

# Chapitre 11

## Espérance conditionnelle et martingales

A l'origine, une martingale est une technique utilisée par les joueurs dans les jeux de hasard pour tenter d'infléchir le hasard en leur faveur, avec cependant peu de réussite en pratique ! Nous allons ici donner une introduction à la théorie des martingales en probabilités.

### 11.1 Espérance conditionnelle

Nous commençons par définir l'espérance conditionnée par une tribu.

**Définition 11.1 (Espérance conditionnelle d'une v.a.r.)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $X$  une v.a.r. intégrable et  $\mathcal{B}$  une tribu incluse dans  $\mathcal{A}$ . On appelle espérance conditionnée par  $\mathcal{B}$  de  $X$  ou espérance conditionnelle de  $X$  par rapport à  $\mathcal{B}$  l'ensemble des applications  $Z$  de  $\Omega$  dans  $\mathbb{R}$ ,  $\mathcal{B}$ -mesurables, intégrables et t.q. :

$$E(ZU) = E(XU) \text{ pour toute application } U : \Omega \rightarrow \mathbb{R}, \mathcal{B}\text{-mesurable, bornée.} \quad (11.1)$$

On note  $E(X|\mathcal{B})$  cette espérance conditionnelle (qui est donc un ensemble de fonctions, on montre plus loin que c'est un élément de  $L^1_{\mathbb{R}}(\Omega, \mathcal{B}, P)$ .) (Noter que dans (11.1), les applications  $ZU$  et  $XU$  sont bien des v.a.r. intégrables.)

Cette définition peut sembler un peu abrupte. On montre dans la proposition 11.2 que, sous les hypothèses de la définition 11.1, l'espérance conditionnelle existe, c'est-à-dire que l'ensemble  $E(X|\mathcal{B})$  est non vide, et que  $E(X|\mathcal{B})$  est unique, ceci signifiant que si  $Z_1, Z_2 \in E(X|\mathcal{B})$ , on a nécessairement  $Z_1 = Z_2$  p.s.. En fait,  $E(X|\mathcal{B})$  est donc un élément de  $L^1_{\mathbb{R}}(\Omega, \mathcal{B}, P)$ .

L'ensemble  $E(X|\mathcal{B})$ , défini dans la définition 11.1, est donc un ensemble de v.a.r. (car  $Z$   $\mathcal{B}$ -mesurable implique  $Z$   $\mathcal{A}$ -mesurable) et, en pratique, on confond, comme d'habitude, cet ensemble avec l'un de ces éléments (on confond un élément de  $L^1_{\mathbb{R}}(\Omega, \mathcal{B}, P)$  avec l'un de ses représentants). Si  $Z$  est une v.a.r.  $\mathcal{B}$ -mesurable intégrable et t.q.  $E(ZU) = E(XU)$  pour toute v.a.r.  $U$   $\mathcal{B}$ -mesurable bornée, on écrira donc  $Z = E(X|\mathcal{B})$  p.s. au lieu d'écrire  $Z \in E(X|\mathcal{B})$ .

Avant de démontrer l'existence et l'unicité de l'espérance conditionnelle, donnons quelques exemples simples. Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé et  $X$  une v.a.r. intégrable.

**Cas  $\mathcal{B}$  est la tribu grossière.** Prenons tout d'abord  $\mathcal{B} = \{\emptyset, \Omega\}$ . Il est alors facile de voir (exercice 11.1) que  $E(X|\mathcal{B})$  est réduit à un seul élément et que cet élément est la fonction constante et égale à  $E(X)$ .

**Cas  $\mathcal{B} = \mathcal{A}$ .** Si maintenant  $\mathcal{B} = \mathcal{A}$ , alors  $E(X | \mathcal{B}) = X$  p.s.. Plus précisément,  $E(X|\mathcal{B})$  est ici l'ensemble des v.a.r.  $Z$  t.q.  $Z = X$  p.s., c'est-à-dire  $X$  en tant que élément de  $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$ .

**Cas où  $\mathcal{B} = \{\emptyset, A, A^c, \Omega\}$ ,  $0 < P(A) < 1$ .** Soit maintenant  $A \in \mathcal{A}$  t.q.  $0 < P(A) < 1$  et  $\mathcal{B} = \{\emptyset, A, A^c, \Omega\}$  (qui est bien une tribu incluse dans  $\mathcal{A}$ ). On peut ici montrer (exercice 11.1) que  $E(X|\mathcal{B})$  est réduit à un seul élément et cet élément est la fonction  $Z$  définie par :

$$Z = \frac{E(X1_A)}{P(A)}1_A + \frac{E(X1_{A^c})}{P(A^c)}1_{A^c}.$$

La quantité  $\frac{E(X1_A)}{P(A)}$  s'appelle espérance de  $X$  sachant  $A$ . On a ainsi fait le lien entre espérance de  $X$  sachant un événement et espérance de  $X$  par rapport à une tribu (ou selon une tribu).

**Cas où  $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$ ,  $P(B) = 1$ .** On prend  $B \in \mathcal{A}$  t.q.  $P(B) = 1$  et  $B^c \neq \emptyset$  (c'est le cas, par exemple, si  $P$  est une mesure diffuse, que  $\mathcal{A}$  contient les singletons et que  $B^c$  est formé d'un nombre fini ou dénombrable de points de  $\Omega$ ). On prend encore  $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$ . Pour  $a \in \mathbb{R}$ , on pose  $Z_a = E(X)1_B + a1_{B^c}$ . On peut alors montrer, pour être précis, que  $E(X|\mathcal{B}) = \{Z_a, a \in \mathbb{R}\}$  (exercice 11.2) c'est-à-dire que  $E(X|\mathcal{B})$  est la fonction constante et égale à  $E(X)$  en tant que élément de  $L^1_{\mathbb{R}}(\Omega, \mathcal{B}, P)$ , ce qu'on écrit, avec la confusion habituelle entre un élément de  $L^1$  et l'un de ses représentants,  $E(X|\mathcal{B}) = E(X)$  p.s..

On montre maintenant l'existence et l'unicité de l'espérance, conditionnée par une tribu, d'une v.a.r. intégrable.

**Proposition 11.2** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé et  $\mathcal{B}$  une tribu incluse dans  $\mathcal{A}$ . Soit  $X$  une v.a.r. intégrable. Alors :

**(Existence)**  $E(X|\mathcal{B}) \neq \emptyset$ .

**(Unicité)**  $Z_1, Z_2 \in E(X|\mathcal{B}) \Rightarrow Z_1 = Z_2$  p.s..

DÉMONSTRATION – On démontre d'abord l'unicité de  $E(X|\mathcal{B})$ . Puis, on démontre l'existence de  $E(X|\mathcal{B})$  si  $X$  est de carré intégrable, puis l'existence si  $X$  est positive (et intégrable) et enfin l'existence si  $X$  est seulement intégrable. En fait, la démonstration de l'existence si  $X$  est de carré intégrable n'est pas utilisée pour la suite de la démonstration mais elle est éventuellement intéressante pour la compréhension de l'espérance conditionnelle.

**Unicité.** Soit  $Z_1, Z_2 \in E(X|\mathcal{B})$ . On pose  $U = \text{sign}(Z_1 - Z_2)$  (on rappelle que la fonction  $\text{sign}$  est définie par  $\text{sign}(s) = -1$  si  $s < 0$ ,  $\text{sign}(s) = 1$  si  $s > 0$  (et par exemple)  $\text{sign}(0) = 0$ ). Comme la fonction  $\text{sign}$  est borélienne de  $\mathbb{R}$  dans  $\mathbb{R}$  et que  $(Z_1 - Z_2)$  est  $\mathcal{B}$ -mesurable, la fonction  $U$  est bien  $\mathcal{B}$ -mesurable. Elle est aussi bornée, on a donc en utilisant la définition de l'espérance conditionnelle (11.1) avec  $Z = Z_1$  et  $Z = Z_2$ , on obtient

$$E(XU) = E(Z_1U) \text{ et } E(XU) = E(Z_2U).$$

Ceci donne  $E((Z_1 - Z_2)U) = 0$  et donc  $E(|Z_1 - Z_2|) = 0$ . On en déduit  $Z_1 = Z_2$  p.s..

**Existence si  $X$  est de carré intégrable.** On note  $P_{\mathcal{B}}$  la restriction de  $P$  (qui est une mesure sur  $\mathcal{A}$ ) à  $\mathcal{B}$  (tribu incluse dans  $\mathcal{A}$ ). La mesure  $P_{\mathcal{B}}$  est donc une probabilité sur  $\mathcal{B}$ . On note  $H$  l'espace de Hilbert  $L^2_{\mathbb{R}}(\Omega, \mathcal{B}, P_{\mathcal{B}})$  et, pour  $V \in H$ , on pose :

$$T(V) = \int_{\Omega} XVdP.$$

Il est clair que  $T(V)$  est bien définie. En étant précis, on remarque que  $T(V) = \int_{\Omega} XvdP$ , où  $v \in \mathcal{L}^2_{\mathbb{R}}(\Omega, \mathcal{B}, P_{\mathcal{B}})$  est un représentant de  $V$  (et cette quantité ne dépend pas du représentant choisi). C'est pour définir  $T$  que nous avons besoin que  $X$  soit de carré intégrable.

L'application  $T$  est linéaire continue de  $H$  dans  $\mathbb{R}$  (et on a  $\|T\| \leq \|X\|_2$ ). On peut donc appliquer le théorème de Riesz dans les espaces de Hilbert (théorème 6.56), qui donne l'existence de  $Z \in \mathcal{L}^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$  t.q. :

$$T(V) = \int_{\Omega} ZVdP \text{ pour tout } V \in H. \tag{11.2}$$

Comme  $Z \in \mathcal{L}^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$ , la fonction  $Z$  est bien  $\mathcal{B}$ -mesurable et intégrable (elle est même de carré intégrable). On montre maintenant que  $Z$  vérifie la propriété (11.1) (et donc que  $Z \in E(X|\mathcal{B})$ ). Soit  $U$  une application  $\mathcal{B}$ -mesurable bornée de  $\Omega$  dans  $\mathbb{R}$ . On a  $U \in \mathcal{L}^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$ , on peut donc utiliser (11.2) avec pour  $V$  la classe de  $U$  et on obtient

$$E(XU) = T(V) = \int_{\Omega} ZV dP = E(ZU).$$

L'application  $Z$  vérifie donc bien la propriété (11.1), ce qui prouve que  $Z \in E(X|\mathcal{B})$ .

Plus précisément, un développement du raisonnement ci-avant (que les courageux peuvent faire) permet d'interpréter l'application  $X \mapsto E(X|\mathcal{B})$  comme l'opérateur de projection orthogonale de  $L^2(\Omega, \mathcal{A}, P)$  dans le sous-espace vectoriel fermé formé à partir de  $L^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$ .

**Existence si  $X$  est positive et intégrable.** On utilise ici le théorème de Radon-Nikodym (théorème 6.78, qui se démontre d'ailleurs avec le théorème de Riesz dans les espaces de Hilbert, théorème 6.56). On note toujours  $p_{\mathcal{B}}$  la restriction de  $P$  à  $\mathcal{B}$  (de sorte que  $P_{\mathcal{B}}$  est une probabilité sur  $\mathcal{B}$ ).

Pour  $B \in \mathcal{A}$ , on pose  $m(B) = \int_{\Omega} X1_B dP$ . On définit ainsi une mesure finie,  $m$ , sur  $\mathcal{A}$ , c'est la mesure de densité  $X$  par rapport à  $P$ . On note maintenant  $m_{\mathcal{B}}$  la restriction de cette mesure à la tribu  $\mathcal{B}$ . La mesure  $m_{\mathcal{B}}$  est absolument continue par rapport à la mesure  $P_{\mathcal{B}}$  (car  $B \in \mathcal{B}$ ,  $P_{\mathcal{B}}(B) = 0$  implique que  $P(B) = 0$  et donc  $X1_B = 0$  p.s. et donc  $m(B) = 0$  et donc  $m_{\mathcal{B}}(B) = 0$ ). Le théorème de Radon-Nikodym (théorème 6.78) donne alors l'existence de  $Z$ ,  $\mathcal{B}$ -mesurable positive, t.q.  $m_{\mathcal{B}} = ZP_{\mathcal{B}}$  (c'est-à-dire que  $m_{\mathcal{B}}$  est la mesure sur  $\mathcal{B}$  de densité  $Z$  par rapport à  $P_{\mathcal{B}}$ ).

La fonction  $Z$  est intégrable car

$$\int_{\Omega} Z dP = \int_{\Omega} Z dP_{\mathcal{B}} = m_{\mathcal{B}}(\Omega) = m(\Omega) = E(X) < +\infty \text{ (et } E(Z) = E(X)\text{)}.$$

Il reste à montrer que  $Z$  vérifie la propriété (11.1) (ce qui donne que  $Z \in E(X|\mathcal{B})$ ). On remarque tout d'abord que

$$E(Z1_B) = \int_{\Omega} Z1_B dP = m_{\mathcal{B}}(B) = \int_{\Omega} X1_B dP = E(X1_B) \text{ pour tout } B \in \mathcal{B}$$

. Par linéarité positive, on a donc, pour toute fonction  $\mathcal{B}$ -étagée positive,

$$E(ZU) = \int_{\Omega} ZU dP = \int_{\Omega} ZU dP_{\mathcal{B}} = \int_{\Omega} U dm = \int_{\Omega} UX dP = E(XU). \quad (11.3)$$

Par convergence monotone, on en déduit alors que (11.3) est encore vraie pour toute fonctions  $\mathcal{B}$ -mesurable positive. Enfin, en utilisant  $U = U^+ - U^-$  (et en remarquant que  $ZU$  et  $XU$  sont intégrables), on conclut que (11.3) est vraie pour toute fonction  $U$   $\mathcal{B}$ -mesurable bornée de  $\Omega$  dans  $\mathbb{R}$ . (On a ici repris un argument vu dans la remarque 6.75.) L'application  $Z$  vérifie donc la propriété (11.1), ce qui prouve que  $Z \in E(X|\mathcal{B})$ .

**Existence si  $X$  est seulement intégrable.** Comme les fonctions  $X^+$  et  $X^-$  sont positives et intégrables, il existe  $Z_1 \in E(X^+|\mathcal{B})$  et  $Z_2 \in E(X^-|\mathcal{B})$ . On pose  $Z = Z_1 - Z_2$ . L'application  $Z$  est  $\mathcal{B}$ -mesurable et intégrable (car  $Z_1$  et  $Z_2$  le sont) et, pour tout fonction  $U$   $\mathcal{B}$ -mesurable bornée, on a :

$$E(ZU) = E(Z_1 U) - E(Z_2 U) = E(X^+ U) - E(X^- U) = E(XU).$$

L'application  $Z$  vérifie donc la propriété (11.1), ce qui prouve que  $Z \in E(X|\mathcal{B})$ . ■

Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé et  $\mathcal{B}$  une tribu incluse dans  $\mathcal{A}$ . On a défini l'espérance, conditionnée par  $\mathcal{B}$ , d'une v.a.r. intégrable. On va maintenant montrer qu'on peut étendre la définition à des v.a.r. qui ne sont pas intégrables mais qui sont positives (la démonstration est déjà essentiellement dans la démonstration de la proposition 11.2). Pour cela, on va commencer par donner une p.s.-caractérisation de  $E(X|\mathcal{B})$  lorsque  $X$  est une v.a.r. positive et intégrable. Cette caractérisation n'utilisant pas l'intégrabilité de  $X$  on aura ainsi une définition de  $E(X|\mathcal{B})$  lorsque  $X$  est une v.a.r. positive. Ceci est fait dans la proposition 11.3 et la définition 11.4.

**Proposition 11.3 (Caractérisation de l'espérance conditionnelle d'une v.a.r. positive)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé et  $\mathcal{B}$  une tribu incluse dans  $\mathcal{A}$ .

1. Soit  $X$  une v.a.r. intégrable positive. Alors,  $Z \in E(X|\mathcal{B})$  si et seulement si  $Z$  est  $\mathcal{B}$ -mesurable, intégrable,  $\geq 0$  p.s. et t.q. :

$$E(ZU) = E(XU), \quad (11.4)$$

pour toute application  $U$  de  $\Omega$  dans  $\mathbb{R}$ ,  $\mathcal{B}$ -mesurable et positive.

2. Soit  $X$  une v.a.r. positive. On note  $\bar{E}(X|\mathcal{B})$  l'ensemble des applications  $\mathcal{B}$ -mesurables positives vérifiant la propriété (11.4). On a alors :

(a) (Existence)  $\bar{E}(X|\mathcal{B}) \neq \emptyset$ .

(b) (Unicité)  $Z_1, Z_2 \in \bar{E}(X|\mathcal{B}) \Rightarrow Z_1 = Z_2$  p.s..

DÉMONSTRATION – On commence par montrer le premier item. Si  $Z \in E(X|\mathcal{B})$ , la fonction  $Z$  est bien  $\mathcal{B}$ -mesurable intégrable et vérifie la propriété (11.1). Elle vérifie donc la propriété (11.4) en ajoutant l'hypothèse  $U$  bornée. Pour montrer que  $Z \geq 0$  p.s., on prend  $U = 1_B$  avec  $B = \{Z < 0\}$  ( $U$  est bien  $\mathcal{B}$ -mesurable bornée). On obtient  $E(ZU) = E(XU) \geq 0$ . Comme  $ZU \leq 0$ , on a donc  $ZU = 0$  p.s. et donc  $Z \geq 0$  p.s.. Enfin, pour montrer que  $Z$  vérifie la propriété (11.4) (c'est-à-dire avec  $U$   $\mathcal{B}$ -mesurable positive mais non nécessairement bornée), il suffit d'utiliser le théorème de convergence monotone (théorème 4.17) en introduisant  $U_n = U1_{B_n}$  avec, pour  $n \in \mathbb{N}$ ,  $B_n = \{U \leq n\}$ .

Réciproquement, si  $Z$  est  $\mathcal{B}$ -mesurable, intégrable,  $\geq 0$  p.s. et vérifie la propriété (11.4) ; il est facile de voir que  $Z$  vérifie (11.1). En effet, si  $U$  est  $\mathcal{B}$ -mesurable bornée, on utilise (11.4) avec les parties positive et négative de  $U$  pour obtenir (11.1). Donc,  $Z \in E(X|\mathcal{B})$ .

On montre maintenant le deuxième item de la proposition.

#### Existence

On reprend la démonstration de la proposition 11.2. On rappelle que  $p_B$  est la restriction de  $P$  à  $\mathcal{B}$ ,  $m = XP$  (c'est-à-dire la mesure de densité  $X$  par rapport à  $P$ ) et  $m_B$  est la restriction de  $m$  à  $\mathcal{B}$ . La mesure  $m_B$  est absolument continue par rapport à la mesure  $P_B$ . La mesure  $m$  n'est pas finie si  $X$  n'est pas intégrable. On ne peut donc pas appliquer directement le théorème 6.78 (qui demande que  $m_B$  soit finie). Pour résoudre cette petite difficulté, on pose, pour  $n \in \mathbb{N}$   $m_n = X1_{\{n \leq X < n+1\}}$  (qui est une mesure sur  $\mathcal{A}$ ) et  $m_{n,B}$  sa restriction  $\mathcal{B}$ . La mesure  $m_{n,B}$  est absolument continue par rapport à la mesure  $P_B$  et est finie. Le théorème de Radon-Nikodym donne alors l'existence de  $Z_n$ ,  $\mathcal{B}$ -mesurable positive, t.q.  $m_{n,B} = Z_n P_B$ . On pose alors  $Z = \sum_{n \in \mathbb{N}} Z_n$ . La fonction  $Z$  est  $\mathcal{B}$ -mesurable positive, il reste à montrer que  $Z$  vérifie la propriété (11.4) (ce qui donnera que  $Z \in \bar{E}(X|\mathcal{B})$ ). Soit  $U$  une application  $\mathcal{B}$ -mesurable positive de  $\Omega$  dans  $\mathbb{R}$ . Pour tout  $n \in \mathbb{N}$ , on a

$$E(Z_n U) = \int_{\Omega} Z_n U dP = \int_{\Omega} Z_n U dP_B = \int_{\Omega} U d m_{n,B} = \int_{\Omega} U X 1_{\{n \leq X < n+1\}} dP.$$

En sommant cette dernière égalité pour  $n \in \mathbb{N}$ , on obtient (par le corollaire 4.18 sur les séries à termes positifs)

$$E(ZU) = E(XU).$$

L'application  $Z$  vérifie donc la propriété (11.4), ce qui prouve que  $Z \in \bar{E}(X|\mathcal{B})$ .

#### Unicité

Soit  $Z_1, Z_2 \in \bar{E}(X|\mathcal{B})$ . prenons  $U = (\text{sign}(Z_1 - Z_2))^+$  (qui est bien  $\mathcal{B}$ -mesurable et positive). On a donc, par la propriété (11.4),

$$E(Z_1 U) = E(Z_2 U) = E(XU),$$

mais on ne peut rien en déduire car il est possible que  $E(XU) = +\infty$ . On va donc modifier légèrement  $U$ . Pour  $n \in \mathbb{N}^*$ , on pose

$$B_n = \{Z_1 \leq n\} \cap \{Z_2 \leq n\} \text{ et } U_n = U1_{B_n}.$$

L'application  $U_n$  est encore  $\mathcal{B}$ -mesurable et positive ; comme  $Z_1, Z_2 \in \bar{E}(X|\mathcal{B})$ , la propriété (11.4) donne  $E(Z_1 U_n) = E(Z_2 U_n)$ . On a donc

$$0 \leq E(Z_1 U_n) = E(Z_2 U_n) \leq n,$$

d'où l'on déduit que  $E((Z_1 - Z_2)U_n) = 0$ . Mais,  $(Z_1 - Z_2)U_n \geq 0$ . En faisant tendre  $n$  vers l'infini, le théorème de convergence monotone (théorème 4.16) donne  $E((Z_1 - Z_2)U) = 0$ , c'est-à-dire  $E((Z_1 - Z_2)^+) = 0$  et donc  $Z_1 \leq Z_2$  p.s.. En changeant les rôles de  $Z_1$  et  $Z_2$  on a aussi  $Z_2 \leq Z_1$  p.s.. D'où  $Z_1 = Z_2$  p.s.. ■

**Définition 11.4 (Espérance conditionnelle d'une v.a.r. positive)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $X$  une v.a.r. positive et  $\mathcal{B}$  une tribu incluse dans  $\mathcal{A}$ . On appelle espérance conditionnée par  $\mathcal{B}$  de  $X$  ou espérance conditionnelle de  $X$  par rapport à  $\mathcal{B}$ , l'ensemble des applications  $Z$  de  $\Omega$  dans  $\mathbb{R}$ ,  $\mathcal{B}$ -mesurables, positives et t.q. :

$$E(ZU) = E(XU), \forall U : \Omega \rightarrow \mathbb{R}, \mathcal{B}\text{-mesurable et positive.} \quad (11.5)$$

On note  $\bar{E}(X|\mathcal{B})$  cette espérance conditionnelle (c'est donc un ensemble de fonctions). (Noter que dans (11.5), les applications  $ZU$  et  $XU$  sont bien des v.a.r. positives, leur intégrale sur  $\Omega$  est donc bien définie et appartient à  $\bar{\mathbb{R}}_+$ ).

La proposition 11.3 nous donne l'existence et l'unicité (p.s.) de l'espérance conditionnelle lorsque  $X$  est une v.a.r. positive. Sous les hypothèses de la définition 11.4, si  $X$  est de plus intégrable, on a donc deux définitions de l'espérance conditionnelle de  $X$  par rapport à  $\mathcal{B}$ , notée  $E(X|\mathcal{B})$  et  $\bar{E}(X|\mathcal{B})$ . La proposition 11.3 montre que  $Z_1 \in E(X|\mathcal{B})$  et  $Z_2 \in \bar{E}(X|\mathcal{B})$  implique  $Z_1 = Z_2$  p.s.. En pratique, comme on confond  $E(X|\mathcal{B})$  avec l'un de ces éléments et  $\bar{E}(X|\mathcal{B})$  avec l'un de ces éléments, on a donc  $E(X|\mathcal{B}) = \bar{E}(X|\mathcal{B})$  p.s.. Il est donc inutile de conserver la notation  $\bar{E}(X|\mathcal{B})$  et on conservera la notation  $E(X|\mathcal{B})$  dans les deux cas, c'est-à-dire "X v.a.r. intégrable" et "X v.a.r. positive".

Nous donnons maintenant quelques propriétés de l'espérance conditionnelle.

**Remarque 11.5 (Linéarité de l'espérance conditionnelle)** Une première propriété de l'espérance conditionnelle est sa linéarité. Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $\mathcal{B}$  une sous tribu de  $\mathcal{A}$  et  $X_1, X_2$  deux v.a.r. intégrables. On pose  $Z_1 = E(X_1|\mathcal{B})$  et  $Z_2 = E(X_2|\mathcal{B})$  (plus précisément,  $Z_1$  et  $Z_2$  sont des représentants de  $E(X_1|\mathcal{B})$  et  $E(X_2|\mathcal{B})$ ). Soit  $U$  une v.a.r.  $\mathcal{B}$ -mesurable bornée, on a

$$E(Z_1 U) = E(X_1 U) \text{ et } E(Z_2 U) = E(X_2 U).$$

Par linéarité de l'espérance, on a donc  $E((Z_1 + Z_2)U) = E((X_1 + X_2)U)$ . Ceci prouve que  $Z_1 + Z_2 = E(X_1 + X_2|\mathcal{B})$  p.s. et donc que

$$E(X_1 + X_2|\mathcal{B}) = E(X_1|\mathcal{B}) + E(X_2|\mathcal{B}) \text{ p.s..}$$

**Proposition 11.6** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $\mathcal{B}$  une tribu incluse dans  $\mathcal{A}$  et  $X$  une v.a.r. Soit  $p \in ]1, \infty]$  et  $q$  le nombre conjugué de  $p$  (i.e.  $q = p/(p-1)$  si  $p < +\infty$  et  $q = 1$  si  $p = \infty$ ). On suppose que  $X \in \mathcal{L}_{\mathbb{R}}^p(\Omega, \mathcal{A}, P)$ . Soit  $Z = E(X|\mathcal{B})$  p.s.. Alors,  $Z \in \mathcal{L}_{\mathbb{R}}^p(\Omega, \mathcal{A}, P)$  et  $E(ZU) = E(XU)$  pour toute application  $U$  (de  $\Omega$  dans  $\mathbb{R}$ )  $\mathcal{B}$ -mesurable telle que  $|U|^q$  soit intégrable.

DÉMONSTRATION – La démonstration fait partie de l'exercice 11.6. En fait, le cas  $p = 2$  a déjà été vu dans la démonstration de la proposition 11.2. ■

**Proposition 11.7 (Inégalité de Jensen généralisée)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $\mathcal{B}$  une tribu incluse dans  $\mathcal{A}$  et  $X$  une v.a.r. de carré intégrable. Soit  $\varphi$  une fonction convexe de  $\mathbb{R}$  dans  $\mathbb{R}$ . On suppose que  $\varphi(X)$  est intégrable. On a alors

$$E(\varphi(X)|\mathcal{B}) \geq \varphi(E(X|\mathcal{B})) \text{ p.s..}$$

DÉMONSTRATION – D'après le lemme 11.8 donné ci-après, comme  $\varphi$  est convexe, il existe  $c$ , fonction croissante de  $\mathbb{R}$  dans  $\mathbb{R}$  (et donc fonction borélienne de  $\mathbb{R}$  dans  $\mathbb{R}$ ) t.q., pour tout  $x, a \in \mathbb{R}$ ,  $\varphi(x) - \varphi(a) \geq c(a)(x - a)$ .

Soit  $Z = E(X|\mathcal{B})$  p.s. (plus précisément,  $Z$  est un représentant de  $E(X|\mathcal{B})$ ). On a donc pour tout  $\omega \in \Omega$  :

$$\varphi(X(\omega)) - \varphi(Z(\omega)) \geq c(Z(\omega))(X(\omega) - Z(\omega)). \quad (11.6)$$

On aimerait intégrer cette inégalité sur un élément (bien choisi) de  $\mathcal{B}$  mais cela n'est pas possible car les v.a.r.  $\varphi(Z)$  et  $c(Z)(X - Z)$  peuvent ne pas être intégrables (bien que  $Z$  et  $X$  soient intégrables). Pour  $p \in \mathbb{N}^*$ , on introduit donc  $A_p = \{|Z| \leq p\}$  de sorte que les v.a.r.  $1_{A_p} c(Z)(X - Z)$  et  $1_{A_p} \varphi(Z)$  sont intégrables (noter que  $c(Z)$  est bornée sur  $A_p$  car  $c$  est croissante). On pose aussi

$$A = \{E(\varphi(X)|\mathcal{B}) - \varphi(Z) < 0\} \text{ et } B_p = A_p \cap A.$$

Soit  $p \in \mathbb{N}^*$ , l'inégalité (11.6) donne  $1_{B_p} (\varphi(X) - \varphi(Z)) \geq 1_{B_p} c(Z)(X - Z)$  et donc, en intégrant sur  $\Omega$  :

$$\int_{B_p} (\varphi(X) - \varphi(Z)) dP \geq \int_{B_p} c(Z)(X - Z) dP. \quad (11.7)$$

Comme  $Z$  et  $E(\varphi(X)|\mathcal{B})$  sont  $\mathcal{B}$ -mesurables, on a  $B_p \in \mathcal{B}$  (et donc  $1_{B_p}$  est  $\mathcal{B}$ -mesurable). On a aussi  $c(Z)$   $\mathcal{B}$ -mesurable (car  $c$  est borélienne) et donc  $1_{B_p} c(Z)$   $\mathcal{B}$ -mesurable. On en déduit :

$$\int_{B_p} c(Z)(X - Z) dP = E(1_{B_p} c(Z)(X - Z)) = 0 \text{ (car } Z \in E(X|\mathcal{B})\text{),}$$

et

$$\int_{B_p} (\varphi(X) - \varphi(Z)) dP = E(1_{B_p} (\varphi(X) - \varphi(Z))) = E(1_{B_p} (E(\varphi(X)|\mathcal{B}) - \varphi(Z))).$$

Avec (11.7), on en déduit :

$$\int_{B_p} (E(\varphi(X)|\mathcal{B}) - \varphi(Z)) dP \geq 0.$$

Comme  $E(\varphi(X)|\mathcal{B}) - \varphi(Z) < 0$  sur  $B_p$  (car  $B_p \subset A$ ), on a donc  $P(B_p) = 0$  et donc  $P(A) = P(\bigcup_{p \in \mathbb{N}^*} B_p) = 0$ , ce qui donne bien  $E(\varphi(X)|\mathcal{B}) \geq \varphi(Z)$  p.s.. ■

Voici maintenant le lemme utilisé dans la démonstration précédente.

**Lemme 11.8** *Soit  $\varphi$  une fonction convexe de  $\mathbb{R}$  dans  $\mathbb{R}$ , il existe alors  $c$ , fonction croissante de  $\mathbb{R}$  dans  $\mathbb{R}$  (et donc fonction borélienne de  $\mathbb{R}$  dans  $\mathbb{R}$ ) t.q., pour tout  $x, a \in \mathbb{R}$ ,  $\varphi(x) - \varphi(a) \geq c(a)(x - a)$ .*

DÉMONSTRATION – Si  $\varphi$  est dérivable sur  $\mathbb{R}$ , la fonction  $c$  existe et est unique, elle est donnée par  $c = \varphi'$ . L'existence de  $c$  est légèrement plus difficile si  $\varphi$  n'est pas dérivable sur tout  $\mathbb{R}$  (et on perd l'unicité de  $c$ ).

Soit  $a \in \mathbb{R}$ , on considère le fonction  $h_a : x \mapsto \frac{\varphi(x) - \varphi(a)}{x - a}$  qui est définie sur  $\mathbb{R} \setminus \{a\}$ . La convexité de  $\varphi$  permet de montrer que  $h_a$  est croissante (c'est-à-dire que  $x, y \in \mathbb{R} \setminus \{a\}$ ,  $x > y \Rightarrow h_a(x) \geq h_a(y)$ ). La fonction  $h_a$  a donc une limite à gauche (et à droite) en tout point, y compris au point  $a$ . On pose (par exemple) :

$$c(a) = \lim_{x \rightarrow a, x < a} h_a(x).$$

Il est facile de vérifier que la fonction  $c$  ainsi définie est croissante de  $\mathbb{R}$  dans  $\mathbb{R}$  et vérifie, pour tout  $x, a \in \mathbb{R}$ ,  $\varphi(x) - \varphi(a) \geq c(a)(x - a)$ . ■

On définit maintenant l'espérance conditionnelle par rapport à une v.a.r. ou un v.a.

**Définition 11.9 (Espérance conditionnelle à une v.a.r)** *Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé et  $X$  une v.a.r. (ou un v.a. de dimension  $d$ ,  $d \geq 1$ ). Soit  $Y$  une v.a.r. intégrable ou une v.a.r. positive. On appelle espérance conditionnelle par  $X$  de  $Y$  ou espérance conditionnelle de  $Y$  par rapport à  $X$  (ou espérance conditionnelle de  $Y$  sachant  $X$ ) l'ensemble  $E(Y|\sigma(X))$ , où  $\sigma(X)$  est la tribu engendrée par  $X$ . On note  $E(Y|X)$  cette espérance conditionnelle, de sorte que  $E(Y|X) = E(Y|\sigma(X))$ . (L'ensemble  $E(Y|X)$  est donc un ensemble de v.a.r. et, comme d'habitude, on confond  $E(Y|X)$  avec l'un de ces éléments.)*

Pour caractériser  $E(Y|X)$  (sous les hypothèses de la définition 11.9) et pour calculer cette espérance conditionnelle, on utilise, en général, le théorème 3.31 que nous rappelons sous une forme légèrement plus précise (donnée dans la démonstration du théorème 3.31).

**Théorème 11.10 (Mesurabilité d'une v.a.r. par rapport à une autre)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé et  $X, Y$  deux v.a.r. On note  $\sigma(X)$  la tribu engendrée par  $X$ . Alors :

- La v.a.r.  $Y$  est  $\sigma(X)$ -mesurable si et seulement s'il existe  $f$ , fonction borélienne de  $\mathbb{R}$  dans  $\mathbb{R}$ , t.q.  $Y = f(X)$ .
- La v.a.r.  $Y$  est  $\sigma(X)$ -mesurable bornée si et seulement s'il existe  $f$ , fonction borélienne bornée de  $\mathbb{R}$  dans  $\mathbb{R}$ , t.q.  $Y = f(X)$ .
- La v.a.r.  $Y$  est  $\sigma(X)$ -mesurable positive si et seulement s'il existe  $f$ , fonction borélienne positive de  $\mathbb{R}$  dans  $\mathbb{R}$ , t.q.  $Y = f(X)$ .

DÉMONSTRATION – La démonstration de ce théorème est donnée dans la démonstration du théorème 3.31. ■

Voici une conséquence immédiate de ce théorème, utilisée pour calculer  $E(Y|X)$

**Proposition 11.11 (Calcul de  $E(Y|X)$ )** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé et  $X, Y$  deux v.a.r. Soit  $Z$  une application de  $\Omega$  dans  $\mathbb{R}$ .

1. On suppose que  $Y$  est intégrable. Alors,  $Z \in E(Y|X)$  si et seulement s'il existe  $\psi$  application borélienne de  $\mathbb{R}$  dans  $\mathbb{R}$  t.q.  $Z = \psi(X)$ ,  $\psi(X)$  est intégrable et

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)), \quad (11.8)$$

pour toute application  $\varphi$  de  $\mathbb{R}$  dans  $\mathbb{R}$ , borélienne bornée.

2. On suppose que  $Y$  est positive. Alors,  $Z \in E(Y|X)$  si et seulement s'il existe  $\psi$  application borélienne positive de  $\mathbb{R}$  dans  $\mathbb{R}$  t.q.  $Z = \psi(X)$  et

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)), \quad (11.9)$$

pour toute application  $\varphi$  de  $\mathbb{R}$  dans  $\mathbb{R}$ , borélienne positive.

DÉMONSTRATION –

La démonstration est une conséquence facile du théorème 11.10. ■

La conséquence de la proposition 11.11 est que (sous les hypothèses de la proposition) l'on cherche  $E(Y|X)$  sous la forme d'une fonction  $\psi(X)$  (avec  $\psi$  de  $\mathbb{R}$  dans  $\mathbb{R}$ ) vérifiant (11.8) (ou (11.9)). On raisonne, en général, par condition nécessaire sur  $\psi$  et, comme on sait que  $E(X|Y)$  existe, il est même inutile de vérifier que la fonction  $\psi(X)$  que l'on trouve (qui est, en général, définie p.s.) est bien intégrable (ou positive).

La proposition 11.11 montre également que (sous les hypothèses de la proposition 11.11) la fonction  $Y$  est une fonction de  $X$  (c'est-à-dire  $Y = \psi(X)$  pour un certain  $\psi$  de  $\mathbb{R}$  dans  $\mathbb{R}$ ) si et seulement si  $E(Y|X) = Y$ . Pour montrer que la v.a.r.  $Y$  est une fonction d'une autre v.a.r.  $X$ , il suffit donc de montrer que  $E(Y|X) = Y$ . En comparaison, le calcul de la covariance entre  $X$  et  $Y$  (après normalisation) s'intéresse seulement à l'existence ou non d'une dépendance affine de  $Y$  en fonction de  $X$ . Voir, à ce propos, l'exercice 11.14.

**Remarque 11.12 (Deux propriétés de l'espérance conditionnelle)** Voici deux propriétés qui nous seront utiles dans la section suivante. Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $\mathcal{B}$  une sous tribu de  $\mathcal{A}$  et  $X$  une v.a.r. intégrable.

1. Soit  $V$  une v.a.r.  $\mathcal{B}$ -mesurable bornée. On a alors  $E(XV|\mathcal{B}) = VE(X|\mathcal{B})$  p.s.. (Voir l'exercice 11.4.)
2. On suppose que  $\sigma(X)$  et  $\mathcal{B}$  sont des tribus indépendantes. On a alors  $E(X|\mathcal{B}) = E(X)$  p.s.. En particulier, si  $Y$  est une v.a.r. indépendante de  $X$  (c'est-à-dire que  $\sigma(X)$  et  $\sigma(Y)$  sont des tribus indépendantes), on a alors  $E(X|Y) = E(X)$  p.s.. (Voir l'exercice 11.5.)

## 11.2 Martingales

**Définition 11.13 (Filtration et processus)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé

1. On appelle filtration une suite de tribus  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  t.q.  $\mathcal{B}_n \subset \mathcal{B}_{n+1} \subset \mathcal{A}$ , pour tout  $n \in \mathbb{N}$ .
2. On appelle processus réel une suite de v.a.r.
3. Soit  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  une filtration et  $(X_n)_{n \in \mathbb{N}}$  un processus réel. On dit que  $(X_n)_{n \in \mathbb{N}}$  est adapté à la filtration  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  si, pour tout  $n \in \mathbb{N}$ ,  $X_n$  est  $\mathcal{B}_n$ -mesurable.

**Définition 11.14 (Martingale)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  une filtration et  $(X_n)_{n \in \mathbb{N}}$  un processus réel (c'est-à-dire une suite de v.a.r.).

**(Martingale)** La suite  $(X_n)_{n \in \mathbb{N}}$  est une martingale par rapport à la filtration  $(\mathcal{B}_n)_n$  si on a, pour tout  $n \in \mathbb{N}$ ,

1.  $X_n$  est  $\mathcal{B}_n$ -mesurable et intégrable,
2.  $E(X_{n+1}|\mathcal{B}_n) = X_n$  p.s..

**(Sous et sur martingale)** La suite  $(X_n)_{n \in \mathbb{N}}$  est une sous-martingale [resp. sur-martingale] par rapport à la filtration  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  si on a, pour tout  $n \in \mathbb{N}$ ,

1.  $X_n$  est  $\mathcal{B}_n$ -mesurable et intégrable,
2.  $E(X_{n+1}|\mathcal{B}_n) \geq X_n$  p.s. [resp.  $E(X_{n+1}|\mathcal{B}_n) \leq X_n$  p.s.].

**Remarque 11.15** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  une filtration et  $(X_n)_{n \in \mathbb{N}}$  une martingale par rapport à la filtration  $(\mathcal{B}_n)_n$ . Pour tout  $n \in \mathbb{N}$ , on a, comme  $1_\Omega$  est  $\mathcal{B}_n$ -intégrable bornée, et  $E(X_{n+1}|\mathcal{B}_n) = X_n$  p.s.,

$$E(X_{n+1}) = \int_{\Omega} X_{n+1} 1_\Omega dP = \int_{\Omega} E(X_{n+1}|\mathcal{B}_n) 1_\Omega dP = \int_{\Omega} X_n 1_\Omega dP = E(X_n).$$

On a donc  $E(X_n) = E(X_0)$  pour tout  $n \in \mathbb{N}$ .

**Exemple 11.16 (Exemples de Martingales)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé.

1. Soit  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  une filtration et  $X$  une v.a.r. intégrable. On pose  $X_n = E(X|\mathcal{B}_n)$ . La suite  $(X_n)_{n \in \mathbb{N}}$  est une martingale (par rapport à la filtration  $(\mathcal{B}_n)_n$ ). (Voir l'exercice 11.24.)
2. Soit  $X_0$  une v.a.r. intégrable et  $(J_n)_{n \in \mathbb{N}^*}$  une suite de v.a.r. intégrable et de moyenne nulle. On suppose que la suite formée de  $X_0$  et  $(J_n)_{n \in \mathbb{N}^*}$  est une suite de v.a.r. indépendantes. On pose alors, pour  $n \in \mathbb{N}$ ,  $X_{n+1} = X_n + J_{n+1}$  et  $\mathcal{B}_n$  la tribu engendrée par  $X_0, \dots, X_n$ . La suite  $(X_n)_{n \in \mathbb{N}}$  est une martingale (par rapport à la filtration  $(\mathcal{B}_n)_n$ ). (Voir l'exercice 11.25.)



**Définition 11.17 (Temps d'arrêt)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  une filtration et  $T$  une v.a. à valeurs dans  $\mathbb{N} \cup \{+\infty\}$  (c'est-à-dire une application mesurable de  $\Omega$ , muni de la tribu  $\mathcal{A}$ , dans  $\mathbb{N} \cup \{+\infty\}$ , muni de la tribu formée de l'ensemble des parties de  $\mathbb{N} \cup \{+\infty\}$ ). L'application  $T$  s'appelle un temps d'arrêt si, pour tout  $n \in \mathbb{N}$ ,  $\{T = n\} \in \mathcal{B}_n$ . Si  $T$  est un temps d'arrêt, on note  $\mathcal{B}_T$  la tribu définie par  $\mathcal{B}_T = \{A \in \mathcal{B}_\infty \text{ t.q., pour tout } n \in \mathbb{N}, A \cap \{T = n\} \in \mathcal{B}_n\}$  où  $\mathcal{B}_\infty$  est la plus petite tribu contenant toutes les tribus  $\mathcal{B}_n$  (c'est-à-dire la tribu engendrée par les  $\mathcal{B}_n$ ).

Le théorème suivant montre qu'une martingale arrêtée est encore une martingale.

**Théorème 11.18 (Martingale arrêtée à un temps d'arrêt)** Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé muni d'une filtration  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  et  $(X_n)_{n \in \mathbb{N}}$  une martingale (par rapport à la filtration  $(\mathcal{B}_n)_{n \in \mathbb{N}}$ ). Soit  $\nu$  un temps d'arrêt. Pour  $n \in \mathbb{N}$ , on pose  $Y_n = X_{\nu \wedge n}$  (On rappelle que  $\nu \wedge n(\omega) = \min\{\nu(\omega), n\}$  et donc que  $X_{\nu \wedge n}(\omega) = X_{\min\{\nu(\omega), n\}}(\omega)$ , pour tout  $\omega \in \Omega$ .) Alors, la suite  $(Y_n)_{n \in \mathbb{N}}$  est encore une martingale (par rapport à la filtration  $(\mathcal{B}_n)_{n \in \mathbb{N}}$ ).

DÉMONSTRATION – Soit  $n \in \mathbb{N}$ , on a  $Y_n = X_n 1_{\{T > n\}} + \sum_{k=0}^n X_k 1_{\{T=k\}}$ . On en déduit tout d'abord que  $Y_n$  est intégrable (comme somme finie de fonctions intégrables). Puis, on montre que  $Y_n$  est  $\mathcal{B}_n$ -mesurable. Pour cela, on remarque que  $\{T = k\} \in \mathcal{B}_k \subset \mathcal{B}_n$ , pour  $k \leq n$ , et que

$$\{T > n\} = \left( \bigcup_{k=0}^n \{T = k\} \right)^c \in \mathcal{B}_n.$$

Enfin, on remarque que  $X_k$  est  $\mathcal{B}_n$ -mesurable pour tout  $k \leq n$ . Grâce à la stabilité des fonctions mesurables par somme et produit, on obtient bien, finalement, que  $Y_n$  est  $\mathcal{B}_n$ -mesurable.

Il reste maintenant à montrer que  $E(Y_{n+1} | \mathcal{B}_n) = Y_n$  p.s., pour tout  $n \in \mathbb{N}$ . Soit  $n \in \mathbb{N}$ , on a

$$Y_{n+1} = X_{n+1} 1_{\{T > n+1\}} + \sum_{k=0}^{n+1} X_k 1_{\{T=k\}} = X_{n+1} 1_{\{T \geq n+1\}} + \sum_{k=0}^n X_k 1_{\{T=k\}}.$$

Par linéarité de l'espérance conditionnelle, on a donc

$$E(Y_{n+1} | \mathcal{B}_n) = E(X_{n+1} 1_{\{T \geq n+1\}} | \mathcal{B}_n) + \sum_{k=0}^n E(X_k 1_{\{T=k\}} | \mathcal{B}_n).$$

Comme  $\{T \geq n+1\} = \left( \bigcup_{k=0}^n \{T = k\} \right)^c \in \mathcal{B}_n$ , la remarque 11.12 (et le fait que  $E(X_{n+1} | \mathcal{B}_n) = X_n$  p.s.) donne

$$E(X_{n+1} 1_{\{T \geq n+1\}} | \mathcal{B}_n) = 1_{\{T \geq n+1\}} E(X_{n+1} | \mathcal{B}_n) = X_n 1_{\{T \geq n+1\}}.$$

Puis comme, pour  $k \in \{0, \dots, n\}$ ,  $X_k 1_{\{T=k\}}$  est  $\mathcal{B}_n$ -mesurable, on a  $E(X_k 1_{\{T=k\}} | \mathcal{B}_n) = X_k 1_{\{T=k\}}$ . On obtient ainsi

$$E(Y_{n+1} | \mathcal{B}_n) = X_n 1_{\{T \geq n+1\}} + \sum_{k=0}^n X_k 1_{\{T=k\}} = Y_n \text{ p.s..}$$

Ce qui termine la démonstration. ■

On conclut cette section par un théorème, sans démonstration, sur la convergence des martingales.

**Théorème 11.19 (Convergence p.s. d'une martingale)**

Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé,  $(\mathcal{B}_n)_{n \in \mathbb{N}}$  une filtration et  $(X_n)_{n \in \mathbb{N}}$  une suite de v.a.r. On suppose que  $(X_n)_{n \in \mathbb{N}}$  est une martingale par rapport à  $(\mathcal{B}_n)_{n \in \mathbb{N}}$ .

1. On suppose que la suite  $(X_n)_{n \in \mathbb{N}}$  est bornée dans  $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, \mathbb{P})$ . Alors il existe une v.a.r. intégrable,  $X$ , t.q.  $X_n \rightarrow X$  p.s., quand  $n \rightarrow +\infty$ .
2. On suppose  $X_n \geq 0$  pour tout  $n \in \mathbb{N}$ . Alors, il existe une v.a.r. intégrable,  $X$ , t.q.  $X_n \rightarrow X$  p.s., quand  $n \rightarrow +\infty$ .

On peut noter que le deuxième item du théorème 11.19 est une conséquence du premier car, pour une martingale, on a toujours  $E(X_n) = E(X_0)$  pour tout  $n \in \mathbb{N}$  (et les  $X_n$  sont toujours intégrales). Si  $X_n \geq 0$ , on a donc  $\|X_n\|_1 = E(X_0) < +\infty$ .