[M. Gubinelli - Processus discrets - M1 MMD 2010/2011 - 20100921 - v.1]

I Espérance Conditionnelle

1 Pré-requis du cours

On donnera pour acquises certaines notions de probabilités et de théorie de l'intégration niveau L3. On rappelle les objets de base des probabilités. Une espace probabilisé est un triplet (Ω, $\mathcal{F}$, $\mathbb{P}$) où Ω est un ensemble, $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ est une tribu (classe de parties de Ω stable par passage au complémentaire et réunion dénombrable) et $\mathbb{P}: \mathcal{F} \to [0,1]$ est une fonction qui satisfait:

1. $\mathbb{P}(\emptyset) = 0$; $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ (et donc $\mathbb{P}(\Omega) = 1$);
2. Si $\{A_n : n \geqslant 1\} \subseteq \mathcal{F}$ est une famille dénombrable tels que $A_i \cap A_j = \emptyset$ pour $i \neq j$, $\mathbb{P}(\cup_{n \geqslant 1} A_n) = \sum_{n \geqslant 1} \mathbb{P}(A_n)$;

On appelle $\mathbb{P}$ une probabilité (mesure positive de masse 1). Les éléments $\omega \in \Omega$ représentent l'issue d'une expérience aléatoire. Les éléments $A \in \mathcal{F}$ sont les événements auxquels on associe une probabilité $\mathbb{P}(A)$. Pour tout $\mathcal{C} \subseteq \mathcal{P}(\Omega)$ la tribu $\sigma(\mathcal{C})$ est la plus petite sous-tribu $\mathcal{G} \subseteq \mathcal{F}$ tel que $\mathcal{C} \subseteq \mathcal{G}$ ou $\sigma(\mathcal{C}) = \cap_{\mathcal{C} \subset \mathcal{G}, \mathcal{G}\,\text{tribu}}\, \mathcal{G}$. On appelle $\sigma(\mathcal{C})$ la tribu engendrée par $\mathcal{C}$. Si $A_1, A_2, \ldots$ sont dans $\mathcal{F}$ alors $\sigma(A_1, A_2, \ldots) = \sigma(\{A_1, A_2, \ldots\})$. Exemple: $\sigma(A) = \{\emptyset, \Omega, A, A^c\}$. Une fonction F: (Ω, $\mathcal{F}$) $\to$ (Ξ, $\mathcal{X}$) est mesurable si pour tout $B \in \mathcal{X}$ on a $F^{-1}(B) \in \mathcal{F}$, autrement dit, si l'ensemble $\{X \in B\} = \{\omega \in \Omega: X(\omega) \in B\}$ est dans $\mathcal{F}$. Pour tout tribu $\mathcal{G}$ on note l'ensemble des fonction $\mathcal{G}$-mesurables par $L^0(\mathcal{G})$, si $f \in L^0(\mathcal{G})$ on écrit aussi $f \hat{\in} \mathcal{G}$. L'espérance $\mathbb{E}[X]$ d'une fonction $\mathcal{F}$-mesurable réelle est l'intégrale de X par rapport à la mesure $\mathbb{P}$: $\mathbb{E}[X] = \int_\Omega X(\omega)\mathbb{P}(\mathrm{d}\omega)$ (il est bien défini si $X \geqslant 0$ où si $\mathbb{E}[|X|] < \infty$). On dit d'une famille $\{\mathcal{A}_i\}_{i \in I}$ des sous-tribus de $\mathcal{F}$ que sont indépendantes ssi $\mathbb{P}(\cap_{i \in J} A_i) = \prod_{i \in J} \mathbb{P}(A_i)$ pour tout choix de $J \subseteq I$ et $A_i \in \mathcal{A}_i$, $i \in J$.

2 L'espérance conditionnelle

Lorsqu'on travaille avec des variables aléatoires (v.a.s) discrètes on introduit la notion de probabilité conditionnelle par la formule:

$$\mathbb{P}(Y = y | X = x) := \frac{\mathbb{P}(Y = y, X = x)}{\mathbb{P}(X = x)}, \text{ si } \mathbb{P}(X = x) > 0, \tag{1}$$

d'où la définition d'espérance conditionnelle de $f(Y)$ sachant que $X = x$ par

$$\sum_y f(y)\mathbb{P}(Y = y | X = x) = u_f(x),$$

pour toute fonction f à valeurs réelles ou complexes mesurable bornée. Souvent on utilise la notation $\mathbb{E}(f(Y)|X = x) := u_f(x)$. Cette définition pose des problèmes lorsque on travaille avec des v.a.s continues où avec des modèles aléatoires plus complexes (suites infinies des v.a.s, espaces des fonctions, etc...). Par exemple, si X est continue, alors $\mathbb{P}(X = x) = 0$ pour tout $x \in \mathbb{R}$, donc la définition ci dessous est inutilisable. La façon plus efficace de contourner cette difficulté est de modifier la définition de probabilité conditionnelle, en effet, il sera plus facile de définir d'abord l'espérance conditionnelle et après la probabilité conditionnelle comme un sous-produit.

Pour motiver la nouvelle définition on observe la propriété suivante de la v.a. $u_f(X)$: pour toute fonction mesurable bornée h:

$$\mathbb{E}[u_f(X)h(X)] = \sum_{x: \mathbb{P}(X = x) > 0} h(x)u_f(x)\mathbb{P}(X = x) = \sum_{x,y} h(x)f(y)\mathbb{P}(Y = y, X = x), \tag{2}$$

ce qui revient à dire que, pour tout h:

$$\mathbb{E}[h(X)u_f(X)] = \mathbb{E}[h(X)f(Y)]. \tag{3}$$

De plus si g est une fonction telle que $\mathbb{E}[h(X)g(X)] = \mathbb{E}[h(X)f(Y)]$ pour tout h mesurable bornée, alors on a que $g(x) = u_f(x)$ pour tout x tel que $\mathbb{P}(X = x) > 0$. En effet, si $r(x) = g(x) - u_f(x)$ on doit avoir que $\mathbb{E}[h(X)r(X)] = 0$ et donc en choisissant $h(x) = \text{sign } r(x)$ on a que $\mathbb{E}[h(X)r(X)] = \mathbb{E}[|r(X)|] = \sum_{x:\mathbb{P}(X=x)>0} |r(x)|\mathbb{P}(X = x) = 0$ ce qu'implique que $r(x) = 0$ pour tout x tq. $\mathbb{P}(X = x) > 0$. De plus on a que la v.a. $\mathbb{P}(u_f(X) \neq g(X)) = 0$ et donc que $u_f(X) = g(X)$ presque sûrement.

On utilisera donc l'eq. (3) pour définir directement l'espérance conditionnelle comme une v.a. $\mathbb{E}[f(Y)|X] = u_f(X)$ telle que (3) soit satisfaite pour tout h mesurable bornée (en effet il suffit que soit valable pour tout $h(x) = 1_A(x)$ fonction caractéristique d'un ensemble A mesurable). Le prix à payer est que l'unicité de l'espérance conditionnelle est valable « à ensemble de mesure nulle prés »: il est clair que on peut changer la définition de $u_f(x)$ pour tout $x \in B$ où B est n'importe quel evenement tel que $\mathbb{P}(X \in B) = 0$ sans affecter la propriété (3). L'avantage majeure de cette nouvelle définition est sa généralité.

3 Sous-tribus

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Soit X une variable aléatoire ($\mathcal{F}$-mesurable) à valeurs dans $\mathbb{R}^d$. La tribu $\sigma(X)$ engendrée par X est la plus petite sous-tribu $\mathcal{B} \subseteq \mathcal{F}$ telle que X soit $\mathcal{B}$-mesurable. Une caractérisation équivalente (pour laquelle on donne pas de preuve) est que $\sigma(X)$ est l'image par X^{-1} de la tribu Borelienne de $\mathbb{R}^d$:

$$\sigma(X) = \left\{A \in \mathcal{F}: \exists\, B \in \mathcal{B}(\mathbb{R}^d), A = \{\omega \in \Omega: X(\omega) \in B\}\right\}.$$

En général si $(X_1, \ldots, X_n)$ est un n-uplet de v.a. réelles alors $\sigma(X_1, \ldots, X_n) = \sigma(X)$ où $X = (X_1, \ldots, X_n)$ est une v.a. à valeurs dans $\mathbb{R}^n$. La généralisation à une suite finie des v.a. vectorielles ne pose pas de problèmes.

On utilisera le résultat suivante (admis) dans la suite.

Théorème 1. *Soit X une v.a. à valeurs dans l'espace mesurable $(\Theta, \mathcal{H})$ et Y une v.a. réelle $\sigma(X)$ mesurable. Alors il existe une fonction mesurable $h: \Theta \to \mathbb{R}$ telle que $Y = h(X)$.*

$$\begin{array}{ccc} (\Omega, \sigma(X)) & \stackrel{X}{\longrightarrow} & (\Theta, \mathcal{H}) \\ {\scriptstyle Y}\searrow & & \swarrow{\scriptstyle h\circ X} \\ & (\mathbb{R}, \mathcal{B}(\mathbb{R})) & \end{array}$$

Exemple 2. Soit $X = 1$ avec probabilité p et $X = 0$ avec probabilité $1 - p$. Alors

$$\sigma(X) = \left\{\emptyset, \Omega, X^{-1}(\{0\}), X^{-1}(\{1\})\right\}.$$

Les sous-tribus $\mathcal{B} \subseteq \mathcal{F}$ que nous allons considérer vont correspondre à des informations partielles sur le modèle. Au sens que l'information nous permet de répondre de façon sûre à la question si l'expérience aléatoire $\omega \in \Omega$ appartient à un ensemble $B \in \mathcal{B}$. Par exemple $\sigma(X)$ doit être interprété comme l'information sur ω donnée par la connaissance de la valeur de X. La tribu triviale $\{\emptyset, \Omega\}$ correspond à une information nulle, et la tribu $\mathcal{F}$ est une information complète.

Exemple 3. Soit $\Omega = [0, 1]$, et $\mathcal{F} = \mathcal{B}([0, 1])$, la tribu borelienne sur $[0, 1]$ (i.e. la tribu engendrée par les sous-ensembles ouverts de $[0, 1]$). Soit $\mathcal{F}_1 = \sigma(0, 1/2],]1/2, 1]) = \{[0, 1/2],]1/2, 1], [0, 1], \emptyset\}$. Alors la tribu $\mathcal{F}_1 \subset \mathcal{F}$, et permet de savoir si le point se trouve à gauche ou à droite de 1/2. En particulier, si $X_1(\omega) = 1_{[0,1/2]}(\omega), \omega \in \Omega$, alors $\mathcal{F}_1 = \sigma(X_1)$. Soit maintenant $X_2(\omega) = 1_{[0,1/4]}(\omega) + 1_{]1/2,3/4]}(\omega)$, et $\mathcal{F}_2 = \sigma(X_1, X_2)$. Alors

$$\mathcal{F}_2 = \sigma([0, 1/4],]1/4, 1/2],]1/2, 3/4],]3/4, 1])$$

mais $\sigma(X_2) \neq \sigma(X_1, X_2)$. Connaître la valeur de X_1 situe ω à droite ou a gauche de $1/2$. Connaître X_2 situe ω soit dans l'ensemble $[0, 1/4] \cup]1/2, 3/4]$ où dans son complémentaire par rapport à $[0, 1]$. La connaissance de $X_1(\omega), X_2(\omega)$ permet de situer ω dans un des ensembles: $[0, 1/4],]1/4, 1/2],]1/2, 3/4],]3/4, 1]$. En passant on remarque que si on considère la probabilité $\mathbb{P}$ uniforme sur $[0, 1]$ les v.a. X_1 et X_2 sont indépendantes et de loi Bernoulli(1/2).

Si $X{:}\,\Omega \to \{x_1, x_2, ...\}$ est une v.a. discrète (avec un nombre fini ou infini de valeurs possibles) et $A_k = \{X = x_k\} = \{\omega \in \Omega{:}\ X(\omega) = x_k\}$ alors $\sigma(X) = \sigma(A_1, A_2, ...)$ est la tribu engendrée par les $\{A_k\}_{k \geqslant 1}$ (la plus petite tribu qui contienne les A_k). Dans ce cas l'espérance conditionnelle $Z = \mathbb{E}[f(Y)|X]$ vérifie

$$Z(\omega) = u_f(X(\omega)) = u_f(x_k) = \sum_y f(y) \frac{\mathbb{P}(A_k, Y = y)}{\mathbb{P}(A_k)} = \frac{\mathbb{E}[f(Y) 1_{A_k}]}{\mathbb{E}[1_{A_k}]}$$

pour tout $\omega \in A_k$ tel que $\mathbb{P}(A_k) > 0$ et donc

$$Z(\omega) = \sum_{k : \mathbb{P}(A_k) > 0} \frac{\mathbb{E}[f(Y) 1_{A_k}]}{\mathbb{E}[1_{A_k}]} 1_{A_k}(\omega) \ p.s.$$

ce qui montre qu'elle dépend seulement de la tribu $\sigma(X)$ et non pas de la v.a. X (deux v.a.s X et X' peuvent générer la même tribu $\sigma(X) = \sigma(X')$, dans ce cas l'espérance conditionnelle est la même). Cet observation justifie la définition générale d'espérance conditionnelle par rapport à une sous-tribu de $\mathcal{F}$.

4 Espérance conditionnelle

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, et soit $\mathcal{B} \subseteq \mathcal{F}$ une sous-tribu de $\mathcal{F}$. Soit $X{:}\,\Omega \to \mathbb{R}$ une variable aléatoire telle que $\mathbb{E}(|X|) < +\infty$ (i.e. $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$) .

Définition 4. *L'espérance conditionnelle de X sachant $\mathcal{B}$ est une variable aléatoire $Y \hat{\in} \mathcal{B}$ telle que*

$$\mathbb{E}[1_A X] = \mathbb{E}[1_A Y] \qquad \forall A \in \mathcal{B} \tag{4}$$

L'assertion (4) est en fait équivalente à

$$\mathbb{E}[Z X] = \mathbb{E}[Z Y] \qquad \forall Z \hat{\in} \mathcal{B} \,\text{bornée} \tag{5}$$

L'existence d'une variable aléatoire Y qui a ces propriétés n'est pas triviale, on l'admettra. Par ailleurs, cette variable aléatoire est unique à l'égalité presque-sûre près (voir la preuve du 2. de la proposition suivante).

On utilisera les notations $Y = \mathbb{E}(X|\mathcal{B})$, ainsi que $\mathbb{E}(X|Z) = \mathbb{E}(X|\sigma(Z))$. La probabilité conditionnelle $\mathbb{P}(\cdot|\mathcal{B})$ sachant $\mathcal{B}$ (ou par rapport à $\mathcal{B}$) est définie par $\mathbb{P}(A|\mathcal{B}) = \mathbb{E}[1_A|\mathcal{B}]$. On remarque que $\mathbb{P}(A|\mathcal{B})$ est une variable aléatoire (donc la probabilité conditionnelle par rapport à une tribu n'est pas une probabilité proprement dite).

Proposition 5.

1. $\mathbb{E}(X|\mathcal{B}) \in L^1(\Omega, \mathcal{B}, \mathbb{P})$.
2. *Soient Y, Y' deux espérances conditionnelles de X sachant $\mathcal{B}$, alors $Y = Y'$ p.s.. En particulier si $X \hat{\in} \mathcal{B}$ alors $\mathbb{E}(X|\mathcal{B}) = X$ p.s.*

Démonstration.

1. Soit $Z = \operatorname{sgn} \mathbb{E}(X|\mathcal{B}) \hat{\in} \mathcal{B}$. Donc $Z \mathbb{E}(X|\mathcal{B}) = |\mathbb{E}(X|\mathcal{B})|$, et d'après b') on a

$$0 \leqslant \mathbb{E}(|\mathbb{E}(X|\mathcal{B})|) = \mathbb{E}(Z \mathbb{E}(X|\mathcal{B})) = \mathbb{E}(Z X) \leq \mathbb{E}(|X|) < +\infty$$

On vient de montrer que l'espérance conditionnelle $X \mapsto \mathbb{E}[X|\mathcal{B}]$ est une contraction dans $L^1(\Omega, \mathcal{F}, \mathbb{P})$.

2. Par définition on a que $\mathbb{E}[1_A(Y - Y')] = 0$ pour tout $A \in \mathcal{B}$. Soit $A_n = \{Y - Y' \geqslant 1/n > 0\} \in \mathcal{F}$. L'événement A étant aussi $\mathcal{B}$-mesurable on a
$$0 = \mathbb{E}[1_{A_n}(Y - Y')] \geqslant n^{-1}\mathbb{E}[1_{A_n}] = \mathbb{P}(A_n)/n$$
et donc $\mathbb{P}(A_n) = 0$. Mais alors $0 = \mathbb{P}(\cup_{n\geqslant 1} A_n) = \mathbb{P}(Y > Y')$ car $\cup_{n\geqslant 1} A_n = \{Y > Y'\}$ (vérifier). Par symétrie on a aussi que $\mathbb{P}(Y < Y') = 0$ et donc que $\mathbb{P}(Y = Y') = 1 - \mathbb{P}(Y \neq Y') = 1 - \mathbb{P}(Y < Y') - \mathbb{P}(Y > Y') = 1$. □

Exemple 6. Pour la tribu triviale $\mathcal{B} = \{\emptyset, \Omega\}$, on a $\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(X)$. Il suffit de vérifier que la constante $\mathbb{E}[X]$ satisfait bien la définition (4).

Si on conditionne par rapport à une v.a. X donnée on trouve bien que la probabilité conditionnelle est une fonction des valeurs de X:

Proposition 7. *Il existe une fonction mesurable h_Z telle que $\mathbb{E}[Z|X] = h_Z(X)$ p.s.*

Démonstration. La v.a. $\mathbb{E}[Z|X]$ est $\sigma(X)$-mesurable. On utilise donc le théorème 1. □

Exercice 1. Montrer que si $X_1 = X_2$ sur $B \in \mathcal{F}$ (c.-à.-d. $X_1(\omega) = X_2(\omega)$ si $\omega \in B$), alors $\mathbb{E}[X_1|\mathcal{F}] = \mathbb{E}[X_2|\mathcal{F}]$ sur $B \in \mathcal{F}$.

Exercice 2. Soit $\{A_1, A_2, \ldots\}$ une partition (finie ou infinie) de Ω. Soit $\mathcal{B} = \sigma(A_1, \ldots)$ la tribu engendrée par cette partition. Alors
$$\mathbb{E}(X|\mathcal{B})(\omega) = \sum_{j:\mathbb{P}(A_j)>0} \frac{\mathbb{E}(X 1_{A_j})}{\mathbb{P}(A_j)} 1_{A_j}(\omega). \tag{6}$$

Exercice 3. [Formule de Bayes] Montrer que si $\mathcal{G}$ est une sous-tribu de $\mathcal{F}$ et $A \in \mathcal{F}$ on a
$$\mathbb{P}(G|A) = \frac{\mathbb{E}[\mathbb{P}(A|\mathcal{G})1_G]}{\mathbb{E}[\mathbb{P}(A|\mathcal{G})]}.$$

Exercice 4. On considère deux v.a. X, Y: X est uniforme sur l'ensemble $\{1, \ldots, 6\}$ et conditionnellement à X la v.a. Y a une loi $\mathrm{Bin}(X, 1/2)$. Calculer $\mathbb{P}(X = i|Y = 0)$ pour $i = 1, \ldots, 6$.

Exercice 5. Soient X_1 et X_2 des v.a. indépendantes et $\sim \mathrm{Poisson}(\lambda)$ avec $\lambda > 0$. Soit $Y = X_1 + X_2$. Calculer $\mathbb{P}(X_1 = k|Y)$.

Exercice 6. Soient (X, Y) une couple des v.a. à valeurs dans $\mathbb{R}^n \times \mathbb{R}^m$ avec densité jointe $f_{X,Y}(x, y)$. Montrer que $\mathbb{E}[g(Y)|X] = h(X)$ où h est n'importe quelle fonction telle que
$$h(x)\int_{\mathbb{R}^m} f_{X,Y}(x, y)\mathrm{d}y = \int_{\mathbb{R}^m} g(y) f_{X,Y}(x, y)\mathrm{d}y.$$
pour tout $x \in \mathbb{R}^n$. En particulier pour tout $x \in \mathbb{R}^n$ tel que $f_X(x) = \int_{\mathbb{R}^m} f_{X,Y}(x, y)\mathrm{d}y > 0$ on a que
$$h(x) = \int_{\mathbb{R}^m} g(y)\frac{f_{X,Y}(x, y)}{f_X(x)}\mathrm{d}y = \int_{\mathbb{R}^m} g(y) f_{Y|X=x}(y)\mathrm{d}y = \mathbb{E}[g(Y)|X = x]$$
selon la définition usuelle d'espérance conditionnelle de $g(Y)$ sachant $X = x$ pour les vecteurs aléatoires avec densité.

Exercice 7. Soient X_1 et X_2 indépendantes avec $\mathbb{P}(X_i > t) = e^{-t}$ pour $t \geqslant 0$. Soit $Y = X_1 + X_2$. Calculer $\mathbb{E}[X_1|Y]$ et $\mathbb{P}(X_1 < 3|Y)$.

Proposition 8. *Pour tout $X, Y \in L^1(\mathcal{F})$ et tout sous-tribu $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ on a les propriétés suivantes:*

1. *Linéarité: $\mathbb{E}[X + Y|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$;*
2. *Positivité: $X \geqslant 0\,p.s. \Rightarrow \mathbb{E}[X|\mathcal{G}] \geqslant 0\,p.s.$*

3. *Convergence monotone:* $0 \leqslant X_n \nearrow X\, p.s. \Rightarrow \mathbb{E}[X_n|\mathcal{G}] \nearrow \mathbb{E}[X|\mathcal{G}]\, p.s.$
4. *Inégalité de Jensen: pour tout fonction* $\varphi: \mathbb{R} \to \mathbb{R}$ *convexe on a* $\mathbb{E}[\varphi(X)|\mathcal{G}] \geqslant \varphi(\mathbb{E}[X|\mathcal{G}])$
5. *Contractivité dans* L^p*:* $\|\mathbb{E}[X|\mathcal{G}]\|_p \leqslant \|X\|_p$.
6. *Emboîtement: Si* $\mathcal{H}$ *est une sous-tribu de* $\mathcal{G}$ *alors* $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]$.
7. *Si* $Z \hat{\in} \mathcal{G}$, $\mathbb{E}[|X|] < \infty$ *et* $\mathbb{E}[|XZ|] < +\infty$ *alors* $\mathbb{E}[XZ|\mathcal{G}] = Z\,\mathbb{E}[X|\mathcal{G}]$.

Démonstration.

1. Exercice.
2. on remarque que si $\mathbb{E}[X|\mathcal{G}] \leqslant \varepsilon < 0$ sur $A \in \mathcal{G}$ tel que $\mathbb{P}(A) > 0$ alors $0 < \mathbb{E}[X 1_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]1_A] \leqslant \varepsilon \mathbb{P}(A) < 0$ ce qui est impossible.
3. Soit $Y_n = \mathbb{E}[X_n|\mathcal{G}]$. Par la positivité de l'esp. cond. on a que Y_n est une suite croissante. Soit $Y = \limsup_n Y_n$ alors $Y \hat{\in} \mathcal{G}$ et le théorème de convergence monotone nous permet de passer à la limite dans l'égalité $\mathbb{E}[X_n 1_A] = \mathbb{E}[Y_n 1_A]$ pour obtenir que $\mathbb{E}[X 1_A] = \mathbb{E}[Y 1_A]$ pour tout $A \in \mathcal{G}$. Donc $Y = \mathbb{E}[X|\mathcal{G}]$ p.s.
4. Admise.
5. On utilise la propriété (4). Exercice.
6. Exercice.
7. Admise. (Facile pour des fonctions étagées, utiliser des limites monotones dans le cas X, $Z \geqslant 0$ et conclure). □

Définition 9. *On dit que* X *est une variable aléatoire indépendante de la tribu* $\mathcal{B}$ *si* $\sigma(X)$ *est indépendante de* $\mathcal{B}$*, ou plus expliciment si* $\forall I \in \mathcal{B}(\mathbb{R})$ *et* $\forall A \in \mathcal{B}$*:* $\mathbb{P}(\{X \in I\} \cap A) = \mathbb{P}(X \in I)\, \mathbb{P}(A)$.

Proposition 10. *Soit* X *indépendante de la sous-tribu* $\mathcal{B}$*, alors* $\mathbb{E}(X|\mathcal{B})$ *est constante presque sûrement et*

$$\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(X) \qquad p.s.$$

Démonstration. $\mathbb{E}(X)$ est une constante, donc $\mathbb{E}(X) \hat{\in} \mathcal{B}$. Pour tout $A \in \mathcal{B}$

$$\mathbb{E}(\mathbb{E}(X|\mathcal{B})1_A) = \mathbb{E}(X 1_A) = \mathbb{E}(X)\mathbb{E}(1_A) = \mathbb{E}(\mathbb{E}(X)1_A).$$ □

Proposition 11. *Si* $\mathcal{H}$ *et* $\mathcal{G}$ *sont indépendantes,* X *est* $\mathcal{G}$*-mesurable et* $\mathcal{G}' \subseteq \mathcal{G}$*, alors*

$$\mathbb{E}[X|\mathcal{H}, \mathcal{G}'] = \mathbb{E}[X|\mathcal{G}'].$$

Démonstration. Admise. □

Proposition 12. *(Conditionnement et indépendance) Si* X_1, ..., X_n *est une famille des v.a. indépendantes et* $f(X_1, ..., X_n) \in L^1$ *alors*

$$\mathbb{E}[f(X_1, ..., X_n)|X_1] = \varphi(X_1)$$

où $\varphi(x) = \mathbb{E}[f(x, X_2, ..., X_n)]$.

Démonstration. Utiliser le théorème de Fubini sur la loi jointe de $X_1, ..., X_n$. □

Exercice 8. Soient X_i, $1 \leq i \leq n$ des v.a. i.i.d et $T = \sum_{i=1}^n X_i$. Montrer que $\mathbb{E}[T_1|T] = T/n$ et $\mathbb{E}[T|T_1] = (n-1)\mathbb{E}[T_1] + T_1$.

Exercice 9. Un modèle discret d'évolution d'actifs. Soit S_0 une constante, $0 < d < u$ et X_n une suite iid à valeurs dans $\{u, d\}$ telle que $\mathbb{P}(X_n = u) = p$. On considère la suite S_n, $n \geqslant 1$ définie par $S_n = X_n S_{n-1}$ pour $n \geqslant 1$ qui est un modèle très simplifié d'évolution d'un actif financier. Soit $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_1 = \sigma(X_1)$, $\mathcal{F}_2 = \sigma(X_1, X_2)$.

a) Montrer que $\sigma(S_2) \neq \sigma(X_1, X_2)$.

b) Calculer $\mathbb{E}[S_2|\mathcal{F}_1]$ et $\mathbb{E}[S_2|\mathcal{F}_0]$ et vérifier que $\mathbb{E}[\mathbb{E}[S_2|\mathcal{F}_1]|\mathcal{F}_0] = \mathbb{E}[S_2]$.

c) Si $\mathcal{F}_n = \sigma(X_1, ..., X_n)$ donner une formule pour $\mathbb{E}[S_n|\mathcal{F}_k]$ pour tout $k \leqslant n$.