Cours 4 : Espérance conditionnelle

1 Marches aléatoires

On rappelle que $\mathcal{P}(\Omega)$ désigne la famille de toutes la parties de Ω. Une sous-famille $\mathcal{T}$ de $\mathcal{P}(\Omega)$ est appelée une *tribu* si elle a la propriété que pour tout $A \in \Omega$ et tout $B \in \Omega$ on a aussi $A^c \in \Omega$ et $A \cap B \in \Omega$. Par exemple, si $C \subseteq \Omega$ alors la famille $\mathcal{T}$ des parties $A \subseteq \Omega$ qui ou bien contiennent C ($C \subseteq A$) ou bien évitent C ($C \cap A = \emptyset$) forment une tribu. Bien entendu $\mathcal{P}(\Omega)$ est une tribu, et $\mathcal{T}$ est une *sous-tribu* de $\mathcal{P}(\Omega)$ ce qui veut dire que $\mathcal{T} \subseteq \mathcal{P}(\Omega)$.

Définition : Soient Ω un ensemble fini, $\mathcal{T}$ une sous-tribu de $\mathcal{P}(\Omega)$, $(\Omega, P, \mathcal{T})$ une espace probabilisé fini et soit $\mathbb{T} = \{0, \delta t, \ldots, n\delta t = T\}$, où $\delta t > 0$ est un réel fixé (petit). On appelle *marche aléatoire (finie)* une application (S) mesurable

$$(S) : \Omega \times \mathbb{T} \to \mathbb{R} \ , \ (\omega, t) \mapsto (S)(\omega, t) =: S_t(\omega)$$

Oublions provisoirement l'adjectif “mesurable” dont nous préciserons le sens plus loin et étudions tout d'abord un exemple :

Exemple : Le modèle CRR que nous avons déjà étudié fournit un premier exemple de marche aléatoire : elle modélise la dynamique de l'actif sous-jacent à une option. L'espace probabilisé fini Ω considéré est l'ensemble à 2^n élements $\Omega = \{-1, 1\}^n$; un “état du monde” $\omega \in \Omega$ est une suite de n valeurs ± 1 qui représente la succession des n mouvements vers le haut (up) ou vers le bas (down) de l'actif. En introduisant la probabilité de calcul, nous avons posé, pour un $\omega \in \Omega$ qui comporte j composantes égales à $+1$ (et $(n-j)$ égales à -1),

$$P(\omega) := p^j (1-p)^{n-j},$$

faisant implicitement une hypothèse d'indépendance des accroissements sur laquelle nous reviendrons. La tribu $\mathcal{T}$ est ici simplement la tribu pleine $\mathcal{T} = \mathcal{P}(\Omega)$. La marche aléatoire CRR, notée (S) est définie sur $\Omega \times \mathbb{T}$ par la formule suivante, lorsque $t = i\delta t$, que $\omega = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i, \ldots, \varepsilon_n)$ comporte j composantes d'indice inférieur à i qui sont égales à $+1$:

$$(\omega, t) \mapsto S_t(\omega) : := S_0 u^j d^{i-j}.$$

Il y a deux façons de voir une marche aléatoire (S) et c'est précisément ce qui fait la richesse de cette notion :

1. Si on fixe un $\omega \in \Omega$, l'application $\mathbb{T} \to \mathbb{R}$ qui à $t \in \mathbb{T}$ associe $S_t(\omega)$ est une fonction de t qu'on appelle la *trajectoire* de l'*état du monde* ω et qui représente l'une des évolutions au cours du temps de la quantité modélisée, celle qui correspond à ω. On peut munir chaque trajectoire $t \mapsto S_t(\omega)$ d'une probabilité en posant
$$P(t \mapsto S_t(\omega)) := P(\overline{\omega})$$
où $\overline{\omega}$ est l'ensemble des $\omega \in \Omega$ qui conduisent à la même trajectoire. Une marche aléatoire peut donc être vue comme *un espace probabilisé de trajectoires*. Par exemple dans le modèle CRR à $n = 3$ étapes, l'espace Ω comporte 8 états du monde, $\omega_1 = (+1, +1, +1)$, $\omega_2 = (+1, +1, -1)$, $\omega_3 = (+1, -1, +1)$, $\omega_4 = (+1, -1, -1)$, $\omega_5 = (-1, +1, +1)$, $\omega_6 = (-1, +1, -1)$, $\omega_7 = (-1, -1, +1)$, et $\omega_8 = (-1, -1, -1)$ et il y a 8 trajectoires
2. Si on fixe un $t \in \mathbb{T}$, l'application $\Omega \to \mathbb{R}$ qui à $\omega \in \Omega$ associe $S_t(\omega)$ est une variable aléatoire ; la marche aléatoire définit donc pour chaque t une v.a.. On peut donc aussi voir une marche aléatoire comme *une famille à un paramètre* t de v.a.. Dans l'exemple du modèle CRR à 3 étapes, ces v.a. sont S_0 (qui est la v.a. certaine égale à la constante S_0), $S_{\delta t}$, $S_{2\delta t}$ et $S_{3\delta t}$ dont les lois sont données respectivement par les tableaux suivants :

$S_{\delta t}$	$S_0 d$	$S_0 u$
$P(S_{\delta t} = \cdot)$	$(1-p)$	p

$S_{2\delta t}$	S_0d^2	S_0du	S_0u^2
$P(S_{2\delta t}=\cdot)$	$(1-p)^2$	$2p(1-p)$	p^2

$S_{3\delta t}$	S_0d^3	S_0d^2u	S_0du^2	S_0u^3
$P(S_{3\delta t}=\cdot)$	$(1-p)^3$	$3p(1-p)^2$	$3p^2(1-p)$	p^3

2 Filtration et information

Soient Ω^1,..., Ω^m des parties non vides de Ω. On dit que $\wp := \{\Omega^1,\ldots,\Omega^m\}$ est une *partition* de Ω si et seulement si les Ω^i sont deux-à-deux disjoints et Ω est la réunion des Ω^i.

La relation $\sim$ sur Ω définie par $\omega' \sim \omega$" si et seulement si ω' et ω" appartiennent à un même $\Omega^i \in \wp$ est une relation d'équivalence. Réciproquement, si $\sim$ est une relation d'équivalence quelconque sur Ω, les classes d'équivalences $\overline{\omega} = \{w' \in \Omega \mid \omega' \sim \omega\}$ des éléments de Ω consituent une partition de Ω.

Définition : Lorsque Ω est l'espace probabilisé sous-jacent à une marche aléatoire (S), on peut définir sur Ω, pour chaque $t \in \mathbb{T}$, des partitions, que nous noterons $\wp_t$, associées à la marche aléatoire, par la relation d'équivalence $\overset{t}{\sim}$ suivante :

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

En d'autres termes, deux états du monde sont é*quivalents jusqu'à l'instant* t si les trajectoires qui leurs sont associées coïncident jusqu'à l'instant t.

Exemple : Si la marche aléatoire est le modèle CRR à 3 étapes, on a :
- si $t = 0$, $\wp_0 = \{\Omega\}$.
- si $t = \delta t$, $\wp_{\delta t} = \{\Omega^1 = \{w_1, \omega_2, \omega_3, \omega_4\}, \Omega^2 = \{w_5, \omega_6, \omega_7, \omega_8\}\}$.
- si $t = 2\delta t$, $\wp_{2\delta t} = \{\Omega^{11} = \{w_1, \omega_2\}, \Omega^{12} = \{\omega_3, \omega_4\}, \Omega^{21} = \{w_5, \omega_6\}, \Omega^{22} = \{\omega_7, \omega_8\}\}$.
- si $t = 3\delta t$, $\wp_{3\delta t} = \{\{w_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{w_5\}, \{\omega_6\}, \{\omega_7\}, \{\omega_8\}\}$.

Définition : Soit Ω un ensemble fini et $\mathcal{T}$ une tribu de parties de Ω. On appelle *atome* de $\mathcal{T}$ tout élément de $\mathcal{T}$ qui ne contient pas d'autre élément de $\mathcal{T}$ que lui même et l'ensemble vide.

Exemple : Si la marche aléatoire est le modèle CRR à 3 étapes, on peut associer à chaque partition $\wp_t$, une tribu, notée $\mathcal{F}_t$:
- si $t = 0$, $\mathcal{F}_0 = \{\phi, \Omega\}$.
- si $t = \delta t$, $\mathcal{F}_{\delta t} = \{\phi, \Omega^1, \Omega^2, \Omega\}$.
- si $t = 2\delta t$, $\mathcal{F}_{2\delta t} = \{\phi, \Omega^{11}, \Omega^{12}, \Omega^{21}, \Omega^{22}, \Omega^1, \Omega^2, \Omega^{11} \cup \Omega^2, \ldots, \Omega\}$.
- si $t = 3\delta t$, $\mathcal{F}_{3\delta t} = \mathcal{P}(\Omega)$.

Et on a évidemment $\mathcal{F}_0 \subset \mathcal{F}_{\delta t} \subset \mathcal{F}_{2\delta t} \subset \mathcal{F}_{3\delta t}$. Cette suite croissante de tribus est appelée une *filtration.*

En généralisant l'exemple précédent, on comprend facilement qu'il est possible d'associer à toute marche aléatoire (S) une filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$ définie de la façon suivante : pour un t donné, les atomes de la tribu $\mathcal{F}_t$ sont constitués des états du monde $\omega \in \Omega$ auxquels sont associés des trajectoires qui coïncident jusqu'à l'instant t.

Lorsque $t = 0$, l'information dont on dispose est que l'un des états du monde $\omega \in \Omega$ va se réaliser (mais on ne sait pas lequel). A l'instant $t = \delta t$, l'actif que l'on modélise à fait soit un mouvement vers le haut soit un mouvement vers le bas et donc on sait, ayant pu observer cette évolution, que l'état du monde qui se réalisera appartient à Ω^1 ou bien à Ω^2. Et à l'instant $t = 2\delta t$, on saura qu'il appartient à Ω^{11}, Ω^{12}, Ω^{21} ou Ω^{22}, et ainsi de suite. A chaque nouvelle étape l'information dont on dispose sur l'actif observé augmente et on peut mesurer la finesse de cette information par la partition $\wp_t$ ou bien, ce qui revient au même, par la tribu $\mathcal{F}_t$. La suite de ces tribus, $(\mathcal{F}_t)_{t\in\mathbb{T}}$ représente donc *l'information dont on dispose à la date t en observant le marché.*

3 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 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* $\mathfrak{P}$, la v.a. notée $\mathbb{E}(X/\mathcal{F})$ définie, pour tout $\omega \in \Omega$, par

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

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 associée à $\mathcal{F}$. On a plus précisément :

Proposition 1 *La v.a. $\mathbb{E}(X/\mathcal{F})$ est mesurable par rapport à la tribu $\mathcal{T}$ et de plus si Y désigne cette v.a., $(Y = \mathbb{E}(X/\mathcal{F}))$, alors Y peut être définie comme l'unique v.a. $\mathcal{F}$-mesurable telle que*

$$\forall \Omega_i \in \mathfrak{P} \quad , \quad \mathbb{E}(Y/\Omega_i) = \mathbb{E}(X/\Omega_i). \tag{1}$$

Preuve : Le fait que $\mathbb{E}(X/\mathcal{F})$ soit mesurable par rapport à la tribu $\mathcal{T}$ est clair par définition puisqu'elle est constante sur les atomes de la partition $\mathfrak{P}$.

Montrons qu'elle vérifie l'équation (1) : on a

$$\mathbb{E}(Y/\Omega_i) = \frac{1}{P(\Omega_i)} \sum_{\alpha \in \Omega_i} Y(\alpha) P(\alpha) = \frac{1}{P(\Omega_i)} Y(\omega) \sum_{\omega \in \Omega_i} P(\omega) = Y(\omega)$$

où ω est un élément quelconque de l'atome Ω_i, puisque Y est constante sur les atomes. D'autre part, on a :

$$\mathbb{E}(X/\Omega_i) = \frac{1}{P(\Omega_i)} \sum_{\alpha \in \Omega_i} X(\alpha) P(\alpha) = Y(\omega)$$

pour tout $\omega \in \Omega_i$, par définition de Y. Réciproquement si Y est $\mathcal{F}$-mesurable, la relation $\mathbb{E}(Y/\Omega_i) = \mathbb{E}(X/\Omega_i)$ définit Y uniquement car, pour tout $\omega \in \Omega$, on posera $Y(\omega) := \mathbb{E}(X/\Omega_i)$, où Ω_i est l'atome contenant ω. □

Voici les principales propriétés de l'espérance conditionnelle :

Proposition 2 *Soient X et Y des v.a. sur $(\Omega, \mathcal{T}, P)$, $\mathcal{F}$ et $\mathcal{G}$ des sous-tribus de $\mathcal{T}$, 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.

4 Application au calcul de prix d'options

Soit $(S_t)_{t \in \mathbb{T}}$ une marche aléatoire et soit $(\mathcal{F}_t)_{t \in \mathbb{T}}$ la filtration associée. On a vu ci-dessus comment associer à une v.a. X son espérance par rapport à une tribu, $Y = \mathbb{E}(X/\mathcal{F})$. 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, indéxée 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 lequel 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 $\overline{\omega}_t$, où $\overline{\omega}_t$ 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) 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 de (S) jusqu'à l'instant t, c'est-à-dire connaissant l'information jusqu'à t. Notons que l'on a $\mathbb{E}(\varphi(S_T)/\mathcal{F}_T) = \varphi(S_T)$ et $\mathbb{E}(\varphi(S_T)/\mathcal{F}_0) = \mathbb{E}(\varphi(S_T))$.

Utilisant la notion d'espérance conditionnelle par rapport aux tribus d'une filtration, 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 3 pour $t = 0$.

Proposition 3 *Dans un marché financier $(S_t, B_t)_{t\in\mathbb{T}}$ où l'actif risqué S_t suit un modèle CRR et l'actif sans risque est donné par $B_t := e^{rt}$, le prix d'une option d'échéance T et de fonction de paiement $\varphi(S_T)$ est la marche aléatoire $(C_t)_{t\in\mathbb{T}}$ définie pour tout $t \in \mathbb{T}$ par*

$$C_t = e^{-r(T-t)}\mathbb{E}(\varphi(S_T)/\mathcal{F}_t) \tag{2}$$

où $(\mathcal{F}_t)_{t\in\mathbb{T}}$ est la filtration associée à la m.a. CRR $(S_t)_{t\in\mathbb{T}}$ et où l'espérance conditionnelle est calculée sous la probabilité de calcul p définie par $p = \frac{e^{r\delta t}-d}{u-d}$.

Calcul pratique dans un modèle de Cox-Ross-Rubinstein

Soit $S(i,j)$ la valeur de S_t pour $t = i\delta t$, $\delta t := \frac{T}{n}$, lorsque la trajectoire a connu j "up" jusqu'à l'instant t. Soit $C(i,j)$ la valeur dans ce même cas d'une option européenne de fonction de paiement φ, ce qui signifie que $C(n,j') = \varphi(S(n,j'))$ pour tout $j' = 0\dots n$. Soit $\omega = (\varepsilon_1,\dots,\varepsilon_i,\dots,\varepsilon_n)$ tel que $S_t(\omega) = S(i,j)$, ce qui impose que $j = \sharp\{i' = 1\dots i \mid \varepsilon_i = +1\}$, où $\sharp A$ désigne le nombre d'éléments de l'ensemble A. Observons que nécessairement $S_T(\omega) = S(n,j')$ avec $j' = j+k$, où $k = 0\dots n-i$ désigne le nombre de "up" ayant lieu *après* t. Par la proposition 3 on a donc

$$\begin{aligned} C(i,j) &= e^{-r(T-t)}\mathbb{E}(\varphi(S_T)|\mathcal{F}_t)(\omega) && (3)\\ &= e^{-r(n-i)\delta t}\sum_{k=0}^{n-i}\begin{pmatrix} n-i \\ k \end{pmatrix} p^k(1-p)^{n-i-k}\varphi(S(n,j+k)), && (4)\end{aligned}$$

puisque la probabilité pour que $S_T = S(n,j+k)$ sachant que $S_t = S(i,j)$ est égale à $\begin{pmatrix} n-i \\ k \end{pmatrix} p^k(1-p)^{n-i-k}$, la probabilité[1] d'avoir k "up" durant les $n-i$ pas de temps situés après t. En particulier, pour $i = 0$, on retrouve la formule obtenue au chapitre 3 pour la prime C_0.

[1]On note $\begin{pmatrix} m \\ k \end{pmatrix} = C_m^k = \frac{m\,!}{k\,!(m-k)\,!}$ le nombre de sous ensemble de k éléments choisis dans un ensemble de m éléments.