Chapitre 3

Espérance conditionnelle

Dans les leçons précédentes nous avons appris à calculer le prix de diverses options à l'instant $t = 0$. Les options étant des actifs financiers négociables, qu'on veut pouvoir acheter et vendre aussi à des instants ultérieurs, on voudrait également savoir décrire et calculer leurs prix à des instants t tels que $0 < t \leq T$. C'est ce que nous allons faire dans cette leçon grâce à la notion d'*espérance conditionnelle par rapport à une tribu*. On découvrira au passage que cette espérance conditionnelle, qui généralise l'espérance conditionnelle usuelle d'une variable aléatoire par rapport à un évènement, est en fait un extraordinaire outil de calcul, un de ceux qui font que la théorie des probabilités s'appelle souvent le *calcul* des probabilités.

3.1 Espérance conditionnelle d'une v.a. sachant un évènement

Voici tout d'abord quelques rappels de probabilités élémentaires. Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé que nous supposerons comme précédemment *fini* et satisfaisant en outre la condition suivante [1] :

$$\forall \omega \in \Omega, \mathbb{P}(\omega) \neq 0.$$

Soit X une variable aléatoire sur Ω. L'espérance de X est le *nombre* $\mathbb{E}\,(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega)$. Notons que, pour tout évènement $A \subseteq \Omega$ appartenant à la tribu $\mathcal{T}$, on peut exprimer la probabilité de A comme l'espérance d'une v.a., en posant $\mathbb{P}(A) = \sum_{\alpha \in A} \mathbb{P}(\alpha) = \mathbb{E}\,\mathbb{I}_A$, où $\mathbb{I}_A$ désigne la v.a. indicatrice de A, égale à 1 sur A et 0 sinon.

Plus généralement, on appelle "espérance de X sachant A" ou "espérance de X conditionnellement à A", le *nombre*, notée $\mathbb{E}(X/A)$, donné par

$$\mathbb{E}(X/A) = \frac{1}{\mathbb{P}(A)} \sum_{\alpha \in A} X(\alpha)\mathbb{P}(\alpha) = \frac{\mathbb{E}\,(X\mathbb{I}_A)}{\mathbb{E}\,\mathbb{I}_A}.$$

En d'autres termes $\mathbb{E}(X/A)$ est la moyenne des éléments de A, pondérée par leur probabilité rapportée à la probabilité de A.

Nous insistons sur le fait que l'espérance conditionnelle d'une v.a. sachant un évènement est un nombre. L'espérance conditionnelle par rapport à une tribu, que nous allons introduire à présent, n'est pas *un* nombre, mais "un nombre qui dépend de l'état du monde", c'est-à-dire une variable aléatoire.

3.2 Espérance conditionnelle d'une v.a. par rapport à une tribu

Soient $\mathcal{F} \subseteq \mathcal{T}$ une tribu et $\wp := \{\Omega_1, \ldots, \Omega_m\}$ la partition de Ω formée par les atomes de $\mathcal{F}$.

Définition : Soit X est une v.a. sur Ω. On appelle *espérance conditionnelle de X par rapport à la tribu $\mathcal{F} \subseteq \mathcal{T}$*, ou encore *espérance conditionnelle de X relativement à la partition $\wp$*, la v.a. notée $\mathbb{E}(X|\mathcal{F})$ définie, pour tout $\omega \in \Omega$, par

$$\mathbb{E}(X|\mathcal{F})(\omega) := \mathbb{E}[X/\overline{\omega}] = \frac{1}{\mathbb{P}(\overline{\omega})} \sum_{\alpha \in \overline{\omega}} X(\alpha)\mathbb{P}(\alpha)$$

1. conformément à l'usage, nous notons $\mathbb{P}(\omega)$ pour $\mathbb{P}(\{\omega\})$

où $\overline{\omega}$ désigne l'atome $\Omega_i \in \mathfrak{P}$ de la partition tel que $\omega \in \Omega_i$.

On voit donc que par définition l'espérance conditionnelle par rapport à une tribu $\mathcal{F}$ est une v.a. constante sur les atomes Ω_i de la partition $\mathfrak{P}$ associée à $\mathcal{F}$. On a plus précisément :

Proposition 3.1 *La v.a.* $\mathbb{E}(X|\mathcal{F})$ *est mesurable par rapport à la tribu* $\mathcal{F}$ *et de plus si* $Y = \mathbb{E}(X|\mathcal{F})$ *désigne cette v.a., alors* Y *peut être définie comme l'unique v.a.* $\mathcal{F}$*-mesurable telle que*

$$\forall \Omega_i \in \mathfrak{P} \quad , \quad \mathbb{E}(Y/\Omega_i) = \mathbb{E}(X/\Omega_i). \tag{3.1}$$

Preuve : Le fait que $\mathbb{E}(X|\mathcal{F})$ soit mesurable par rapport à la tribu $\mathcal{F}$ est clair par définition puisqu'elle est constante sur les atomes de la partition $\mathfrak{P}$ associée à $\mathcal{F}$.

Montrons qu'elle vérifie l'équation (3.1) : Soit ω est un élément quelconque de l'atome Ω_i, et donc $\overline{\omega} = \Omega_i$; comme Y est constante sur les atomes on a $Y(\omega) = Y(\alpha)$ pour tout $\alpha \in \Omega_i$. Par définition de $Y = \mathbb{E}(X|\mathcal{F})$, on a

$$\mathbb{E}(Y/\Omega_i) = \frac{1}{\mathbb{P}(\Omega_i)} \sum_{\alpha \in \Omega_i} Y(\alpha)\mathbb{P}(\alpha) = Y(\omega)\frac{1}{\mathbb{P}(\Omega_i)} \sum_{\alpha \in \Omega_i} \mathbb{P}(\alpha) = \mathbb{E}(X|\mathcal{F})(\omega)\frac{\mathbb{P}(\Omega_i)}{\mathbb{P}(\Omega_i)} = \mathbb{E}[X/\overline{\omega}] = \mathbb{E}(X/\Omega_i).$$

Réciproquement si Y est $\mathcal{F}$-mesurable, elle est constante sur les $\Omega_i \in \mathfrak{P}$, et la relation $\mathbb{E}(Y/\Omega_i) = \mathbb{E}(X/\Omega_i)$ définit Y uniquement car, pour tout $\omega \in \Omega$, on aura $Y(\omega) := \mathbb{E}(X/\Omega_i)$, où Ω_i est l'atome contenant ω. $\square$

Voici les principales propriétés de l'espérance conditionnelle :

Proposition 3.2 *Soient* X *et* Y *des v.a. sur* $(\Omega, \mathcal{T}, \mathbb{P})$*,* $\mathcal{F}$ *et* $\mathcal{G}$ *des sous-tribus de* T*, et* x_0*,* a *et* b *des nombres réels. On a :*

1. $\mathbb{E}(X|\mathcal{T}) = X$*, et* $\mathbb{E}(X|\{\emptyset, \Omega\}) = \mathbb{E}(X)$*.*
2. $\mathbb{E}(aX + bY|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + b\mathbb{E}(Y|\mathcal{F})$*.*
3. *Si* $X \geq 0$*, on a* $\mathbb{E}(X|\mathcal{F}) \geq 0$*, avec égalité si et seulement si* $X = 0$*.*
4. $\mathbb{E}(x_0|\mathcal{F}) = x_0$*.*
5. *Si* Y *est* $\mathcal{F}$*-mesurable, on a* $\mathbb{E}(XY|\mathcal{F}) = Y\mathbb{E}(X|\mathcal{F})$
6. *Si* $\mathcal{G} \subseteq \mathcal{F}$*, on a* $\mathbb{E}(\ \mathbb{E}(X|\mathcal{F}) \ | \ \mathcal{G}) = \mathbb{E}(X|\mathcal{G})$*. En particulier* $\mathbb{E}(\mathbb{E}(X|\mathcal{F})) = \mathbb{E}(X)$*.*

Cette sixième propriété s'appelle la transitivité des espérances conditionnelles. Elle est d'un usage fréquent en Finance.

3.3 L'espace euclidien $L^2(\Omega)$

On désigne par $L^2(\Omega)$ l'ensemble des variables aléatoires sur Ω dans le cas où on munit cet ensemble d'une structure euclidienne (c'est-à-dire d'un produit scalaire) que nous allons définir à présent. Cet espace de v.a. joue un rôle fondamental en économétrie et en calcul stochastique.

Tout d'abord, il est facile de munir l'ensemble des v.a. sur Ω d'une structure d'espace vectoriel : si X et Y sont deux éléments de cet ensemble, $X + Y$ et, pour tout $\lambda \in \mathbb{R}$, λX sont encore des v.a. sur Ω. D'autre part on peut définir un produit scalaire sur cet ensemble de la façon suivante :

$$< X, Y > := \mathbb{E}(XY) \tag{3.2}$$

et la norme associée par :

$$\|X\| := \sqrt{\mathbb{E}(X^2)}$$

En effet on vérifie facilement la bilinéarité et la symétrie ; de plus on a $\|X\| \geq 0$ puisque X^2 est une v.a. positive ou nulle ; enfin, si $\|X\| = 0$, alors $X = 0$ car $\mathbb{E}(X^2)$ est une combinaison linéaire à coefficients strictement positifs (puisque $\mathbb{P}(\omega) > 0$ pour tout $\omega \in \Omega$) des nombres positifs.

On vérifie sans peine que l'on a $\mathbb{E}(X) =< X, 1 >$, $Cov(X, Y) =< X - \mathbb{E}(X), Y - \mathbb{E}(Y) >$ et donc $Var(X) =< X - \mathbb{E}(X), X - \mathbb{E}(X) >= \|X - \mathbb{E}(X)\|^2$.

L'espace $L^2(\Omega)$ des variables aléatoires sur Ω, muni du produit scalaire (3.2), est donc bien un espace euclidien. On sait que dans un espace euclidien on peut associer à tout vecteur sa projection orthogonale sur un sous espace. Rappelons la définition de la projection orthogonale :

Définition : Soient $G \subseteq L^2(\Omega)$ un sous espace vectoriel et $X \in L^2(\Omega)$ un vecteur quelconque. On appelle *projection orthogonale* de X sur G, notée $\pi(X)$ l'unique élément de G tel que

$$\forall Z \in G \ , \ < X - \pi(X), Z >= 0$$

On peut vérifier facilement que si $\mathcal{F}$ est une tribu alors l'ensemble des v.a. de $L^2(\Omega)$ $\mathcal{F}$-mesurables forment un sous espace vectoriel. En effet, si X et Y sont deux v.a. $\mathcal{F}$-mesurables, donc constantes sur les atomes de la partition associée à la tribu $\mathcal{F}$, toute combinaison linéaires $\lambda X + \mu Y$, pour λ et μ réels quelconques, est encore $\mathcal{F}$-mesurable.

Proposition 3.3 *Soient X une v.a. et $\mathcal{F}$ une tribu. L'espérance conditionnelle de X par rapport à la tribu $\mathcal{F}$ est la projection orthogonale de X sur le sous espace G des v.a. $\mathcal{F}$-mesurables.*

Preuve : Soit $Z \in G$. En utilisant les propriétés de l'espérance conditionnelle du théorème 3.2, on a :

$$\mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}(XZ|\mathcal{F})) = \mathbb{E}(\mathbb{E}(X|\mathcal{F})Z)$$

donc $< X - \mathbb{E}(X|\mathcal{F}), Z >= \mathbb{E}(XZ) - \mathbb{E}(\mathbb{E}(X|\mathcal{F})Z) = 0$. □

3.4 Application au calcul de prix d'options

Soit $X = (X_t)_{t\in\mathbb{T}}$ une marche aléatoire et soit $(\mathcal{F}_t^X)_{t\in\mathbb{T}}$ la filtration associée; rappelons que la filtration associée à une m.a. est par définition la suite croissante de tribus

$$\{\phi, \Omega\} = \mathcal{F}_0^X \subset \mathcal{F}_{\delta t}^X \subset \ldots \subset \mathcal{F}_T^X \subseteq \mathcal{P}(\Omega)$$

définie de la façon suivante : pour chaque $t \in \mathbb{T}$,les atomes de la tribu $\mathcal{F}_t^X$ sont les classes d'équivalence pour la relation :

$$\omega' \overset{t}{\sim} \omega'' \text{ si et seulement si } X_s(\omega') = X_s(\omega'') \text{ pour tout } s \in [0..t].$$

L'exemple de filtration importante pour nous et la filtration associée à la marche CRR est détaillé au chapitre 3.

On a vu ci-dessus comment associer à une v.a. X son espérance conditionnellement à une tribu, $\mathbb{E}(X|\mathcal{F}) =: Y$. Si l'on dispose non plus d'une seule tribu mais de toute une filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$, on peut associer alors, à toute v.a. X, une famille, indexée par $t \in \mathbb{T}$, de variables aléatoires $Y_t := \mathbb{E}(X|\mathcal{F}_t)$, c'est-à-dire une nouvelle marche aléatoire $(Y_t)_{t\in\mathbb{T}}$.

Prenons l'exemple d'une option enropéenne standard $(T, \varphi(S_T))$ souscrite sur une actif $(S_t)_{t\in\mathbb{T}}$. La fonction de paiement $X := \varphi(S_T)$ est une v.a. sur l'ensemble des états du monde Ω sur lesquel est définie la m.a. (S_t). On peut donc associer à l'option une nouvelle m.a. donnée par $Y_t := \mathbb{E}(\varphi(S_T)|\mathcal{F}_t)$. Que représente Y_t par rapport à $X := \varphi(S_T)$? Pour chaque état du monde $\omega \in \Omega$, $Y_t(\omega)$ est l'espérance conditionnelle de X sachant $\overline{\omega}$, où $\overline{\omega}$ désigne l'atome de la tribu $\mathcal{F}_t$, c'est-à-dire l'ensemble des états du monde correspondant à des trajectoires de la marche (S_t) qui coïncident jusqu'à l'instant t. En d'autres termes, $Y_t(\omega)$ est la moyenne des paiements attendus sur toutes les trajectoires qui coïncident avec ω jusqu'à l'instant t, ou la moyenne des paiements futurs sachant la trajectoire S_s jusqu'à l'instant t, c'est-à-dire connaissant l'information jusqu'à t. Notons que l'on a $\mathbb{E}(\varphi(S_T)|\mathcal{F}_T^S) = \varphi(S_T)$ (car $\varphi(S_T)$ est $\mathcal{F}_T^S$-mesurable) et $\mathbb{E}(\varphi(S_T)|\mathcal{F}_0^S) = \mathbb{E}(\varphi(S_T))$ (car $\mathcal{F}_0^S = \{\emptyset, \Omega\}$).

En fait, par le choix de $p = \frac{R-d}{u-d}$ et des lois des $\delta J = (\delta J_i)_{i=1..n}$, le nombre $\mathbb{E}(\varphi(S_T)|\mathcal{F}_t^S)(\omega)$ est précisément le prix du portefeuille de couverture à l'instant t, lorsque l'action vaut $S_t(\omega)$. En d'autres termes, en utilisant la notion d'espérance conditionnelle par rapport aux tribus de la filtration des prix du sous-jacent, il est possible de calculer le prix d'une option, non seulement à l'instant $t = 0$ mais à tout instant $t \in \mathbb{T}$: c'est la *formule fondamentale* qui généralise celle donnée au chapitre **??** pour $t = 0$. La seule hypothèse que nous faisons sur l'option est qu'elle soit européenne d'échéance T, c'est à dire qu'elle paye sa valeur Π_T à la date T et que cette valeur ne dépendent que des valeurs de S jusqu'à cette date, c'est-à-dire qu'elle soit $\mathcal{F}_T^S$-mesurable; par exemple $\Pi_T = \varphi(S_T)$ pour une option vanille de fonction de paiement φ.

Théorème 3.4 *Considérons un marché financier* $(S_t, B_t)_{t\in\mathbb{T}}$ *où l'actif risqué* S *suit un modèle CRR et l'actif sans risque est donné par* $B_t := e^{rt}$. *Considérons une option européenne d'échéance* T *et de valeur* Π_T *qui est* $\mathcal{F}_T^S$*-mesurable. Pour tout* $t \in \mathbb{T} = [0..T]$ *la valeur* Π_t *du portefeuille de couverture est donnée par*

$$\Pi_t = e^{-r(T-t)}\mathbb{E}(\Pi_T|\mathcal{F}_t^S) \tag{3.3}$$

où $(\mathcal{F}_t^S)_{t\in\mathbb{T}}$ *est la filtration associée à la m.a.* $S = (S_t)_{t\in\mathbb{T}}$ *et où l'espérance conditionnelle est calculée sous la probabilité de calcul définie par* $p = \frac{e^{r\delta t}-d}{u-d}$. *En particulier, la prime de l'option est egale à* $\Pi_0 = e^{-rT}\mathbb{E}(\Pi_T)$.

Rappelons que le calcul du prix $\Pi_t(\omega)$ du portefeuille de couverture s'effectue par récurrence rétrograde à partir de sa valeur finale Π_T, supposée $\mathcal{F}_T^S$-mesurable, c'est à dire que $\Pi_T(\omega)$ est fonction des valeur $s_{\delta t}$, $s_{2\delta t}$, ..., s_T des $S_t(\omega)$ pour $t \in \mathbb{T}$. De fait, plus généralement, Π_t est $\mathcal{F}_t^S$-mesurable, c'est à dire une fonction (dépendant de t) des $s_{\delta t}$, ..., s_t ; en effet supposons que ce soit le cas pour $t + \delta t$:

$$\Pi_{t+\delta t}(\omega) = \pi(t + \delta t, s_{\delta t}, \dots s_t, s_{t+\delta t}),$$

et plaçons-nous à la date t, lorsque nous savons que $S_t(\omega) = s_{\delta t}$, ..., $S_t(\omega) = s_t$. Dans ce cas $S_{t+\delta t}(\omega) = s_{\delta t}u$ ou $S_{t+\delta t}(\omega) = s_{\delta t}d$ et le calcul élémentaire vu au chapitre 2 nous montre que la valeur du portefeuille de couverture à la date t se déduit des deux valeurs correspondantes $\pi(t, s_{\delta t}, \dots s_t, s_t u)$ et $\pi(t, s_{\delta t}, \dots s_t, s_t d)$ à $t + \delta t$ par la formule

$$\Pi_t(\omega) = \frac{1}{R}\left(p\pi(t + \delta t, s_{\delta t}, \dots s_t, s_t u) + (1 - p)\pi(t + \delta t, s_{\delta t}, \dots s_t, s_t d)\right) =: \pi(t, s_{\delta t}, \dots s_t) \tag{3.4}$$

avec $R := e^{r\delta t}$ et $p := \frac{R-d}{u-d}$, et qui est bien une fonction des seules valeurs de $S_s(\omega)$ pour $s \le t$. Reformulons ce résultat en termes d'espérance conditionnelle :

Lemme 3.5

$$\Pi_t = \frac{1}{R}\mathbb{E}(\Pi_{t+\delta t} \mid \mathcal{F}_t^S). \tag{3.5}$$

Preuve : Pour tout $\omega \in \Omega$, soient $s_{i\delta t} := S_{i\delta t}(\omega)$, $i = 1..n$. Rappelons que $\overline{\omega}_t$ est l'atome de ω dans la tribu$\mathcal{F}_t^S$,

$$\overline{\omega}_t = \{\omega' \in \Omega \mid S_{\delta t}(\omega') = s_{\delta t}, \dots, S_t(\omega') = s_t\}.$$

Par définition de l'espérance conditionnelle nous avons

$$\begin{aligned}
& \mathbb{E}(\Pi_{t+\delta t} \mid \mathcal{F}_t^S)(\omega) \\
= \;& \mathbb{E}(\Pi_{t+\delta t}\mathbb{I}_{\overline{\omega}_t})\frac{1}{\mathbb{P}(\overline{\omega}_t)} \\
= \;& \mathbb{E}\left(\pi\left(t + \delta t, S_{\delta t}, \dots, S_t, S_t d\,(u/d)^{\delta J_{t+\delta t}}\right)\mathbb{I}_{\{S_{\delta t}=s_{\delta t}\}}\cdots\mathbb{I}_{\{S_t=s_t\}}\right)\frac{1}{\mathbb{E}(\mathbb{I}_{\{S_{\delta t}=s_{\delta t}\}}\cdots\mathbb{I}_{\{S_t=s_t\}})} \\
= \;& \mathbb{E}\left(\pi\left(t + \delta t, s_{\delta t}, \dots, s_t, s_t d\left(\frac{u}{d}\right)^{\delta J_{t+\delta t}}\right)\mathbb{I}_{\{S_{\delta t}=s_{\delta t}\}}\cdots\mathbb{I}_{\{S_t=s_t\}}\right)\frac{1}{\mathbb{E}(\mathbb{I}_{\{S_{\delta t}=s_{\delta t}\}}\cdots\mathbb{I}_{\{S_t=s_t\}})} \\
= \;& \mathbb{E}\left(\pi\left(t + \delta t, s_{\delta t}, \dots, s_t, s_t d\left(\frac{u}{d}\right)^{\delta J_{t+\delta t}}\right)\right)\frac{\mathbb{E}(\mathbb{I}_{\{S_{\delta t}=s_{\delta t}\}}\cdots\mathbb{I}_{\{S_t=s_t\}})}{\mathbb{E}(\mathbb{I}_{\{S_{\delta t}=s_{\delta t}\}}\cdots\mathbb{I}_{\{S_t=s_t\}})} \text{ , par independance,} \\
= \;& p\pi(t + \delta t, s_{\delta t}, \dots, s_t, s_t u) + (1 - p)\pi(t + \delta t, s_{\delta t}, \dots, s_t, s_t d) \text{ , puisque } \delta J_{t+\delta t} \rightsquigarrow \mathcal{B}(1, p), \\
= \;& R\,\pi(t, s_{\delta t}, \dots, s_t) \text{ , par (3.4),} \\
= \;& R\,\pi(t, S_{\delta t}, \dots, S_t)(\omega) = R\,\Pi_t(\omega).
\end{aligned}$$

□

La preuve du théorème s'en déduit facilement par transitivité de l'espérance conditionnelle : posons $t = i\delta t$; par $n - i$ applications du lemme on obtient

$$\Pi_t = \frac{1}{R}\mathbb{E}\left(\frac{1}{R}\mathbb{E}\left(\dots\frac{1}{R}\mathbb{E}(\Pi_T \mid \mathcal{F}_{T-\delta t}^S)\dots \mid \mathcal{F}_{t+\delta t}^S\right) \mid \mathcal{F}_t^S\right) = \frac{1}{R^{n-i}}\mathbb{E}(\Pi_T|\mathcal{F}_t^S) = e^{-r(T-t)}\mathbb{E}(\Pi_T|\mathcal{F}_t^S).$$