient :

$$\cap_n K_n = \cap_{n,i} K_{n,i} = \cap_m \cap_{n\leq m, i\leq m} K_{n,i} \neq \emptyset \text{ car } \cap_{n\leq m,i\leq m} K_{n,i} \supseteq \cap_{n\leq m} K_n = K_m \neq \emptyset$$

3.1.7 Définition : projections et coupes

Soitent E, F deux ensembles non vides et $\pi : E \times F \to F$ le projecteur sur F défini par $\pi(x, y) = y$. Pour tout ensemble $H \subseteq E \times F$ la projection $\pi(H)$ de H sur F est :

$$\pi(H) = \{y \in F \mid \exists x \in F, (x, y) \in H\}$$

et la coupe $H(y)$ de H en $y \in F$ est :

$$H(y) = \{x \in E \mid (x, y) \in H\}$$

3.1.8 Proposition : propriétés des projections et coupes

1. Appartenance à la projection et coupe non vide : $\forall H \in E \times F$ et $\forall y \in F$

$$y \in \pi(H) \Longleftrightarrow H(y) \neq \emptyset$$

2. Projection et coupe d'une union : $\forall (H_n) \subseteq E \times F$ et $\forall y \in F$

$$\pi(\cup_n H_n) = \cup_n \pi(H_n) \text{ et } (\cup_n H_n)(y) = \cup_n H(y)$$

3. Projection et coupe d'une intersection : $\forall (H_n) \subseteq E \times F$ et $\forall y \in F$

$$\pi(\cap_n H_n) \subseteq \cap_n \pi(H_n) \text{ et } (\cap_n H_n)(y) = \cap_n H(y)$$

Démonstration
Il suffit d'expliciter les définitions §3.1.7.

1. $y \in \pi(H) \Longleftrightarrow \exists x \in E, (x,y) \in H \Longleftrightarrow \exists x \in E, x \in H(y) \Longleftrightarrow H(y) \neq \emptyset$
2. $\pi(\cup_n H_n) = \{y \in F \mid \exists x \in E, (x,y) \in \cup_n H_n\} = \cup_n \{y \in F \mid \exists x \in E,$
$(x,y) \in H_n\} = \cup_n \pi(H_n)$
$(\cup_n H_n)(y) = \{x \in E \mid (x,y) \in \cup_n H_n\} = \cup_n \{x \in E \mid (x,y) \in H_n\} = \cup_n H_n(y)$
3. $\pi(\cap_n H_n) = \{y \in F \mid \exists x \in E, (x,y) \in \cap_n H_n\}$ et
$\cap_n \pi(H_n) = \cap_n \{y \in F \mid \exists x_n \in E, (x_n, y) \in H_n\}$
$\Longrightarrow \pi(\cap_n H_n) \subseteq \cap_n \pi(H_n)$
$(\cap_n H_n)(y) = \{x \in E \mid (x,y) \in \cap_n H_n\} = \cap_n \{x \in E \mid (x,y) \in H_n\} = \cap_n H_n(y)$

Exemple : inégalité stricte dans (3)
On suppose $E = \mathbb{R}^+$ et $F = \mathbb{R}$, soit $H_n =]0, 2^{-n}] \times \mathbb{R}$ la suite (H_n) est décroissante, $\pi(H_n) = \mathbb{R}$ et $\cap_n H_n = \emptyset$ on obtient donc :

$$\pi(\cap_n H_n) = \emptyset \subset \mathbb{R} = \cap_n \pi(H_n)$$

3.1.9 Lemme de projection compacte

Soit E et F deux ensembles non vides et π la projection de $E \times F$ sur F. Si $\mathcal{H}$ est un pavage de $E \times F$ tel que pour tout $y \in F$ l'ensemble $\mathcal{K}_y$ des coupes $H(y)$ des pavés $H \in \mathcal{H}$ suivant $y \in F$ est un pavage compact sur E alors pour toute suite (H_n) décroissante dans H_δ on a :

$$\pi(\cap_n H_n) = \cap_n \pi(H_n)$$

Corollaire
Si le pavage $\mathcal{H}$ sur $E \times F$ est le pavage produit $\mathcal{K} \otimes \mathcal{F}$ où $\mathcal{K}$ est un pavage compact sur E et $\mathcal{F}$ un pavage quelconque sur F les coupes $H(y)$ sont des pavés de $\mathcal{K}$ donc l'ensemble $\mathcal{K}_y$ est un pavage compact de E la condition de compacité du lemme est satisfaite et on a :

$$(H_n) \subseteq (\mathcal{K} \otimes \mathcal{F})_\delta, H_n \downarrow \Longrightarrow \pi(\cap_n H_n) = \cap_n \pi(H_n)$$

Démonstration du lemme
On a déja $\pi(\cap_n H_n) \subseteq \cap_n \pi(H_n)$ il faut donc montrer $\cap_n \pi(H_n) \subseteq \pi(\cap_n H_n)$ il vient :

1. $y \in \cap_n \pi(H_n) \Longleftrightarrow \forall n, y \in \pi(H_n) \Longleftrightarrow \forall n, H_n(y) \neq \emptyset$
2. $y \in \pi(\cap_n H_n) \Longleftrightarrow (\cap_n H_n)(y) = \cap_n H_n(y) \neq \emptyset$
3. $H_n \in \mathcal{H}_\delta \Longrightarrow H_n = \cap_m H_{n,m}$ où $H_{n,m} \in \mathcal{H}$
4. $H_n(y) = \cap_m H_{n,m}(y)$ où $H_{n,m}(y) \in \mathcal{K}_y$
5. $H_n(y) \in \mathcal{K}_{y\delta} \Longrightarrow \cap_n H_n(y) \neq \emptyset$ car $\mathcal{K}_{y\delta}$ est un pavage compact et pour tout n, $H_n(y) \neq \emptyset$. $\square$

3.2 Théorème d'approximation par dedans

3.2.1 Définition : famille ascendante ou capacitance

Une famille $\mathcal{C}$ de parties d'un ensemble non vide E est une *famille ascendante* dans E ou *capacitance* si :

1. $A \in \mathcal{C},\ B \subseteq E$ et $A \subseteq B \Longrightarrow B \in \mathcal{C}$
2. Si (A_n) est une suite croissante de parties de E telle que $\cup_n A_n \in \mathcal{C}$ il existe $k \in \mathbb{N}$ tel que $A_k \in \mathcal{C}$ (compte tenu de (1) la suite croissante (A_n) est dans $\mathcal{C}$ à partir du rang k).

3.2.2 Théorème d'approximation par dedans

Soit $(E, \mathcal{E})$ un espace pavé, $m(\mathcal{E})$ la mosaïque engendrée par $\mathcal{E}$ et $\mathcal{C}$ une famille ascendante dans E. Si $A \in \mathcal{C} \cap m(\mathcal{E})$ il existe une suite décroissante (B_n) d"éléments de $\mathcal{C} \cap m(\mathcal{E})$ telle que $\cap_n B_n \subseteq A$.

La suite de ce paragraphe est consacrée à réunir les éléments pour montrer ce résultat. L'étape essentiel est le théorème de Sierpinski (1882-1969). On notera $(E, \mathcal{E})$ un espace pavé et $\mathcal{C}$ une famille ascendante dans E.

3.2.3 Définition : enveloppe

Soit (A_n) une suite décroissante de parties quelconques de E. Un ensemble $A \subseteq E$ est appelé *enveloppe* de la suite (A_n) s'il existe une suite décroissante (B_n) de pavés qui enveloppe la suite (A_n) et dont la limite est dans A :

$$\forall n, A_n \subseteq B_n \text{ et } \cap_n B_n \subseteq A$$

3.2.4 Définitions : lissage, suite lissée

Soit $\mathcal{P}_E$ l'ensemble des parties de E. Un *lissage* dans E associé à $\mathcal{C}$ est une suite $(L_n)_{n\in\mathbb{N}^*}$ d'applications de $\mathcal{P}_E^n$ dans $\mathcal{P}_E$ telle que $\forall n \geq 1, \forall (X_1, \ldots X_n) \in \mathcal{P}_E^n$

$$L_n(X_1, \ldots X_n) \subseteq X_n \text{ et } X_n \in \mathcal{C} \Longrightarrow L_n(X_1, \ldots X_n) \in \mathcal{C}$$

On dit qu'un lissage $(L_n)_{n\in\mathbb{N}^*}$ *lisse* une suite $(A_n)_{n\in\mathbb{N}^*}$ d'éléments de $\mathcal{C}$ si $\forall n \geq 1$,

$$A_{n+1} \subseteq L_n(A_1, \ldots A_n)$$

Une suite $(A_n)_{n\in\mathbb{N}^*} \subseteq \mathcal{C}$ est dite *lissée* s'il existe un lissage $(L_n)_{n\in\mathbb{N}^*}$ dans E associé à $\mathcal{C}$ qui la lisse. Notons qu'une suite lissée est décroissante.

Exemples
L'exemple le plus simple de lissage est la suite définie par $L_n(X_1, \ldots X_n) = X_n$ c'est le *lissage identité*.

Si $A \in \mathcal{C}$ et si $(L_n)_{n\in\mathbb{N}^*}$ est un lissage associé à $\mathcal{C}$ alors la suite $(A_n)_{n\in\mathbb{N}^*}$ définie par $A_1 = A$ et $A_{n+1} = L_n(A_1, \ldots A_n)$ est une suite de $\mathcal{C}$ lissée par $(L_n)_{n\in\mathbb{N}^*}$. On l'appelle *suite déduite* de A par $(L_n)_{n\in\mathbb{N}^*}$.

3.2.5 Définitions : lissage compatible, ensembles lisses

On dit qu'un lissage $(L_n)_{n\in\mathbb{N}^*}$ est *compatible* avec une partie A de E si A est enveloppe de toute suite décroissante $(A_n)_{n\in\mathbb{N}^*} \subseteq \mathcal{C}$ lissée par $(L_n)_{n\in\mathbb{N}^*}$ et partant dans A c'est-à-dire $A_1 \subseteq A$. On dit aussi dans ce cas que A est *compatible* avec $(L_n)_{n\in\mathbb{N}^*}$
Une partie A de E est *lisse* s'il existe un lissage compatible avec A.

Notation
On note $\alpha : (p,q) \to \alpha(p,q)$ une bijection de $\mathbb{N}^{*2}$ sur $\mathbb{N}^*$ strictement croissante pour chacune des variables p et q. Par exemple $\alpha(p,q) = 2^{p-1}(2q-1)$

3.2.6 Définition et proposition : mélange de lissages

Soit $(L^p)_{p\in\mathbb{N}^*}$ une suite de lissages. Pour tout $n \geq 1$ il existe un unique couple (p,q) tel que $\alpha(p,q) = n$ et on a $\alpha(p,j) \leq n$ pour tout $1 \leq j \leq q$ puisque α est croissante. On peut donc définir pour tout $(X_1, \dots X_n) \in \mathcal{P}_E^n$

$$L_n(X_1, \dots X_n) = L_q^p(X_{\alpha(p,1)}, \dots X_{\alpha(p,q)})$$

La suite d'applications $L = (L_n)_{n\in\mathbb{N}^*}$ est un lissage appelé *mélange* de la suite de lissages $(L^p)_{p\in\mathbb{N}^*}$ pour la bijection α. De plus si une partie $A \subseteq E$ est compatible avec un des lissages L^p alors elle est compatible avec le mélange L.

Démonstration
La suite d'applications $L = (L_n)_{n\in\mathbb{N}^*}$ est un lissage car L^p est un lissage donc

1. $L_n(X_1, \dots X_n) = L_q^p(X_{\alpha(p,1)}, \dots X_{\alpha(p,q)}) \subseteq X_{\alpha(p,q)} = X_n$
2. $X_n \in \mathcal{C} \Longrightarrow X_{\alpha(p,q)} \in \mathcal{C} \Longrightarrow L_n(X_1, \dots X_n) = L_q^p(X_{\alpha(p,1)}, \dots X_{\alpha(p,q)}) \in \mathcal{C}$

Soit $A \subseteq E$ compatible avec le lissage L^p et $(A_n)_{n\in\mathbb{N}^*}$ une suite de $\mathcal{C}$ lissée par L telle que $A_1 \subseteq A$ il faut montrer que A est une enveloppe de $(A_n)_{n\in\mathbb{N}^*}$. On va établir :

1. $(A_{\alpha(p,q)})_{q\in\mathbb{N}^*} \subseteq \mathcal{C}$ est une suite lissée par L^p
2. A est une enveloppe de $(A_{\alpha(p,q)})_{q\in\mathbb{N}^*}$
3. A est une enveloppe de $(A_n)_{n\in\mathbb{N}^*}$

Montrons (1), il vient avec $n = \alpha(p,q)$

$$A_{\alpha(p,q+1)} \subseteq A_{\alpha(p,q)+1} = A_{n+1} \subseteq L_n(A_1, \dots A_n) = L_q^p(A_{\alpha(p,1)}, \dots A_{\alpha(p,q)})$$

Montrons (1) implique (2), si A est compatible avec L^p, A est une enveloppe de toute suite de $\mathcal{C}$ lissée par L^p de premier terme inclus dans A or la suite $(A_{\alpha(p,q)})_{q\in\mathbb{N}^*}$ de $\mathcal{C}$ est lissée par L^p et $A_{\alpha(p,1)} \subseteq A$ car $A_{\alpha(p,1)} \subseteq A_1 \subseteq A$ donc A est une enveloppe de $(A_{\alpha(p,q)})_{q\in\mathbb{N}^*}$.

Montrons (2) implique (3), si A est une enveloppe de $(A_{\alpha(p,q)})_{q\in\mathbb{N}^*}$ il existe une suite décroissante $(B_q)_{q\in\mathbb{N}^*}$ de pavés qui enveloppe $(A_{\alpha(p,q)})_{q\in\mathbb{N}^*}$ et telle que $\cap_q B_q \subseteq A$. On peut alors construire une suite décroissante $(C_n)_{n\in\mathbb{N}^*}$ de pavés qui enveloppe $(A_n)_{n\in\mathbb{N}^*}$ et telle que $\cap_n C_n \subseteq A$. Il suffit en effet de poser $C_n = E$ pour $1 \leq n < \alpha(p,1)$ et $C_n = B_q$ pour $\alpha(p,q) \leq n < \alpha(p,q+1)$. $\square$

3.2.7 Corollaire : lissage compatible avec une suite de parties lisses

Si $(A^p)_{p\in\mathbb{N}^*}$ est une suite de parties lisse de E il existe un lissage compatible avec tous les ensembles A^p.

Démonstration
Soit pour tout $p \geq 1$, L^p un lissage compatible avec A^p et L le mélange des lissages L^p pour une bijection α quelconque. D'après la proposition §3.2.6 ci-dessus comme pour tout $p \geq 1, A^p$ est compatible avec L^p le mélange L est compatible avec A^p. □

Lemme I
L'ensemble des enveloppes d'une suite décroissante (A_n) de parties de E est stable par intersections dénombrable (stabilité $\cap_\delta$).

Démonstration
Si (A^p) est une suite d'enveloppes de la suite décroissante (A_n) de parties de E on va montrer que $A = \cap_p A^p$ est aussi une enveloppe de la suite (A_n). On va pour cela construire une suite décroissante (C_n) de pavés qui enveloppe la suite (A_n) et dont la limite est dans A. L'ensemble A^p est pour tout p une enveloppe de la suite décroissante (A_n). Il existe donc une suite $(B_{p,n})$ décroissante en n de pavés qui enveloppe la suite (A_n) et dont la limite est dans A^p. Posons $C_n = \cap_{p\leq n} B_{p,n}$ on obtient une suite décroissante de pavés qui enveloppe la suite (A_n) et dont l'intersection est dans A.

Lemme II
Tout pavé $A \in \mathcal{E}$ est lisse.

Démonstration
Si $A \notin \mathcal{C}$ alors pour tout lissage il n'existe pas de suite lissée (A_n) dans $\mathcal{C}$ telle que $A_1 \subseteq A$. Donc tout lissage est compatible et par conséquent A est lisse. Sinon comme $A \in \mathcal{C}$ est un pavé le lissage identique est compatible avec A.

3.2.8 Théorème de Sierpinski : une mosaïque est lisse

L'ensemble des parties lisses de E est stable par union croissante (stabilité $\cup_c$) et intersection dénombrable (stabilité $\cap_\delta$) par conséquent tout les éléments de la mosaïque $m(\mathcal{E})$ sont lisses.

Démonstration
La stabilité des parties lisses de E par intersection dénombrable (stabilité $\cap_\delta$) entraîne la stabilité par intersection décroissante (stabilité $\cap_d$). Comme tout pavé $A \in \mathcal{E}$ est lisse (lemme II ci-dessus) et que la mosaïque $m(\mathcal{E})$ est la plus petite classe de parties de E stable par $\cup_c$ et $\cap_d$ et contenant $\mathcal{E}$ (§3.1.3) tout les éléments de la mosaïque $m(\mathcal{E})$ sont lisses.

Montrons la stabilité $\cap_\delta$ des parties lisses de E, soit $(A^p)_{p\in\mathbb{N}^*}$ une suite de parties lisses et $L = (L_n)_{n\in\mathbb{N}^*}$ un lissage compatible avec tous les A^p (qui existe d'après le corollaire §3.2.6). Posons $A = \cap_p A^p$, il faut montrer que A est lisse. On va pour cela montrer que L est compatible avec A. Si (A_n) est une suite décroisssante dans $\mathcal{C}$ lissée par L telle que $A_1 \subseteq A$ il faut vérifier que A est une enveloppe de (A_n). Comme $A_1 \subseteq A = \cap_p A^p \Longrightarrow \forall p, A_1 \subseteq A^p$ et que L est compatible avec A^p l'ensemble A^p est une enveloppe de (A_n). Or l'ensemble des enveloppes d'une suite décroissante (A_n) de