INITIATION AUX MATHÉMATIQUES FINANCIÈRES

J.Azéma

Février 2004

Introduction

Ces notes , destinées aux étudiants de la maitrise MASS de l'université Paris V , ne prétendent à aucune originalité. Elles ont pour but d'aider les étudiants de ce cursus, possèdant pour certains d'entre eux un bagage mathématique réduit ,à comprendre le début du remarquable petit livre de **Lamberton** et **Lapeyre** intitulé *"Introduction au calcul stochastique appliqué à la finance" Ed.Ellipses (97)* ;la lecture des 40 premières pages de cet ouvrage est , bien entendu , fortement recommandée.

Le problème principal que nous nous poserons est celui de l'évaluation du prix d'une *option (pricing)* ; voici en gros de quoi il s'agit. Et tout d'abord de quoi il ne s'agit pas : on ne cherche en aucune façon à prédire l'évolution du cours d'une action ; ce travail est celui des *analystes financiers* qui auscultent le bilan de la firme , ses avantages ou ses handicaps vis à vis de la concurrence , regardent si elle est positionnée sur un marché porteur , etc...Cela fait, ils prodiguent à leurs clients ou publient dans la presse financière des conseils d'achat ou de vente.
Mais cette même action peut servir de *support* à des options d'achat (*call*) ou de vente (*put*) dont le fonctionnement peut être décrit de la manière suivante.
L'acheteur d'une option achète un droit à son émetteur (en général un grand établissement financier) qu'il pourra ou non *exercer* à une date convenue à l'avance appelée *date d'exercice*. S'il s'agit d'un call, (resp.d'un put), ce droit consiste en la possibilité pour le détenteur de l'option d'acheter, (resp.de vendre), l'action support à son émetteur à un prix, lui aussi déterminé à l'avance, appelé *prix d'exercice*.
Il est à noter

- que l'acheteur de l'option paye immédiatement son dû , à savoir le prix de l'option, à son vendeur ; celui ci dipose donc de liquidités qu'il peut faire fructifier a sa guise jusqu'à la date d'exercice.
- que si le détenteur de l'option peut choisir d'exercer ou non son droit, son vendeur est tenu de respecter le contrat en cas d'exercice de l'option.

Pour se rendre compte du caractère très spéculatif des marchés où se négocient les options , souvent appelés *marchés dérivés*, imaginons l'exemple suivant :
Au 1^{er} Otobbre 2000 ,l'action Michelin étant cotée 200F,la Société Générale émet des options d'achat sur ce support , au prix de 10F , la date d'exercice étant fixée au 1^{er} Octobre 2001 , et le prix d'exercice à 220F ; vous disposez d'une somme de 20.000F et avez toute confiance dans les perspectives de la société Michelin ; vous pouvez manifester cette confiance de 2 façons

- soit acheter 100 actions Michelin
- soit acheter a la Société Générale 2000 call sur Michelin

Si les dieux vous sont favorables, et que l'action gagne 20% en un an, elle cotera 240F à la date d'exercice ; dans le premier cas il ne se passe rien que de très banal : vous avez gagné 4000F ; mais dans le deuxième cas, vous pouvez acheter 2000 actions Michelin à la S.G. qui est tenue de vous les vendre au prix d'exercice (220F), puis les revendre en bourse 240F, ce qui vous procure une plus value de 40.000F ; si l'on déduit de cette somme votre dépense initiale,votre bénéfice est de 20.000F, soit 100%.
Mais ce petit miracle a une contrepartie : si l'action Michelin, contrairement à votre attente, perd 5% au cours de l'année, elle cotera 190F à la date d'exercice ; le droit que vous avez acquis (la possibilité de l'acheter a 220F) n'a plus aucune utilité et donc aucune valeur ; il ne vous reste plus qu'à jeter

votre option à la corbeille et à constater, en gardant le sourire, que vous avez perdu la totalité de votre investissement initial de 20.000F
Le mécanisme de l'option de vente est exactement inverse : c'est un outil de spéculation à la baisse avec un risque de ruine en cas de hausse ; si vous achetez un put doté des memes caractéristiques, et si le cours de l'action support a perdu 20%, elle cote 160F à la date d'exercice ; vous achetez alors sur le marché boursier 2000 actions à 160F que vous revendez 220F à la Société Générale ; je vous laisse calculer votre bénéfice...
Ces options ont une autre fonction ; supposons qu'un fonds de pension américain détienne 30% du capital de Michelin et s'attende à une forte baisse du cours de l'action ; il peut, bien entendu , vendre sa participation avant la catastrophe annoncée, mais, pour des raisons fiscales, ou pour ne pas perdre son droit de regard sur la gestion de l'entreprise, il préfèrera souvent recourir à une stratégie de couverture des risques (*hedging*) fondée sur l'achat d'options de vente. En cas d'effondrement du marché, les gains procurés par la possession de ces put, amplifiés par *l'effet de levier* qui vient d'être décrit, permettront de compenser la perte provoquée par la chute des cours. Plus que le goût de la spéculation prêté aux anglo-saxons, c'est vraisemblablement ce souci de sécurité qui explique le développement spectaculaire des marchés dérivés constaté au cours des 20 dernières années aux Etats Unis.

Il reste à expliquer le rôle que jouent les mathématiques dans cet univers ; une étude statistique des cours de bourse conduit tout d'abord a une estimation de la loi de probabilité qui gouverne leur évolution ; répétons qu'il ne s'agit pas ici de formuler des prédictions ; pour prendre un exemple analogue , il s'agit , après observation d'un grand nombre de jets d'une pièce de monnaie , d'établir que la probabilité de tomber sur pile est (à peu près) 1/2 ; dans ces notes , nous supposerons ce problème résolu à l'aide des techniques classiques de la *statistique des processus* , et nous supposerons donc cette loi connue. La tâche qui incombe alors à l'équipe de probabilistes salariés par la Société Générale est de déterminer à quel prix vendre ses options (les 10F de notre exemple) ; vendues trop chères, personne n'en voudra, vendues trop bon marché , elles peuvent conduire à la faillite des établissements réputés indestructibles , (toujours ce dévastateur *effet de levier*) . Ce dernier point n'est pas une exagération : les faillites retentissantes (Baring , LTCB,etc..) qui ont récemment semé la panique sur les marchés financiers sont presque toutes dues a des spéculations malheureuses sur les marchés dérivés .

Nous nous bornerons, pour l'essentiel , à l'étude du modèle discret de **Cox, Ross et Rubinstein** dans lequel on suppose que le marché financier est réduit a un actif risqué (par exemple une action) et à un actif sans risque (par exemple un carnet d'épargne à taux fixe) . On fait en outre les hypothèses simplificatrices suivantes :

- Le temps est discret (on ne s'intéresse par exemple au cours des actifs qu'à la clôture de la bourse chaque jour a 17H)
- l'actif risqué n'a que 2 comportements possibles : son cours est multiplié chaque jour par une quantité aléatoire qui ne peut prendre que 2 valeurs fixées (par exemple, à chaque clôture ce cours est soit multiplié par 2 soit divisé par 3) .

Nous verrons plus loin les raisons mathématiques qui conduisent à faire cette dernière hypothèse (fort éloignée du comportement réel d'une action !)
En dépit de son caractère simpliste, l'étude de ce modèle est instructive pour plusieurs raisons ; tout d'abord les méthodes employées sont en tout point analogues à celles qui conduisent aux résultats de **Black and Scholes** et ceci dans un cadre élémentaire . On peut déjà y discerner le rôle fondamental joué par la théorie des martingales ainsi que les raisons d'un certain nombre de résultats assez surprenants du modèle de B. and S. , le fait que le prix d'une option dépende uniquement de la volatilité du

marché et non pas de sa tendance (haussière ou baissière) par exemple.

Mais il va nous falloir faire un peu de mathématiques ; au travail !

Chapter 1

Rappels

Notation : Si E est un ensemble fini , le nombre d'éléments de E sera noté $cardE$ ou $\sharp E$ ou plus simplement encore $|E|$

1.1 L'espace des épreuves

Soit Ω un ensemble fini (destiné à représenter l'ensemble des résultats possibles d'une expérience aléatoire) . Les points de Ω seront appelés *épreuves* et les sous ensembles de Ω *événements* (Raisonner sur des ensembles finis nous suffira pour la suite , ce qui évitera de réveiller les traumatismes causés par un apprentissage précoce de la théorie de la mesure)

1.1.1 Exemples

a) L'expérience aléatoire consistant en un jet de 2 dés peut être décrite par l'ensemble $\{1,2,3,4,5,6\}^2$ et l'événement "la somme des deux dés a été de 4", par le sous-ensemble $\{(1,3),(2,2),(3,1)\}$
b) Dans une population E comportant n individus , on prélève au hasard pour un sondage un échantillon de taille p ;on pourra prendre pour espace Ω l'ensemble E^p des suites $(x_1, x_2, ..., x_p)$ de taille p à valeurs dans E ; on a $|\Omega| = n^p$; S'il s'agit d'un sondage pour les élections présidentielles , $(n = 20.000.000, p = 2.000)$, on voit que l'espace des épreuves , bien que restant fini , est gigantesque.
c) Cet exemple , voisin du précédent , peut servir de modèle au jeu de loto .
Soient n et p deux entiers > 0 avec $p < n$, (au loto français , $n = 50$, $p = 6$) ; une épreuve sera un sous ensemble de cardinal p de l'ensemble $E = \{0, 1, 2, \ldots, n-1\}$. L'espace Ω des épreuves est alors de cardinal $\left(\begin{array}{c} n \\ p \end{array}\right)$
d) E est un ensemble fini de cardinal n , p et q deux entiers > 0 inférieurs à n ; on choisit "au hasard" deux sous ensembles F et G de E tels que $|F| = p$ et $|G| = q$; l'ensemble Ω des épreuves est $\{(F,G)|\ |F| = p\ ;\ |G| = q\}$; on a dans ce cas

$$|\Omega| = \left(\begin{array}{c} n \\ p \end{array}\right)\left(\begin{array}{c} n \\ q \end{array}\right)$$

1.1.2 Espaces probabilisés

- Une **probabilité** sur Ω sera une application $\mathbf{P} : \Omega \mapsto [0,1]$ telle que $\sum_{\omega \in \Omega} \mathbf{P}(\omega) = 1$
- La **probabilité d'un événement** A sera définie par $\mathbf{P}A = \sum_{\omega \in A} \mathbf{P}(\omega)$
- Un couple $(\Omega, \mathbf{P})$ sera appelé **espace probabilisé** , ou encore **espace de probabilité**
- l'ensemble des épreuves ayant une probabilité non nulle est appelé **support** de $\mathbf{P}$; une propriété qui est vérifiée en tout point du support de $\mathbf{P}$ est dite vraie **presque sûrement (p.s.)**

Si A et B sont 2 événements , on ne dispose d'aucune formule permettant de calculer $\mathbf{P}(A \cup B)$ ou $\mathbf{P}(A \cap B)$; en revanche , si A et B sont disjoints , on note leur réunion par $A + B$ et l'on a $\mathbf{P}(A + B) = \mathbf{P}A + \mathbf{P}B$.
On dira que $\mathbf{P}$ est la **probabilité uniforme** sur Ω si la fonction $\mathbf{P}$ est constante ; on a alors

$$\mathbf{P}(\omega) = \frac{1}{|\Omega|} \quad \forall \omega \quad ; \quad \mathbf{P}A = \frac{|A|}{|\Omega|} \tag{1.1}$$

1.1.3 Retour aux exemples

Les espaces des épreuves introduits au paragraphe **1.1.1** étant supposés munis de la probabilité uniforme , calculer la probabilité des événements suivants :

a) la somme des deux dés est un multiple de 3
b) Les premier et dernier noms de l'échantillon sont identiques
c) Au jeu du loto tous les nombres tirés sont ≥ 30
d) $F \subset G$; faire le calcul pour $n = 6, p = 2, q = 3$

Réponses : $\frac{1}{3}$; $\frac{1}{n}$; 0,0024391625 ; $\frac{1}{5}$.

1.2 Les variables aléatoires et leurs lois

Comme leurs noms ne l'indiquent pas , les *variables* aléatoires sont en réalité des *fonctions* ; plus précisément :

1.2.1 Définition

On appellera variable aléatoire à valeurs dans un ensemble E une application de Ω dans E

Si X est une v.a. et Γ un sous ensemble de$E$,l'événement $\{\omega \in \Omega \mid X(\omega) \in \Gamma\} = X^{-1}(\Gamma)$ est souvent noté par les probabilistes $[X \in \Gamma]$.
Puisque Ω est fini, l'ensemble $X(\Omega)$ des valeurs que peut prendre X est lui aussi fini .

1.2.2 Retour aux exemples

a) Les variables aléatoires X_1 et X_2 définies par $X_1(i,j) = i$; $X_2(i,j) = j$ représentent les résultats affichés par le premier et le second dé ; une autre v.a.intéressante est $S = X_1 + X_2$; ainsi , $S(3,4) = 7$

b) On définit de la même façon p variables aléatoires à valeurs dans E :
si $\omega = (x_1, x_2 ...., x_i, ..., x_p)$, on posera $X_i(\omega) = x_i \quad (i \in \{1, 2, ..., p\})$. Si l'on imagine l'épreuve ω comme une liste ordonnée de p noms recueillie par un enquêteur, $X_i(\omega)$ représentera le $i - eme$ nom

de cette liste (et n'a aucun intérêt pour le sondage) . Pour rentrer dans le vif du sujet , installons une partition sur E constituée de 2 sous ensembles E_1 et E_2 représentant respectivement l'ensemble des partisans des candidats 1 et 2 ; les 2 variables aléatoires que l'on examinera avec attention seront

$$Z_1(\omega) = \sharp\{i \,| X_i(\omega) \in E_1\} \;\; ; \quad Z_2(\omega) = \sharp\{i \,| X_i(\omega) \in E_2\} \tag{1.2}$$

Elles sont à valeurs dans l'ensemble $\{0, 1, 2, ..., p\}$ et vérifient l'identité $Z_1 + Z_2 = p$

c) Au jeu du loto , la somme des éléments figurant dans le tirage ω

d) $(F, G) \to \sharp(F \cap G)$

1.2.3 Les indicateurs

Soient E un ensemble et F un sous ensemble de E ; l'indicateur de F sera par définition la fonction $1_F : E \to \{0, 1\}$ définie par

$$1_F(x) = 1 \text{ si } x \in F \;\; ; \;\; 1_F(x) = 0 \text{ si } x \in F^c$$

Sur un espace de probabilité , l'indicateur d'un événement A est donc une variable aléatoire á valeursdans $\{0, 1\}$; leur utilisation permet de remplacer les opérations booléennes sur les ensembles (intersection , réunion etc ...) par des calculs algébriques ordinaires . Par exemple ,

$$1_{A\cap B} = 1_A.1_B = 1_A \wedge 1_B \;\; ; \;\; 1_{A\cup B} = 1_A + 1_B - 1_A.1_B = 1_A \vee 1_B (= 1_A + 1_B \; si \; A \cap B = \phi)$$

$$1_{A\cup B\cup C} = 1_A + 1_B + 1_C - (1_A.1_B + 1_A 1_C + 1_B 1_C) + 1_A 1_B 1_C$$

Cela justifie les notations $A \cap B = AB$ et $A \cup B = A + B \; si \; A \cap B = \phi$.

Deux propriétés à retenir

1. Soient X une v.a. à valeurs dans E et Γ un sous ensemble de E ; on a l'égalité

$$1_{[X \in \Gamma]} = 1_\Gamma \circ X \tag{1.3}$$

   Ainsi , dans l'exemple **1.2.2 b)** , définissons les p indicateurs $Y_1, Y_2, ..., Y_p$ par $Y_i = 1_{[X_i \in E_1]} = 1_{E_1} \circ X_i$; on a alors l'égalité $Z_1 = \sum_{i=1}^{i=p} Y_i$.

2. Soit X une v.a. réelle ; on note $(a_1, a_2, ..., a_n)$ une énumération de $X(\Omega)$ où les a_i sont distincts et A_i l'événement $[X = a_i]$. La suite (A_i) constitue une partition de Ω et l'on peut écrire

$$X = \sum_{i=1}^{i=n} a_i 1_{A_i} \tag{1.4}$$

La démonstration de ces 2 points est immédiate : il suffit de vérifier que les fonctions figurant aux 2 membres de (1.3) et de (1.4) prennent la même valeur en tout point ω . Nous invitons le lecteur inexpérimenté à prendre quelques minutes pour écrire la démonstration complète de manière à s'assurer qu'il a assimilé les définitions et notations du début de ce cours .

1.2.4 Espérance d'une variable aléatoire réelle

Definition

Soit X une variable aléatoire à valeurs réelles ;on posera

$$\mathbf{E}[X] = \sum_{\omega \in \Omega} X(\omega)\mathbf{P}(\omega) \tag{1.5}$$

Rappelons quelques propriétés essentielles de l'espérance résultant imméditement de sa définition :

- **Espérance d'un indicateur :** On a $\mathbf{E}[1_A] = \mathbf{P}A$ quelque soit l'événement A
- **Linéarité :** Si X et Y sont 2 v.a.r et si α et β sont 2 nombres réels , on a

$$\mathbf{E}[\alpha X + \beta Y] = \alpha \mathbf{E}[X] + \beta \mathbf{E}[Y]$$

- **Positivité :** Si X est positive , son espérance est positive et n'est nulle que si X est presque sûrement nulle
- **Inegalité de convexité :** Si $\phi : \mathbb{R} \mapsto \mathbb{R}$ est une fonction convexe , alors $\mathbf{E}[\phi(X)] \geq \phi(\mathbf{E}[X])$
- **Exemple :** Retournons encore une fois à l'exemple **b)** ; compte tenu de la remarque faite à la fin du paragraphe précédent , on a

$$\mathbf{E}[Z_1] = \sum_1^p \mathbf{E}[Y_i] = \sum_1^p \mathbf{P}[X_i \in E_1] = \sum_1^p \frac{n_1}{n} = p\frac{n_1}{n}$$

  quand on a posé $\sharp E_1 = n_1 \;\; , \;\; \sharp E_2 = n_2$
- **Exercice :** Soient A, B, C trois événements d'un espace de probabilité Ω ; montrer l'égalité

$$\mathbf{P}[A \cup B \cup C] = \mathbf{P}A + \mathbf{P}B + \mathbf{P}C - (\mathbf{P}[A \cap B] + \mathbf{P}[A \cap C] + \mathbf{P}[B \cap C]) + \mathbf{P}[A \cap B \cap C]$$

  Généraliser à un nombre quelconque d'événements .

1.2.5 Loi d'une variable aléatoire

Soit X une variable aléatoire à valeurs dans E ; nous appellerons loi (on dit aussi souvent distribution) de X la probabilité Π_X définie sur l'ensemble fini $X(\Omega)$ par

$$\Pi_X(k) = \mathbf{P}[X = k] \quad \forall k \in X(\Omega)$$

Deux v.a. admettant la même loi seront dites équidistribuées

Regardons quelles sont les lois des v.a. que nous avons rencontrées

1. **loi d'un indicateur ; loi de Bernoulli** . Soit A un événement de probabilité α ; la v.a. 1_A ne prend que les deux valeurs 0 et 1 , de sorte que sa loi Π est la probabilité sur$\{0, 1\}$ vérifiant $\Pi(0) = 1 - \alpha \;\; , et \; \Pi(1) = \alpha$

   La probabilité Π que l'on vient de définir sur $\{0, 1\}$ s'appelle *loi de Bernoulli de paramètre α*

   Dans l'exemple **b)**, les variables aléatoires Y_i définies à la fin du paragraphe **1.2.3.** sont équidistribuées : chacune d'entre elles suit une loi de Bernoulli de paramètre $\frac{n_1}{n}$

2. **loi binomiale** . Restons dans l'exemple **b)** et calculons la loi Π de la v.a. Z_1 introduite en **1.2.2** ; notons tout d'abord que Z_1 prend ses valeurs dans l'ensemble $\{0, 1, 2, ..., p\}$; si k est un entier $\leq p$, l'événement $[Z_1 = k]$ est l'ensemble des suites de taille p dont k termes appartiennent à E_1 ; il contient donc $\frac{p!}{k!(p-k)!}{n_1}^k {n_2}^{p-k}$ éléments ,ce qui nous mène à l'égalité

$$\Pi(k) = \mathbf{P}[Z_1 = k] = \frac{p!}{k!(p-k)!}(\frac{n_1}{n})^k(\frac{n_2}{n})^{p-k}$$

On appellera *loi binomiale de paramètres* (α, p) $\quad$ $(\alpha \in]0,1[;\quad p \in \mathbb{N})$ *, la probabilite* Π *sur* $\{0, 1, 2, ..., p\}$ *déterminée par*

$$\Pi(k) = \frac{p!}{k!(p-k)!}\alpha^k(1-\alpha)^{p-k}$$

Ainsi Z_1 (resp. Z_2) suit une loi binomiale de paramètres $(\frac{n_1}{n}, p)$, (resp. $(\frac{n_2}{n}, p)$) .

3. **loi hypergéométrique** Déterminons la loi de la v.a. $X(F,G) = \sharp(F \cap G)$ de l'exemple **d)** : X prend ses valeurs dans l'intervalle $I = [(p+q-n)^+,\ \ p \wedge q]$; si k appartient à cet intervalle , on a

$$\sharp\{(F,G) | X(F,G) = k\} = \begin{pmatrix} n \\ p \end{pmatrix} \begin{pmatrix} p \\ k \end{pmatrix} \begin{pmatrix} n-p \\ q-k \end{pmatrix},$$

de sorte que

$$\mathbf{P}[X = k] = H(n,p,q,k) = \frac{\begin{pmatrix} p \\ k \end{pmatrix} \begin{pmatrix} n-p \\ q-k \end{pmatrix}}{\begin{pmatrix} n \\ q \end{pmatrix}} \tag{1.6}$$

La probabilité Π portée par I donnée par $\Pi(k) = H(n,p,q,k)$ est appelée loi hypergéométrique de paramètres n, p, q .

4. **Exercices**
   (a) Vérifier par un calcul purement algébrique que $\sum_k H(n,p,q,k) = 1$; (on pourra utiliser l'identité $(1+x)^p(1+x)^{n-p} = (1+x)^n$ et égaler les coefficients de x^q figurant dans les 2 membres) .
   (b) Dans l'exemple **d)** , calculer $\mathbf{P}[F \subset G]$ et $\mathbf{P}[F \cap G = \phi]$.
   (c) Au jeu du loto , (exmple **c)**) , calculer la probabilité d'avoir k numéros gagnants .

La connaissance de la loi d'une variable aléatoire permet de calculer son espérance , ses moments , sa fontion génératrice etc... ; plus généralement , on a le résultat suivant :

Proposition :

Soient X une v.a. à valeurs dans E et ϕ une application de E dans $\mathbb{R}$; $\phi(X)$ est alors une variable aléatoire réelle dont l'espérance est donnée par la formule :

$$\mathbf{E}[\phi(X)] = \sum_{k \in X(\Omega)} \phi(k)\Pi_X(k). \tag{1.7}$$

Démonstration : Ce résultat découle immédiatement de l'égalité (à peu près évidente , si l'on a réfléchi à (1.4))

$$\Phi(X) = \sum_{k \in X(\Omega)} \Phi(k) 1_{\{X=k\}}$$

et de la linéarité de l'espérance

Dans le cas où X est elle réelle , on a en particulier

$$\mathbf{E}[X] = \sum_{k \in X(\Omega)} k \Pi_X(k) \quad , \quad \mathbf{E}[X^r] = \sum_{k \in X(\Omega)} k^r \Pi_X(k) \quad , \quad \mathbf{E}[s^X] = \sum_{k \in X(\Omega)} s^k \Pi_X(k) \qquad (1.8)$$

Par exemple , si X suit une loi binomiale de paramètres (α, p) , on a

$$\mathbf{E}[X] = \sum_{k=0}^{k=p} k \begin{pmatrix} p \\ k \end{pmatrix} \alpha^k (1-\alpha)^{p-k}.$$

Le calcul du second membre est un exercice classique sur la formule du binôme (développer $(x+a)^p$, dériver , puis faire $a = 1 - x$ dans l'égalité obtenue) ; on obtient $\mathbf{E}[X] = p\alpha$; on retrouve ainsi d'une autre façon la valeur trouvée en **1.2.4** pour l'espérance de Z_1 (dans ce cas , $\alpha = \frac{n_1}{n}$)

Exercices

1. Calculer , dans le même esprit , le moment d'ordre 2 et la variance de Z_1
2. Dans l'exemple **d)** , montrer que $\mathbf{E}[\ \sharp(F \cap G)] = \frac{pq}{n}$.
   Indications : Il s'agit de calculer (Cf (1.6)) , $\sum_k k\ H(n,p,q,k)$; après avoir montré l'identité $p(1+x)^{p-1} = \sum_0^{p-1} (k+1) \begin{pmatrix} p \\ k+1 \end{pmatrix} x^k$, on identifiera les termes en x^{q-1} dans l'égalité entre polynômes $(1+x)^{p-1}(1+x)^{n-p} = (1+x)^{n-1}$

1.2.6 Loi d'un couple de variables aléatoires

Soient $X_1 : \Omega \mapsto E_1$ et $X_2 : \Omega \mapsto E_2$ deux variables aléatoires définies sur le même espace de probabilité $(\Omega, \mathbf{P})$.

Posons , pour tout $\omega \in \Omega$, $X(\omega) = (X_1(\omega), X_2(\omega))$; X est alors une variable aléatoire à valeurs dans l'espace produit $E = E_1 \times E_2$, que l'on appellera le couple (X_1, X_2) . Posons $X_1(\Omega) = F_1$, $X_2(\Omega) = F_2$; on peut appliquer les définitions et les résultats du paragraphe précédent à la variable aléatoire X ; il en résulte que :

- La loi Π_X du couple (X_1, X_2) est une probabilité sur l'ensemble fini $F_1 \times F_2$
- Soit $f : E_1 \times E_2 \mapsto \mathbb{R}$; $f(X_1, X_2)$ est alors une variable aléatoire réelle et

$$\mathbf{E}[f(X_1, X_2)] = \sum_{x_1 \in F_1 ; x_2 \in F_2} f(x_1, x_2) \Pi_X(x_1, x_2)$$

avec les cas particuliers suivants , obtenus en choisissant convenablement la fontion f ,

– Si $A_1 \subset F_1$,

$$\mathbf{P}[X_1 \in A_1] = \Pi_{X_1} A_1 = \Pi_X(A_1 \times F_2) = \sum_{x_1 \in A_1; x_2 \in F_2} \Pi_X(x_1, x_2)$$

– Soit $\phi_1 : E_1 \mapsto \mathbb{R}$; on a

$$\mathbf{E}[\phi_1(X_1)] = \sum_{x_1 \in F_1; x_2 \in F_2} \phi_1(x_1) \Pi_X(x_1, x_2)$$

Les loi de chacune des deux variables $X_1 et X_2$ sont appelées *lois marginales* du couple ; il resulte des 2 points qui précèdent qu'on les déduit aisément de la loi Π_X ; par exemple ,

$$\forall x_1 \in F_1 \quad , \quad \Pi_{X_1}(x_1) = \sum_{x_2 \in F_2} \Pi_X(x_1, x_2)$$

Inversement , on ne peut pas déduire la loi d'un couple de la seule connaissance de ses lois marginales. Pour s'en convaincre , on réfléchira à l'exercice suivant :

Exercice

On pose $\Omega = \{1, 2, 3, 4, 5, 6\}^2$, $X_1(\omega) = i$, $X_2(\omega) = j \;\; si \;\; \omega = (i, j)$. On munit Ω de 2 probabilites : $\mathbf{P}$ est la probabilité uniforme sur Ω , tandis que $\mathbf{P}$' est la probabilité uniforme sur la diagonale de Ω

1) Comparer les lois marginales du couple (X_1, X_2) selon que Ω est muni de $\mathbf{P}$ ou $\mathbf{P}$'

2) Comparer $\mathbf{E}[X_1^2 + X_2^2]$ et $\mathbf{E}'[X_1^2 + X_2^2]$; pouvait-on prévoir ce résultat sans calcul ?

3)Comparer $\mathbf{E}[(X_1 + X_2)^2]$ et $\mathbf{E}'[(X_1 + X_2)^2]$

$\mathbf{P}$' sert de modèle à une expérience parfaitement idiote qui comporte 2 résultats mais où on ne lance qu'un dé ; on prend ensuite comme 2^{ieme} résultat , la copie du premier ; c'est évidemment une expérience très différente du jet de 2 dés qui est modélisé par $\mathbf{P}$; les considérer sous cet angle expérimental permet de comprendre pourquoi ces 2 modèles donnent naissance aux mêmes lois marginales.

La notion de $n - uplet$ $(X_1, X_2, ..., X_n)$ **de variables aléatoires** est une généralisation facile de la notion de couple ; les adaptations qui s'imposent sont laissées au lecteur .

Un exemple : la loi multinomiale . Sans le dire , nous avons supposé que notre sondage préélectoral concernait le second tour des élections présidentielles (il ne restait que 2 candidats) ; on peut, sans changer d'espace de probabilité, revenir 15 jours en arrière et s'intéresser au premier tour , à ses q candidats ,et surtout à leurs électeurs qui forment une partition $E_1, E_2, ...E_q$ de la population E ; on posera $n_j = cardE_j \quad (1 \leq j \leq q) \qquad (2 \leq q \leq n)$ et l'on se rappellera que $n_1 + n_2 + ... + n_q = n$. On pose ensuite

$$Z_j = card\{i \,| X_i \in E_j\} \quad (1 \leq j \leq q)$$

($Z_j(\omega)$ représente le nombre d'électeurs du candidat j figurant sur la liste ω de p électeurs ramenée par notre enquêteur) . Notons au passage que $Z_1 + Z_2 + ...Z_q = p$, et posons $Z = (Z_1, Z_2, ..., Z_q)$.

Le calcul des lois marginales de Z est facile et sera laissé au lecteur (la loi de Z_j , est binomiale de paramètres $(\frac{n_j}{n}, p))$; intéressons nous maintenant à la loi de Z , c'est à dire aux quantités

$$\Pi_Z(k_1, k_2, ..., k_q) = \mathbf{P}[Z_1 = k_1; Z_2 = k_2; ....; Z_q = k_q]$$

quand $k_1 + k_2 + ...k_q = p$.
Nous avons donc à calculer le cardinal N de l'ensemble des suites de taille p qui ont k_1 termes dans E_1 ,k_2 termes dans E_2 etc...Pour faire cela , on remarque d'abord que le nombre de suites dont les k_1 *premiers* termes sont dans E_1 , les k_2 *suivants* dans E_2,...,et les k_q *derniers* dans E_q est égal à $n_1^{k_1} n_2^{k_2} ....n_q^{k_q}$; N s'obtient en multipliant cette quantité par le nombre P de partitions de $\{1, 2, ..., p\}$ en sous ensembles de cardinaux $k_1, k_2, ..., k_q$; Or , P est égal à

$$\begin{pmatrix} p \\ k_1 \end{pmatrix} \begin{pmatrix} p - k_1 \\ k_2 \end{pmatrix} \begin{pmatrix} p - k_1 - k_2 \\ k_3 \end{pmatrix} .... \begin{pmatrix} p - k_1 - k_2 - ... - k_{q-1} \\ k_q \end{pmatrix} = \frac{p!}{k_1!k_2!....k_q!}.$$

On obient donc finalement :

$$\Pi_Z(k_1, k_2, ..., k_q) = \frac{N}{n^p} = \frac{p!}{k_1!k_2!....k_q!} (\frac{n_1}{n})^{k_1} (\frac{n_2}{n})^{k_2} .... (\frac{n_q}{n})^{k_q} 1_{[k_1 + k_2 + ... + k_q = p]}$$

Cette loi de probabilité sur $F = \{0, 1, 2, ..., p\}^q$ est appelée loi multinomiale de paramètres $(\frac{n_1}{n}, \frac{n_2}{n}, ..., \frac{n_q}{n}, p, q)$. On remarquera que son support est le sous ensemble strict de F :

$$\{k \in F \mid k_1 + k_2 + ... + k_q = p\}$$

1.3 Indépendance

Rappelons brièvement les définitions essentielles et quelques résultats faciles dont nous ne donnerons pas les démonstrations

Indépendance d'une famille d'événements :

Une famille finie $(A_1, A_2, ..., A_n)$ d'événements est dite indépendante si $\forall J \subset \{1, 2, ..., n\}$ on a

$$\mathbf{P}[\bigcap_{j \in J} A_j] = \prod_{j \in J} \mathbf{P} A_j$$

Indépendance de deux classes :

Soient $\mathcal{C}$ et $\mathcal{D}$ deux classes d'événements; on dira que $\mathcal{C}$ et $\mathcal{D}$ sont indépendantes si

$$\forall C \in \mathcal{C} \ \ \forall D \in \mathcal{D} \quad \mathbf{P}[C \cap D] = \mathbf{P} C \mathbf{P} D.$$

Une v.a. Y sera dite indépendante d'une classe $\mathcal{D}$ si les classes $\{Y \in \Gamma; \Gamma \subset E\}$ *et* $\mathcal{D}$ sont indépendantes .

Indépendance d'une famille de variables aléatoires

Soient $(X_i : \Omega \mapsto E_i)$ une famille finie de variables aléatoires ; on dit que cette famille est indépendante si ,
$\forall \Gamma_1 \subset E_1, \forall \Gamma_2 \subset E_2 ..., \forall \Gamma_n \subset E_n$,

$$\mathbf{P}[X_1 \in \Gamma_1; X_2 \in \Gamma_2; ...X_n \in \Gamma_n] = \mathbf{P}[X_1 \in \Gamma_1]\mathbf{P}[X_2 \in \Gamma_2]...\mathbf{P}[X_n \in \Gamma_n]$$

Pour qu'il en soit ainsi,il suffit que les égalités précédentes soient vérifiées quand les sous ensembles Γ_i sont reduits à des points .

Si $f_i : E_i \mapsto F_i$ sont n applications , et si $(X_1, X_2, ...X_n)$constitue une famille indépendante , il en est de même de la famille $(f_1(X_1), f_2(X_2), ..., f_n(X_n))$

Indépendance et produit de lois :

Soient (F_1, Π_1) et (F_2, Π_2) 2 espaces probabilisés ; on appellera *probabilité produit* de Π_1 etΠ_2 , que l'on notera $\Pi_1 \times \Pi_2$, la probabilité Π sur $F_1 \times F_2$ definie par $\Pi(x_1, x_2) = \Pi_1(x_1)\Pi_2(x_2)$. On a alors , si $\Gamma_1 et \Gamma_2$ désignent respectivement 2 sous ensembles de $F_1 et F_2$, $\Pi(\Gamma_1 \times \Gamma_2) = \Pi_1(\Gamma_1) \times \Pi_2(\Gamma_2)$
Si , maintenant , $X = (X_1, X_2)$ est un couple de variables aléatoires , $X_1 et X_2$ sont indépendantes si et seulement si $\Pi_X = \Pi_{X_1} \times \Pi_{X_2}$.
Vous constatez donc , que dans le cas d'un couple de v.a. indépendantes , les lois marginales déterminent la loi du couple ; mais toutes les v.a. ne sont pas indépendantes ...

Existence et unicité en loi d'une suite de variables aléatoires indépendantes équidistribuées.

Soient E un ensemble fini et μ une loi de probabilité sur E ; la proposition qui suit a pour objet la construction d' une suite $(X_1, X_2, ..., X_n)$ de variables aléatoires indépendantes et équidistribuées de loi μ . A cet effet , posons

- $\Omega = E^n$
- $X_i(\omega) = x_i$ si $\omega = (x_1, x_2, ..., x_i, ..., x_n)$
- $\mathbf{P} = \mu \times \mu \times ... \times \mu = \mu^n$

On a alors le résultat suivant dont la démonstration , facile , est laissée au lecteur .

Proposition

1. $\mathbf{P}$ est l'unique probabilité sur Ω sous laquelle les v.a. $X_1, X_2, ..., X_n$ sont indépendantes de loi μ
2. $\mathbf{P}$ est la probabilité uniforme sur Ω si et seulement si μ est la loi uniforme sur E .

Exemple : une autre approche de la loi binomiale

Soit$(A_1, A_2, \ldots, A_p)$ une suite de p événements indépendants de même probabilité α ; on pose $Z = \sum_{i=1}^{i=p} 1_{A_i}$. Nous allons voir que Z suit une loi binomiale de paramètres α, p . En vertu de l'indépendance des A_i , on a en effet , si I est une partie de $\{1, 2, ..., p\}$ de cardinal k ,

$$\mathbf{P}\{\bigcap_{i \in I} A_i \bigcap_{i \in I^c} A_i^c\} = \alpha^k (1 - \alpha)^{p-k} \;\; , \;\; \mathbf{P}[Z = k] = \begin{pmatrix} p \\ k \end{pmatrix} \alpha^k (1 - \alpha)^{p-k}$$

Le fait que Z admette même loi que la v.a. Z_1 du sondage (exemple **b)**), n'est pas dû au hasard. Dans ce modèle, en effet, $\mathbf{P}$ est la loi uniforme sur $\Omega = E^p$ et les variables aléatoires X_i sont les applications coordonnées ; d'après la proposition précédente , ces v.a. sont donc indépendantes et suivent une loi uniforme sur E. Les événements $[X_i \in E_1]$ sont donc indépendants de probabilité $\frac{n_1}{n}$, ce qui permet d'appliquer la formule ci-dessus à la v.a. $Z_1 = \sum_i 1_{[X_i \in E_1]}$.

Espérance d'un produit de variables aléatoires indépendantes :

Soient X et Y sont 2 variables aléatoires indépendantes à valeurs dans E , f et g deux fonctions numériques sur E alors

$$\mathbf{E}[f(X)g(Y)] = \mathbf{E}[f(X)]\mathbf{E}[g(Y)].$$

En particulier , si X et Y sont réelles , $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$.
Après avoir remarqué qu'il suffit de montrer la deuxième assertion , posons $Z = (X, Y)$; on a

$$\mathbf{E}[XY] = \sum_{x,y} xy\Pi_Z(x,y) = \sum_{x,y} xy\Pi_X(x)\Pi_Y(y) = (\sum_x x\Pi_x(x))(\sum_y y\Pi_Y)(y) = \mathbf{E}[X]\mathbf{E}[Y]$$

On notera que si l'une des 2 v.a. est centrée , $\mathbf{E}[XY] = 0$

Variance d'une somme de v.a.r. indépendantes

Soit $(X_1, X_2, ..., X_n)$ une famille de n v.a.r. indépendantes ; on pose $S = X_1 + X_2 + ... + X_n$.
On a alors

$$Var(S) = Var(X_1) + Var(X_2) + ... + Var(X_n).$$

On se ramène en effet aisément au cas où les v.a. X_i sont centrées ;la fin du calcul va ensuite de soi , compte tenu du point précédent . L'hypothèse d'indépendance est ici essentielle .

Exercice :

En déduire un nouveau mode de calcul de la variance d'une loi binomiale .

Chapter 2

Conditionnement

Une compréhension approfondie de ce chapitre est indispensable pour la suite ; nous commencerons par un rappel des notions élémentaires .

2.1 Conditionnement par un événement

Soient A et B deux événements d'un espace probabilisé $(\Omega, \mathbf{P})$ avec $\mathbf{P}B > 0$; La probabilité conditionnelle $\mathbf{P}[A|B]$ de A quand B (on dit aussi "sachant B ") sera , par définition

$$\mathbf{P}[A|B] = \frac{\mathbf{P}AB}{\mathbf{P}B}$$

On a évidemment les égalités

$$\mathbf{P}(A + A'|B) = \mathbf{P}(A|B) + \mathbf{P}(A'|B) \ ; \ \mathbf{P}[\Omega|B] = 1 \ ; \ \mathbf{P}[\phi|B] = 0.$$

La fonction d'ensemble $\mathbf{P}_B$ définie par $\mathbf{P}_B(A) = \frac{\mathbf{P}AB}{\mathbf{P}B}$ est donc une nouvelle probabilité sur Ω .
Les deux propriétés suivantes sont conformes à l'idée intuitive que l'on se fait d'un conditionnement .

- $\mathbf{P}_B[B] = 1$; (la probabilité qu'il pleuve , sachant qu'il pleut , est égale à 1) .
- Les événements invariants par ce changement de probabilité sont exactement les événements indépendants de B ; (la probabilité de A : *il va pleuvoir* est modifiée par la connaissance de l'événement : *le temps est couvert* , mais ne l'est pas par celle de l'événement indépendant : *j'ai un brelan servi*) .

Exercices :

1. On dispose de 3 cartes ainsi constituées : la première est blanche sur ses 2 faces , la seconde noire sur ses 2 faces et la troisième possède une face blanche et une face noire ; un ami tire au hasard l'une d'elles derrière son dos et vous montre une de ses faces : elle est blanche ; quelle est la probabilité que l'autre face soit blanche ?
2. Dans une population de un million d'individus , on compte mille personnes atteintes d'une certaine maladie . On dispose pour la déceler d'un test qui n'est pas totalement fiable ; plus précisément :

– appliqué à une personne malade , il donne un résultat positif avec une probabilité de $0,9$
– appliqué à un individu bien portant , il est positif avec la probabilité $0,1$
On tire au hasard un individu dans cette population ; il est testé positif ; quelle est la probabilité qu'il soit malade ? (réponse : $1/112$)

3. L'animateur d'un jeu télévisé américain est posté devant trois portes closes dissimulant 2 chèvres et une cadillac ; un joueur , évidemment désireux de gagner la cadillac , choisit au hasard une des 3 portes . L'animateur , qui sait où se trouve l'automobile , ouvre alors une des 2 portes non choisies et fait en sorte de montrer une chèvre ; il indique alors au candidat qu'il peut encore modifier son choix s'il le désire . Celui-ci a-t-il intérêt à le faire ? (réponse : oui ; s'il désigne l'autre porte , il a 2 fois plus de chances de gagner la voiture) .

2.1.1 Loi conditionnelle d'une variable aléatoire

Soit X une variable aléatoire à valeurs dans E ; la *loi conditionnelle* $\Pi_{X/B}$ de X quand B sera , par définition , la loi de X quand Ω est muni de la probabilité $\mathbf{P}_B$; on a donc

$$\Pi_{X/B}(x) = \mathbf{P}_B[X = x] = \mathbf{P}[X = x \,|B] = \frac{\mathbf{P}[X = x \,; B]}{\mathbf{P}B} \qquad \forall x \in X(\Omega)$$

Montrons sur un exemple que le conditionnement par un événement modifie la loi d'une v.a.

Exemple

On joue n fois à un jeu de pile ou face biaisé ; le $k^{\text{ième}}$ coup est représenté par une v.a. X_k de Bernoulli vérifiant $\mathbf{P}[X_k = 1] = \alpha$; la suite de v.a. $(X_k \quad 1 \leq k \leq n)$ est supposée indépendante. Si $p \leq n$, $S_p = \sum_{k=1}^{p} X_k$ représente le nombre de piles tombés au cours des p premiers coups ; nous savons que la loi de S_p est binomiale de paramètres (α, p) .
Donnons nous maintenant un autre entier $q \leq n$ et examinons la loi de S_p quand $[S_n = q]$. On a

$$\mathbf{P}[S_p = k \mid S_n = q] = \frac{\mathbf{P}[S_p = k \,; S_n = q]}{\mathbf{P}[S_n = q]} = \frac{\mathbf{P}[S_p = k \,;\; S_n - S_p = q - k]}{\mathbf{P}[S_n = q]} = \frac{\mathbf{P}[S_p = k]\mathbf{P}[S_{n-p} = q - k]}{\mathbf{P}[S_n = q]}$$

La dernière égalité provenant de ce que $S_n - S_p$ est indépendante de S_p et a même loi que S_{n-p} . Un calcul facile montre alors que

$$\mathbf{P}[S_p = k \mid S_n = q] = \frac{\binom{p}{k}\binom{n-p}{q-k}}{\binom{n}{q}} \qquad ((p + q - n)_+ \leq k \leq p \wedge q) \tag{2.1}$$

En particulier , pour $p = 1$,

$$\mathbf{P}[X_1 = 1 \mid S_n = q] = \mathbf{P}[S_1 = 1 \mid S_n = q] = \frac{q}{n} \quad , \quad \mathbf{P}[X_1 = 0 \mid S_n = q] = 1 - \frac{q}{n} \tag{2.2}$$

Dans le conditionnement par l'événement $[S_n = q]$ la loi binomiale de S_p est devenue la loi hypergéométrique rencontrée au paragraphe (**1.2.5.**) ; celle de X_1 est quant à elle restée une loi de Bernoulli mais avec un paramètre différent . On remarquera que les lois conditionnelles trouvées ne dépendent plus de α

2.1.2 Exercice : Conditionnement d'une loi multinomiale

Soient p et q deux entiers > 0 *et* $\alpha_1, \alpha_2, ..., \alpha_q$
q *reels* > 0 ; on pose $H = \{k \in \{0, 1, , ..., p\}^q \mid k_0 + k_1 + k_2 + ... + k_q = p\}$; On dit qu'une v.a. $Z = (Z_1, Z_2, ...Z_q)$ à valeurs dans $\{0, 1, 2, ..., p\}^q$ est multinomiale de parametres $(\alpha_1, \alpha_2, ..., \alpha_q, p, q)$ (cf **1.2.6**) si elle admet pour loi la probabilité sur $\{0, 1, ..., p\}^q$, portée par H définie par

$$\mathbf{P}[Z_1 = k_1; Z_2 = k_2; ...Z_q = k_q] = \frac{p!}{k_1!k_2!...k_q!}\alpha_1^{k_1}\alpha_2^{k_2}...\alpha_q^{k_q} \quad \forall k \in H$$

Soit Z une telle variable aléatoire ; montrer que la loi conditionnelle de $(Z_1, Z_2, ..., Z_{q-1})$ quand $[Z_q = j]$ est la loi multinomiale de paramètres $(\frac{n_1}{n-n_q}, \frac{n_2}{n-n_q}, ..., \frac{n_{q-1}}{n-n_q} \quad , p-j \quad , q-1)$.

2.1.3 Espérance conditionnelle d'une variable aléatoire réelle

Soient X une variable aléatoire réelle et B un événement de probabilité non nulle ; l' espérance de X relativement à la probabilité $\mathbf{P}_B$ (introduite au début de ce chapitre) , sera notée $\mathbf{E}_B[X]$ ou $\mathbf{E}[X|B]$ et sera appelée espérance conditionnelle de X quand B . L'égalité

$$\mathbf{E}[X|B] = \frac{\mathbf{E}[X1_B]}{\mathbf{P}B} \tag{2.3}$$

se démontre aisément si l'on a retenu (1.4) dont nous reprenons les notations ; on a en effet

$$\mathbf{E}_B[X] = \sum_i a_i \mathbf{P}_B A_i = \frac{1}{\mathbf{P}B}\sum_i a_i \mathbf{E}[1_{A_i} 1_B] = \frac{1}{\mathbf{P}B}\mathbf{E}[1_B \sum_i a_i 1_{A_i}] = \frac{1}{\mathbf{P}B}\mathbf{E}[X1_B] \; .$$

Le calcul pratique de l'espérance conditionnelle d'une variable aléatoire s'effectue à l'aide de sa loi conditionnelle ; appliquant (1.7) à la probabilité $\mathbf{P}_B$, on obtient en effet

$$\mathbf{E}[X|B] = \sum_{k \in X(\Omega)} k \Pi_{X/B}(k) \tag{2.4}$$

Exemples :

Revenons aux lois conditionnelles rencontrées au paragraphe précédent données par les formules (2.1) et (2.2) ; on obtient en appliquant (2.4)

$$\mathbf{E}[S_p \mid S_n = q] = \frac{p\,q}{n} \quad , \quad \mathbf{E}[X_1 \mid S_n = q] = \frac{q}{n} \tag{2.5}$$

(Voir à ce sujet l'exercice 2 du paragraphe **1.2.5**)

2.1.4 Exercice :

Dans l'exemple **a)** (jet de 2 dés) , déterminer la loi conditionnelle de X_1 quand $[S = q]$
$(2 \leq q \leq 12 \;\; , \;\; S = X_1 + X_2)$; en déduire que $\mathbf{E}[X_1 \, |S = q] = \frac{q}{2}$

2.1.5 Conditionnement et indépendance

On dira que X est indépendante de B si les v.a. X et 1_B sont indépendantes ; pour qu'il en soit ainsi, il faut et il suffit que $[X = k]$ soit indépendant de $B \quad \forall k \in X(\Omega)$ ce qui s'écrit encore

$$\mathbf{P}[X = k \mid B] = \mathbf{P}[X = k] \quad \forall k.$$

L'égalité

$$\Pi_{X/B} = \Pi_X \tag{2.6}$$

est donc une condition nécessaire et suffisante assurant l'indépendance de X et B. On notera que (2.6) entraîne

$$\mathbf{E}[X|B] = \mathbf{E}[X] \tag{2.7}$$

2.1.6 Exercice :

On lance 8 fois un dé ; les résultats successifs sont notés $X_1, X_2, ..., X_8$, et l'on pose $S = \sum_1^8 X_i$; montrer que

$$\mathbf{E}[S \mid X_1 + X_2 = k] = k + 21. \tag{2.8}$$

2.1.7 Exercice

Soient X_1, X_2, X_3 trois v.a.r. indépendantes , $S = X_1 + X_2 + X_3$; on appelle μ la loi de X_3 ; montrer que la loi conditionnelle de S quand $[X_1 = i; X_2 = j]$ est donnée par

$$\mathbf{P}[S = k \mid X_1 = i\,; X_2 = j] = \mu(k - i - j); \tag{2.9}$$

en déduire l'égalité

$$\mathbf{E}[S \mid X_1 = i\,;\; X_2 = j] = i + j + \mathbf{E}[X_3]. \tag{2.10}$$

Montrer (2.10) directement (i.e. sans calculer la loi conditionnelle de S) .

2.1.8 Exercice :

On suppose l'égalité (2.7) vérfiée pour une v.a.r. prenant 2 valeurs ; montrer que X est alors indépendante de B . Trouver un contre-exemple montrant que cela n'est plus nécessairement le cas quand X prend plus de 2 valeurs .

2.2 Conditionnement par une variable aléatoire

2.2.1 Une notation condensée

Au paragraphe précédent, nous avons souvent été amenés à considérer des expressions de la forme $\mathbf{E}[X|Z = i]$ où X est une v.a. réelle et Z une v.a. quelconque à valeurs dans un ensemble fini F . En général le résultat dépend de i, ce qui nous amène à définir une fonction $\phi : F \to \mathbb{R}$ par

$$\mathbf{E}[X|Z = i] = \phi(i) \tag{2.11}$$

(On posera, par convention, $\phi(i) = 0$ si $\mathbf{P}[Z = i] = 0$).
Il est alors naturel de poser, de manière plus ramassée,

$$\mathbf{E}[X|Z] = \phi(Z) \tag{2.12}$$

Puisque le second membre de (2.12) est une v.a. qui vaut $\phi(i)$, (c'està dire $\mathbf{E}[X|Z = i]$) "quand $Z = i$ " . Revenons un instant aux exemples du paragraphe **2.1** pour s'habituer à cette notation . Les égalités (2.1) , (2.2) , (2.5) , (2.8) deviennent ainsi respectivement

$$\mathbf{P}[S_p = k|S_n] = \frac{\begin{pmatrix} p \\ k \end{pmatrix}\begin{pmatrix} n-p \\ S_n - k \end{pmatrix}}{\begin{pmatrix} n \\ S_n \end{pmatrix}}$$

$$\mathbf{P}[X_1 = 1|S_n] = \frac{S_n}{n} \quad ; \quad \mathbf{P}[X_1 = 0|S_n] = 1 - \frac{S_n}{n}$$

$$\mathbf{E}[S_p|S_n] = p\frac{S_n}{n} \quad ; \quad \mathbf{E}[X_1|S_n] = \frac{S_n}{n}$$

$$\mathbf{E}[S|X_1 + X_2] = X_1 + X_2 + 21$$

On aura noté que la v.a. conditionnante Z peut prendre ses valeurs dans un espace produit ; c'est le cas dans la formule (2.10) où l'on posera $Z = (X_1, X_2)$, ce qui conduit à l'égalité

$$\mathbf{E}[S|X_1, X_2] = \mathbf{E}[S|Z] = X_1 + X_2 + \mathbf{E}[X_3]$$

2.2.2 Une caractérisation de $\mathbf{E}[X|Z]$

Proposition

$\mathbf{E}[X|Z]$ est l'unique v.a.r Y vérifiant les 2 propriétés suivantes :

1. Y est une fonction déterministe de Z
2. $\mathbf{E}[Y\psi(Z)] = \mathbf{E}[X\psi(Z)] \quad \forall\psi\ : E \to \mathbb{R}$

Démonstration : Il est clair d'après sa définition que $Y = \mathbf{E}[X|Z]$ vérifie la propriété 1. D'autre part, les égalités

$$\mathbf{E}[\phi(Z)\psi(Z)] = \sum_i \phi(i)\psi(i)\mathbf{P}[Z = i] = \sum_i \mathbf{E}[X1_{[Z=i]}]\psi(i) = \mathbf{E}[X\sum_i 1_{[Z=i]}\psi(i)] = \mathbf{E}[X\psi(Z)]$$

montrent que Y vérifie 2. Réciproquement, soit $Y = h(Z)$ une v.a.r. vérifiant 1. et 2. ; cette dernière propriété écrite pour $\psi = 1_{\{j\}}$ conduit à l'égalité (Cf. (1.3)) $\quad h(j)\mathbf{P}[Z = j] = \mathbf{E}[X1_{[Z=j]}]$, si bien que $h = \phi$, d'où le résultat.

Énonçons maintenant quelques propriétés simples de cette espérance conditionnelle.

2.2.3 Proposition :

1. $\mathbf{E}[aX + bX'|Z] = a\mathbf{E}[X|Z] + b\mathbf{E}[X'|Z] \ \ \forall a \in \mathbb{R} \ \ \forall b \in \mathbb{R}$
2. $X \geq 0 \Longrightarrow \mathbf{E}[X|Z] \geq 0$
3. $\mathbf{E}[\mathbf{E}[X|Z]] = \mathbf{E}[X]$
4. Si X et Z sont indépendantes , $\mathbf{E}[X|Z]$=$\mathbf{E}$[X]
5. Soit $\psi : F \to \mathbb{R}$ une fonction numérique sur F ; on a
$$\mathbf{E}[X\psi(Z)|Z] = \psi(Z)\mathbf{E}[X|Z]. \tag{2.13}$$
En particulier
$$\mathbf{E}[\psi(Z)|Z] = \psi(Z).$$
6. (2.13) peut être généralisée comme suit ; on associe à la fonction $H : \Omega \times F \to \mathbb{R}$ la variable aléatoire réelle $Y = H(., Z)$; si l'on pose alors $K(., i) = \mathbf{E}[H(., i)|Z]$, on a
$$\mathbf{E}[Y|Z] = K(., Z) \tag{2.14}$$
7. Soient X une v.a à valeurs dans E indépendante de Z et $h : E \times F \to \mathbb{R}$; on pose $v(i) = \mathbf{E}[h(X, i)] \quad (i \in F)$. On a alors
$$\mathbf{E}[h(X, Z)|Z] = v(Z) \tag{2.15}$$

Démonstrations

1. et 2. étant évidents, passons à 3 ; on a
$$\mathbf{E}[\phi(Z)] = \sum_{i \in Z(\Omega)} \phi(i)\mathbf{P}[Z = i] = \sum_i \mathbf{E}[X1_{Z=i}] = \mathbf{E}[X] \ .$$

Le point 4. provient de ce que $\phi(i) = \mathbf{E}[X] \quad \forall i \in Z(\Omega)$. Pour montrer 5., on remarquera que $\mathbf{E}[X\psi(Z)|Z = i] = \psi(i)\phi(i)$; on a donc $\mathbf{E}[X\psi(Z)|Z] = \psi(Z)\phi(Z) = \psi(Z)\mathbf{E}[X|Z]$.
Montrons enfin la formule (2.14). Posons pour cela $u(i, j) = \mathbf{E}[H(., i)|Z = j]$; on a alors
$$\mathbf{E}[Y|Z = i] = \mathbf{E}[H(., i)|Z = i] = u(i, i)$$
d'où $\mathbf{E}[Y|Z] = u(Z, Z) = K(., Z)$. La formule (2.15) en résulte aisément : si l'on pose
$$H(., i) = h(X, i),$$
la fonction u vérifie alors $u(i, j) = v(i) \ \ \forall i \, \forall j$ en vertu de l'hypothèse d'indépendance .

On aura noté que les v.a. U qui sont des fonctions déterministes de Z jouent un rôle particulier dans le conditionnement par Z . Elles sont caractérisées par la propriété suivante : quelque soit l'épreuve ω , la connaissance de $Z(\omega)$ entraîne celle de $U(\omega)$. Intuitivement, on peut se représenter cela comme le contraire de l'indépendance : si, en effet, une v.a. V est indépendante de Z, la connaissance de $Z(\omega)$ n'est d'aucun secours pour prédire les valeurs possibles de $V(\omega)$.

2.2.4 Exercice : Conditionnement et projection orthogonale

Supposons $\mathbf{P}(\omega) > 0 \quad \forall \omega$ et posons $|\Omega| = n$; On notera $\mathcal{V}$ l'espace vectoriel des v.a.r sur Ω et $\mathcal{V_Z}$ le sous ensemble de $\mathcal{V}$ formé des fonctions déterministes de Z ; on posera enfin $|Z(\Omega)| = p$.

1. On munit $\mathcal{V}$ du produit scalaire défini par $< X, Y >= \mathbf{E}[XY]$; donner la dimension de l'espace euclidien ainsi défini ainsi que celle du sous espace $\mathcal{V_Z}$.
2. Soit $X \in \mathcal{V}$; montrer que $\mathbf{E}[X|Z]$ est la projection eucidienne de X sur $\mathcal{V_Z}$.
3. On note $\mathcal{R_Z}$ le sous espace de $\mathcal{V}$ formé des v.a.r. qui sont des fonctions affines de Z . Quelle est sa dimension ? Montrer qur la projection eucidienne $R[X|Z]$ de X sur $\mathcal{R_Z}$ est donnée par l'égalité
$$Y = \mathbf{E}[X] + r(Z - \mathbf{E}[Z])$$
quand on a posé $r = \frac{cov(X,Z)}{var(Z)}$
4. Montrer sans calculs l'inégalité $\mathbf{E}[(X - \mathbf{E}[X|Z])^2] \leq \mathbf{E}[(X - R[X|Z])^2]$.
5. Montrer que $\mathbf{E}[X|Z] = R[X|Z]$ si $|Z(\Omega)| = 2$

2.3 Algèbres et partitions

Nous supposerons dans ce paragraphe que le support de $\mathbf{P}$ est égal à Ω . Avant de commencer , interrogeons nous sur la dépendance en Z de $\mathbf{E}[X|Z]$. Introduisons d'abord la notion suivante :

Partition engendrée par une variable aléatoire.

Soit Z une v.a. á valeurs dans E ; la famille finie d'événements $\{[Z = k] \ \ , \ \ k \in Z(\Omega)\}$ constitue une partition finie de Ω appelée partition engendrée par Z ; elle sera notée $\mathcal{P}_Z$. Revenons alors à la définition de l'epérance conditionnelle ; on a, quand on a posé $A_k = [Z = k] \ \ , \ \ c_k = \frac{\mathbf{E}[X;A_k]}{\mathbf{P}A_k}$

$$\mathbf{E}[X|Z] = \sum_{k \in Z(\Omega)} c_k 1_{A_k}$$

On voit sur cette formule que $\mathbf{E}[X|Z]$ ne dépend de Z que par l'intermédiaire de la partition qu'elle engendre . Ainsi,
$$\mathbf{E}[X|Z] = \mathbf{E}[X|Z+1] = \mathbf{E}[X|Z^3]$$
puisque les partitions engendrées par $Z, Z+1, Z^3$ sont identiques ; en revanche, si Z est à valeurs dans $\{-1 \ , \ +1\}$, $\mathbf{E}[X|Z]$ et $\mathbf{E}[X|Z^2]$ sont différentes. La véritable notion intéressante est donc celle de conditionnement par une partition ; on est conduit aux définitions suivantes :

Algèbre engendrée par une partition .

- si $\mathcal{P}$ est une partition de Ω, on appellera *algèbre engendrée par* $\mathcal{P}$ et l'on notera $\mathcal{F}(\mathcal{P})$ (ou plus simplement $\mathcal{F}$ si aucune confusion n'est à craindre), la classe d'événements formée des réunions d'éléments de $\mathcal{P}$. La donnée d'une partition est équivalente à celle de l'algèbre qu'elle engendre ; on préfère néammoins raisonner sur les algèbres car elles possèdent des propriétés de stabilité relativement aux opérations booléennes que n'ont pas les partitions. Ainsi, si $\mathcal{F}$ est l'algèbre engendrée par $\mathcal{P}$,

$$F_1 et F_2 \in \mathcal{F} \Longrightarrow F_1 \cup F_2 \quad , \quad F_1 \cap F_2 \quad , \quad F_1^c \in \mathcal{F}$$

On notera également que Ω et ϕ appartiennent à $\mathcal{F}$
Les éléments de $\mathcal{P}$ seront appelés *atomes* de l'algèbre $\mathcal{F}$.

- Soient $\mathcal{P}$ et $\mathcal{P}'$ deux partitions de Ω ; on dira que $\mathcal{P}$ est *plus fine* que $\mathcal{P}'$, ce que l'on notera $\mathcal{P} \succ \mathcal{P}'$, si tout élément de $\mathcal{P}$ est contenu dans un élément de $\mathcal{P}'$; on vérifie aisément que cette propriété est équivalente á l'inclusion $\mathcal{F} \supset \mathcal{F}'$. Il en est ainsi si et seulement si tout atome de $\mathcal{F}'$ est dans $\mathcal{F}$.

2.3.1 Algèbre engendrée par une famille de variables aléatoires

Soit Z une v.a. à valeurs dans E . L'algèbre $\mathcal{F}(Z)$ engendrée par Z sera, par définition, l'algèbre engendrée par la partition $\mathcal{P}_Z$. Il est facile de voir que

$$\mathcal{F}(Z) = \{[Z \in \Gamma] \quad ; \quad \Gamma \subset E\}$$

Si Z est un couple (Z_1, Z_2), tout atome $[Z_1 = k_1]$ de $\mathcal{F}(Z_1)$ appartient à $\mathcal{F}(Z)$ puisqu'on peut l'écrire $\sum_{k_2}[Z_1 = k_1; Z_2 = k_2]$; il en résulte que $\mathcal{F}(Z_1) \subset \mathcal{F}(Z_1, Z_2)$
On généralise aisément par récurrence cette situation au cas de N variables aléatoires : soit $Z_1, Z_2, \ldots, Z_N$ une suite de v.a. ; notons $\mathcal{F}_n$ l'algèbre $\mathcal{F}(Z_1, Z_2, \ldots, Z_n)$. Alors

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_N$$

2.3.2 Variables aléatoires $\mathcal{F}$-mesurables

Soit $\mathcal{P}$ une partition engendrant une algèbre $\mathcal{F}$ une v.a. Y sera dite *$\mathcal{F}$-mesurable* si elle est constante sur chaque atome de $\mathcal{F}$ (i.e. sur chaque événement de $\mathcal{P}$).
Les remarques suivantes nous seront utiles

- $Y et Y'$ sont 2 v.a.r. $\mathcal{F}$-mesurables vérifiant $\mathbf{E}[Y 1_F] = \mathbf{E}[Y' 1_F] \ \ \forall F \in \mathcal{F}$, alors $Y = Y'$.
- Un événement $F \in \mathcal{F}$ si et seulement si 1_F est $\mathcal{F}$-mesurable

Variables aléatoires $\mathcal{F}(\mathcal{Z})$-mesurables

Nous allons montrer que ce sont tout simplement les v.a. que nous avons appelées jusqu'ici de manière peu élégante "fonctions déterministes de Z" .

Proposition : Soit Y une v.a.r. ; les deux propriétés suivantes sont équivalentes

1. Y est $\mathcal{F}(Z)$-mesurable .
2. Il existe une fonrtion numérique ϕ sur E telle que $Y = \phi(Z)$.

Démonstration : Si 2. est vérifiée, Y est constante sur les atomes de $\mathcal{F}(Z)$, d'où 1. ; réciproquement, posons $E = \{z_1, z_2, ..., z_p\}$; si Y est $\mathcal{F}(Z)$- mesurable, il existe, par définition, p constantes $c_1, c_2, ..., c_p$ telles que $Y = \sum_1^p c_i 1_{\{Z=z_i\}}$; il en résulte que Y vérifie 2., avec $\phi = \sum_1^p c_i 1_{\{z_i\}}$, (Cf.(1.3))

2.3.3 Conditionnement par une algèbre .

Définition : Soient X une v.a.r. et $\mathcal{F}$ l'algèbre engendrée par une partition $\mathcal{P}$ de Ω . On appellera espérance (conditionnelle) de X quand $\mathcal{F}$ la *variable aléatoire* notée $\mathbf{E}[X|\mathcal{F}]$ définie par

$$\mathbf{E}[X|\mathcal{F}] = \sum_{B\in\mathcal{P}} \mathbf{E}[X|B]1_B$$

On notera que $\mathbf{E}[X|\mathcal{F}(Z)]$ est égale à la v.a. $\mathbf{E}[X|Z]$ introduite en **2.2.1**. Voici maintenant les propriétés les plus importantes de l'espérance conditionnelle.

1. **Linéarite et positivité** Appelons $\mathcal{V}$ l'espace vectoriel (de dimension $|\Omega|$) des variables aléatoires réelles ; $X \mapsto \mathbf{E}[X|\mathcal{F}]$ definit alors une application linéaire et positive de $\mathcal{V}$ dans le sous espace (de dimension $|\mathcal{P}|$) formé par les v.a.r. $\mathcal{F}$-mesurables .
2. **Inegalité de convexité** Soit $u : \mathbb{R} \mapsto \mathbb{R}$ une application convexe ; on a l'inégalité

$$\mathbf{E}[u(X)|\mathcal{F}] \geq u(\mathbf{E}[X|\mathcal{F}]) \tag{2.16}$$

En particulier ,

$$\mathbf{E}[|X||\mathcal{F}] \geq |\mathbf{E}[X|\mathcal{F}]| \tag{2.17}$$

$$\mathbf{E}[X^2|\mathcal{F}] \geq \mathbf{E}[X|\mathcal{F}]^2. \tag{2.18}$$

Démonstration Soit $B \in \mathcal{P}$. La variable aléatoire figurant au premier membre de (2.16) est égale sur B à $\mathbf{E}[u(X)|B]$ et le second membre à $u(\mathbf{E}[X|B])$; on termine la démonstration en appliquant l'inégalité de convexité à la probabilité $\mathbf{P}_B$ introduite en **2.1**

3. **Une caractérisation fondamentale :** $\mathbf{E}[X|\mathcal{F}]$ est l'unique v.a.r. vérifiant les 2 propriétés suivantes :
a) Y est $\mathcal{F}$-mesurable b) $\forall F \in \mathcal{F},\ \ \mathbf{E}[X1_F] = \mathbf{E}[Y1_F]$
Laissant au lecteur le soin de démontrer la partie directe, nous nous bornerons à montrer qu'une v.a.r. Y satisfaisant aux points a) et b) est nécessairement égale a $\mathbf{E}[X|\mathcal{F}]$
En vertu de a), il existe une famille de scalaires $(c_A; A \in \mathcal{P})$ telle que $Y = \sum_{A\in\mathcal{P}} c_A 1_A$; l'égalité b) , applquée à un événement B de $\mathcal{P}$ permet de calculer c_B ; on a en effet : $\mathbf{E}[Y1_B] = c_B\mathbf{P}B$ ce qui entraine $c_B = \mathbf{E}[X|B]$; c'est le résultat cherché .
Remarque : On peut remplacer b) par
b') : $\forall Z\quad \mathcal{F}-$ mesurable, $\mathbf{E}[XZ] = \mathbf{E}[YZ]$;
(on utilise simplement le fait que Z est une combinaison linéaire d'indicateurs d'atomes et la linéarite de l'espérance .)
4. **Conditionnements successifs :** Soient $\mathcal{F}$ et $\mathcal{G}$ deux algèbres avec $\mathcal{F} \supset \mathcal{G}$; on a alors, quelquesoit la v.a.r. X ,

$$\mathbf{E}[X|\mathcal{F}|\mathcal{G}] = \mathbf{E}[X|\mathcal{G}] \quad ; \quad \mathbf{E}[\mathbf{E}[X|\mathcal{F}]] = \mathbf{E}[X] \tag{2.19}$$

Il nous suffira de montrer la première égalité ; les premiers et second membres de celle-ci , que nous appellerons respectivement U et V étant $\mathcal{G}$-mesurables, tout revient à montrer les égalités

$$\mathbf{E}[U1_G] = \mathbf{E}[V1_G] \quad \forall G \in \mathcal{G}.$$

Mais $\mathbf{E}[U1_G] = \mathbf{E}[\mathbf{E}[X|\mathcal{F}]1_G] = \mathbf{E}[X1_G]$, la deuxième égalite venant du fait que $G \in \mathcal{F}$. Retenons une conséquence pratique de ce résultat : avec les notations de **2.3.1.** , on a par exemple

$$\mathbf{E}[X|Z_1, Z_2, \ldots, Z_N|Z_1, Z_2] = \mathbf{E}[X|Z_1, Z_2]$$

5. **Si Y est $\mathcal{F}$-mesurable ,** $\mathbf{E}[Y|\mathcal{F}] = Y$; plus généralement, quelquesoit la v.a.r X,

$$\mathbf{E}[XY|\mathcal{F}] = Y\mathbf{E}[X|\mathcal{F}]. \tag{2.20}$$

Démonstration : Désignant par $UetV$ les deux membres de (10) , nous avons à montrer que

$$\mathbf{E}[U1_B] = \mathbf{E}[V1_B] \ \ \forall B \in \mathcal{F}.$$

Vous laissant calculer le premier membre, notons que $\mathbf{E}[V1_B] = \mathbf{E}[(Y1_B)\mathbf{E}[X|\mathcal{F}]] = \mathbf{E}[XY1_B]$ en vertu de 3.b') .
Voici maintenant une généralisation souvent utilisée de l'égalité (2.20) . Les hypothèses sont les suivantes : X et Y sont deux v.a. à valeurs respectives dans E et F , u est une fontion numérique sur $E \times F$; on suppose Y $\mathcal{F}$-mesurable ; si l'on pose $H(\omega, y) = \mathbf{E}[f(X, y)|\mathcal{F}](\omega)$ on a :

$$\mathbf{E}[f(X, Y)|\mathcal{F}](\omega) = H(\omega, Y(\omega)) \tag{2.21}$$

Démonstration : Montrons tout d'abord que la v.a. figurant au second membre de (2.21) , (que nous noterons suivant l'usage des probabilistes , $H(., Y)$) est $\mathcal{F}$-mesurable ; choisissons pour cela un atome B de $\mathcal{F}$, désignons par y et par c les valeurs (constantes) prises respectivement par Yet $\mathbf{E}[f(X, y)|\mathcal{F}]$ sur B ;on voit, en regardant sa définition, que $H(., Y)$ est égale à la constante c sur B ; elle est donc $\mathcal{F}$-mesurable ; en outre , on peut écrire :

$$\mathbf{E}[1_B H(., Y)] = \mathbf{E}[1_B H(., y)] = \mathbf{E}[1_B \mathbf{E}[f(X, y)|\mathcal{F}]] = \mathbf{E}[1_B f(X, y)] = \mathbf{E}[1_B f(X, Y)],$$

ce qui montre l'égalité (2.21) .

6. **Si X est indépendante de $\mathcal{F}$,** la variable aléatoire $\mathbf{E}[X|\mathcal{F}]$ se réduit à une constante ; plus précisement :

$$\mathbf{E}[X|\mathcal{F}] = \mathbf{E}[X] \tag{2.22}$$

Dans ce cas , la formule (2.21) se simplifie ainsi : si l'on pose $h(y) = \mathbf{E}[f(X, y)]$, on a, en appliquant (2.21) et (2.22),

$$\mathbf{E}[f(X, Y)|\mathcal{F}] = h(Y) \tag{2.23}$$

2.3.4 Lois conditionnelles

Dans ce paragraphe , X désignera une variable aléatoire à valeurs dans E .

1. *Conditionnement par une algèbre :* Soit $\mathcal{F}$ une algèbre (engendrée par une partition $\mathcal{P}$); si $x \in E$, on posera

$$\Pi_X(x|\mathcal{F}) = \mathbf{P}[\{X = x\}|\mathcal{F}] = \mathbf{E}[1_{\{X=x\}}|\mathcal{F}]$$

(Bien noter qu'il s'agit d'égalités entre *variables aléatoires* $\mathcal{F}$-mesurables) . Les v.a. $\Pi_X(x|\mathcal{F})$ possèdent les 2 propriétés suivantes resultant immédiatement de ce que nous avons vu en **2.3.3**

(a) $\forall x \in E$ $\qquad \Pi_X(x|\mathcal{F})$ est $\mathcal{F}$-mesurable et ≥ 0

(b) $\forall \omega \in \Omega$ $\qquad \sum_{x\in E} \Pi_X(x|\mathcal{F}) = \mathbf{E}[1_\Omega|\mathcal{F}] = 1$

Pour y voir plus clair, notons provisoirement Q la fonction de 2 variables définie par $Q(x,\omega) = \Pi_X(x|\mathcal{F})(\omega)$; les propriétés (a) et (b) s'expriment sur Q de la façon suivante :

(a) $\forall x \in E$ $\qquad \omega \mapsto Q(x,\omega)$ est $\mathcal{F}$-mesurable et ≥ 0

(b) $\forall \omega \in \Omega$ $\qquad x \mapsto Q(x,\omega)$ est une probabilité sur E

Cette fonction Q , que l'on notera $\Pi_X(.|\mathcal{F})$ sera appelée *loi conditionnelle de X quand $\mathcal{F}$* . On se rappellera qu'il s'agit , (cf (b)) , d'une famille de lois de probabilité sur E indexée par Ω

2. *Conditionnement par une variable aléatoire* Soit Z une variable aléatoire à valeurs dans F ; on supposera comme d'habitude que $Z(\Omega) = F$. La *loi conditionnelle de X quand Z* sera par définition $\Pi_X(.|\mathcal{F}(Z))$. Cet objet est donc , comme au point 1. précédent dont nous reprenons les notations, une fonction Q des 2 variables (x,ω) ; mais d'après (b) , et compte tenu de **2.3.2** , il existe $q : E \times F \mapsto [0 \quad 1]$ telle que $Q(x,\omega) = q(x, Z(\omega))$, q se calculant très simplement par la formule $q(x,z) = \mathbf{P}[X = x|Z = z]$; on a donc l'égalite $\Pi_{(X,Z)}(x,z) = q(x,z)\Pi_Z(z)$ qui permet de montrer de facon élémentaire une relation que vous avez sans doute déjà vue dans votre jeunesse :
$$\mathbf{E}[f(X,Z)] = \sum_{z\in F} \Pi_Z(z) \sum_{x\in E} f(x,z)q(x,z)$$
On peut retrouver cette formule par une voie plus abstraite , (ce que je vous conseille de faire pour vous habituer au maniement des espérances conditionnelles) ; c'est une conséquence simple de l'égalité $\mathbf{E}[f(X,Z)] = \mathbf{E}[\mathbf{E}[f(X,Z)|Z]\,]$ et de (2.21) .

3. *Lois conditionnelles et indépendance* : On a vu en (2.22) que l'indépendance d'un couple (X,Z) de v.a.r. entrainait l'égalité $\mathbf{E}[X|Z] = \mathbf{E}[X]$; la réciproque est grossièrement inexacte (trouver un contre-exemple) ; il existe en revanche une condition nécessaire et suffisante d'indépendance en termes de loi conditionnelle ; on a le résultat suivant :
$$X \text{ et } Z \text{ sont indépendantes si et seulement si } \Pi_X(.|Z) = \Pi_X$$
Cette égalité signifie que la famille de probabilités indexée par Ω figurant au premier membre est en réalité réduite à la seule loi marginale de X . La démonstration de ce résultat,très simple, est laissée au lecteur .

2.3.5 Exercice

1. **Conditionnement d'une loi multinomiale .** Soient p et q deux entiers > 0 *et* $\alpha_1, \alpha_2, ..., \alpha_q$ *q reels* > 0 ; on pose $H = \{k \in \{0,1,,...,p\}^q \mid k_0 + k_1 + k_2 + ... + k_q = p\}$; On dit qu'une v.a. $Z = (Z_1, Z_2, ...Z_q)$ à valeurs dans $\{0,1,2,...,p\}^q$ est multinomiale de parametres $(\alpha_1, \alpha_2, ..., \alpha_q, p, q)$ (cf **1.2.6**) si elle admet pour loi la probabilité sur $\{0,1,...,p\}^q$, portée par H définie par
$$\mathbf{P}[Z_1 = k_1; Z_2 = k_2; ...Z_q = k_q] = \frac{p!}{k_1!k_2!...k_q!}\alpha_1^{k_1}\alpha_2^{k_2}...\alpha_q^{k_q} \quad \forall k \in H$$
Soit Z une telle variable aléatoire ; montrer que la loi conditionnelle de $(Z_1, Z_2, ..., Z_{q-1})$ quand Z_q est la loi multinomiale de paramètres $(\frac{n_1}{n-n_q}, \frac{n_2}{n-n_q}, ..., \frac{n_{q-1}}{n-n_q} \quad , p - Z_q \quad , q-1)$; on notera que le support de cette loi conditionnelle est aléatoire .

2. **Sans titre** Soit $(X_1, X_2, ..., X_n)$ une suite de n variables aléatoires réelles, indépendantes, equidistribuées ; on pose $S_n = X_1 + X_2 + ...X_n$;
   (a) Soit $f : \mathbb{R} \mapsto \mathbb{R}$; montrer que $\mathbf{E}[X_1 f(S_n)] = \mathbf{E}[X_2 f(S_n)] = ... = \mathbf{E}[X_n f(S_n)]$
   (b) En déduire que $\mathbf{E}[X_1|S_n] = \frac{S_n}{n}$

2.3.6 Proriétés markoviennes d'une marche aléatoire

Chaines de Markov

Revenons rapidement sur la notion de *Chaine de Markov* qui a été étudiée au premier semestre . Dans ce paragrahe , on désigera par

- E un ensemble fini comportant k éléments que nous appellerons *espace d'états* de la chaine .
- $P = (p_{ij})$ une matrice carrée d'ordre k markovienne , (i.e $p_{ij} \geq 0 \quad \forall i, j$ et $\sum_j p_{ij} = 1 \ \ \forall i$) ; si f est une foncion numérique sur E , on posera $Pf(i) = \sum_j p_{ij} f(j)$

Une suite $(X_0, X_1, X_2, \cdots, X_N)$ de v.a. à valeurs dans E sera appelée *chaine de Markov de matrice de transition* P si quelque soit la fonction numérique f sur E ,

$$\mathbf{E}[f(X_{n+1})|\mathcal{F}_n] = Pf(X_n) \qquad \forall n \in \{0, 1, 2, \cdots, (N-1)\}$$

Le lecteur vérifiera , à titre d'exercice , une conséquence de cette égalité qui lui rappellera peut être quelques souvenirs :

$$\mathbf{P}[X_{n+1} = j | X_0 = i_0; X_1 = i_1; \cdots; X_n = i] = \mathbf{P}[X_{n+1} = j | X_n = i] = p_{ij}$$

Matrice de transition d'une marche aléatoire

Soient μ une loi de probabilité sur $\mathbb{R}^d$ de support fini F et $((Y_n) \ ; \ \ n \geq 1)$ une suite de v.a. indépendantes équidistribuées de loi μ à valeurs dans $\mathbb{R}^d$. On pose

$$X_0 = 0 \quad , \quad X_n = \sum_{i=1}^{i=n} Y_i \quad \forall n \geq 1$$

(X_n) *est alors une chaine de Markov de matrice de transition* $P_{ij} = \mu(j - i)$

Démontrons cela ; on a , en vertu de (2.23)

$$\mathbf{E}[f(X_{n+1})|\mathcal{F}_n] = \mathbf{E}[f(X_n + Y_{n+1})|\mathcal{F}_n] = \sum_{k \in F} f(X_n + k)\mu(k) = Pf(X_n)$$

quand on a posé $Pf(i) = \sum_{k \in F} f(i + k)\mu(k) = \sum_{j \in F+i} f(j)\mu(j - i)$; on en déduit immédiatement le résultat annoncé .

Chapter 3

Martingales

Soit $(\Omega, \mathbf{P})$ un espace de probabilité (fini) ; on supposera dans tout ce chapitre que le support de $\mathbf{P}$ est égal a Ω ; N sera un nombre entier strictement positif fixé dans toute la suite .

3.1 Filtrations ; suites adaptées et prévisibles

3.1.1 Définitions :

- On appellera *filtration* sur Ω une suite croissante d'algèbres $(\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, ...\mathcal{F}_N)$ sur Ω . Un espace probabilisé muni d'une filtration sera appelé *espace filtré*

- On dira qu'une suite $(X_0, X_1, X_2, ..., X_N)$ de v.a. est *adaptée* à cette filtration si , $\forall i \in \{0, 1, 2, ..., N\}$, X_i est $\mathcal{F}_i$-mesurable .

- On dira que cette suite est *prévisible* si $\forall i \in \{0, 1, 2, ..., N\}$, X_i est $\mathcal{F}_{i-1}$-mesurable ; (en convenant que $\mathcal{F}_{-1} = \mathcal{F}_0$).

- Soit $(X_0, X_1, X_2, ..., X_N)$ une suite de v.a. . On appellera *filtration naturelle* engendrée par cette suite la filtration $\mathcal{F}_n = \mathcal{F}(X_0, X_1, X_2, ..., X_n)$; (nous avons vu en **2.3.1** qu'il s'agissait bien d'une filtration).
  La suite(X_n) est adaptée a cette filtration tandis que la suite (X_{n-1}) ,(avec la convention $X_{-1} = X_0$) ,est prévisible. On peut construire à l'aide de (X_n) d'autres suites adaptées intéressantes : supposons, pour simplifier, que les v.a. X_n soient à valeurs dans le même espace E et donnons nous une suite $(\phi_n : E^{n+1} \mapsto \mathbb{R})$ d'applications numériques ; pour tout n, la v.a.r. $Y_n = \phi_n(X_0, X_1, X_2, ..., X_n)$ est $\mathcal{F}_n$-mesurable , (elle est constante sur chaque atome de $\mathcal{F}_n$) , de sorte que la suite (Y_n) est adaptée ; dans le cas ou les v.a. X_n sont réelles , les suites $S_n = X_0 + X_1 + X_2 + ... + X_n \; et \; \Pi_n = X_0 X_1 X_2 ... X_n$ sont ainsi des suites adaptées .

3.2 Martingales

Soit $(\Omega, (\mathcal{F}_n), \mathbf{P})$ un espace filtré .

3.2.1 Définitions

Soit $(M_n\ ;\ 0 \leq n \leq N)$ une suite adaptée à la filtration $(\mathcal{F}_n)$; nous dirons que cette suite est une $(\mathcal{F}_n)$-martingale (ou plus brièvement une martingale), si elle vérifie les égalités

$$\mathbf{E}[M_{n+1}|\mathcal{F}_n] = M_n \quad \forall n \in \{0, 1, 2, ..., N-1\}. \tag{3.1}$$

Nous dirons que c'est une *surmartingale* (resp. une *sousmartingale*) si elle vérifie les inégalités

$$\mathbf{E}[M_{n+1}|\mathcal{F}_n] \leq M_n \qquad\qquad (\text{resp.}\mathbf{E}[M_{n+1}|\mathcal{F}_n] \geq M_n)$$

Si le contexte ne comporte pas de filtration , il sera sous-entendu que l'on travaille dans la filtration naturelle.

3.2.2 Premières propriétés

Soit (M_n) une martingale ;

1. Si $n \leq n+p \leq N$, $\mathbf{E}[M_{n+p}|\mathcal{F}_n] = M_n$, en particulier $M_n = \mathbf{E}[M_N|\mathcal{F}_n]$
2. Si $\phi : \mathbb{R} \mapsto \mathbb{R}$ est convexe , $(\phi(M_n))$ est une sousmartingale ; c'est en particulier le cas pour $|M_n|$ *et* M_n^2
3. Une combinaison linéaire d'une famille finie de martingales est une martingale .
4. $\mathbf{E}[M_n] = \mathbf{E}[M_0] \quad \forall n$; on retiendra que $\mathbf{E}[M_n]$ ne dépend pas de n
   Le premier point résulte immédiatement de (2.19) ; on a en effet

   $$\mathbf{E}[M_{n+2}|\mathcal{F}_n] = \mathbf{E}[M_{n+2}|\mathcal{F}_{n+1}|\mathcal{F}_n] = \mathbf{E}[M_{n+1}|\mathcal{F}_n] = M_n,$$

   d'où le résultat annoncé par récurrence.
   Le second est une conséquence de l'inégalité de convexité (2.16) ; on a en effet

   $$\mathbf{E}[\phi(M_{n+1})|\mathcal{F}_n] \geq \phi(\mathbf{E}[M_{n+1}|\mathcal{F}_n]) = \phi(M_n),$$

   le troisième est évident ; si vous ne savez pas montrer le quatrième , c'est qu'un point important, (sans doute (2.19)), vous a échappé ; il est plus que temps d'y remédier .

3.2.3 Premiers exemples

1. **Sommes de v.a.r. centrées indépendantes :** Soit $(X_0, X_1, X_2, ..., X_N)$ une suite de v.a.r. centrées indépendantes ; posons $S_n = X_0 + X_1 + X_2 + ... + X_n$; $\mathcal{F}_n = \mathcal{F}(X_0, X_1, X_2, ..., X_n)$; je dis que (S_n) est une martingale , en effet , puisque X_{n+1} est indépendante de $\mathcal{F}_n$,

   $$\mathbf{E}[S_{n+1}|\mathcal{F}_n] = X_0 + X_1 + ... + X_n + \mathbf{E}[X_{n+1}|\mathcal{F}_n] = S_n + \mathbf{E}[X_{n+1}] = S_n$$

Supposons de plus les v.a. X_n équidistribuées ; elles ont alors toutes la même variance que nous appellerons σ^2 ; dans cette situation , $S_n^2 - (n+1)\,\sigma^2$ est aussi une martingale comme le prouve le petit calcul suivant :

$$\mathbf{E}[S_{n+1}^2|\mathcal{F}_n] = \mathbf{E}[(S_n + X_{n+1})^2|\mathcal{F}_n] = S_n^2 + 2S_n\mathbf{E}[X_{n+1}] + \mathbf{E}[X_{n+1}^2] = S_n^2 + \sigma^2$$

d'où il résulte aisément que

$$\mathbf{E}[S_{n+1}^2 - (n+2)\sigma^2|\mathcal{F}_n] = S_n^2 - (n+1)\sigma^2$$

2. **Martingales produit :** Soit $(T_0, T_1, T_2, ..., T_N)$, une suite de v.a. positives , indépendantes , d'espérance 1 ; la suite $M_n = T_0T_1T_2...T_n$ est alors une martingale par rapport à la filtration naturelle engendrée par les v.a. T_n . En effet ,

$$\mathbf{E}[M_{n+1}|\mathcal{F}_n] = T_0T_1...T_n\mathbf{E}[T_{n+1}] = M_n$$

Il y a d'autres exemples de martingales intéressantes associées à des suites indépendantes, mais ce n'est pas là le plus important ; le succès de cette théorie est dû au fait qu'elle permet d'aller très au delà de l'indépendance : un certain nombre de transformations simples détruisent en effet celle ci tout en préservant la propriété de martingale ; nous allons étudier deux d'entre elles .

3.3 Transformées de martingales

Sur l'espace filtré $(\Omega, \mathcal{F}_n, \mathbf{P})$, on se donne pour $0 \le n \le N$:
- une suite (M_n) adaptée à $(\mathcal{F}_n)$
- une suite prévisible (U_n)

Introduisons d'autre part la notation suivante : si $(X_n\ ;\ 0 \le n \le N)$ est une suite de v.a.r. , on désignera par $(\Delta X_n\ ;\ 1 \le n \le N)$ la suite des accroissements de (X_n) définie par

$$\Delta X_n = X_n - X_{n-1}.$$

On peut reconstituer (X_n) connaissant la suite de ses accroissements et sa valeur initiale : on a de manière évidente ,

$$X_n = X_0 + \sum_{k=1}^{k=n} \Delta X_k. \tag{3.2}$$

On notera qu'une suite adaptée (M_n) est une martingale si et seulement si

$$\mathbf{E}[\Delta M_n|\mathcal{F}_{n-1}] = 0 \ \ \forall n \ge 1 \tag{3.3}$$

3.3.1 Définition :

On appellera transformée de (M_n) par(U_n) et l'on notera $(U*M)_n$ la suite définie par

$$(U*M)_n = U_0M_0 + \sum_{k=1}^{k=n} U_k\Delta M_k \tag{3.4}$$

Il résulte de ce qui vient d'etre dit que la suite $(U*M)_n$ est caracterisée par les relations :

$$(U*M)_0 = U_0M_0 \qquad\qquad \Delta(U*M)_n = U_n\Delta M_n \quad \forall n \geq 1 \tag{3.5}$$

Un week-end à Monte Carlo

Imaginons que vous jouiez a pile ou face dans un casino : vous gagnez 1F si le coup est pile et perdez 1F dans le cas contraire ; pour arriver a ce résultat dans un environnement rappelant celui d'une salle de jeux , disons plutôt que vous misez uniquement 1F sur pile ; vous gagnez une fois votre mise en cas de succès , et la perdez en cas d'échec . On supposera que vous disposez à votre arrivée dans le jeu d'une somme d'argent M_0 .
Quittons smoking et robe de bal , le temps d'introduire

- la loi symétrique sur l'ensemble $\{-1,+1\}$ que l'on notera Π $\quad(\Pi(-1) = \Pi(1) = 1/2)$
- une suite de v.a. indépendantes , équidistribuées de loi Π que nous noterons $(X_n\,; 1 \leq n \leq N)$

X_n représentera le gain procuré par le n-ieme coup tandis que $M_n = M_0 + X_1 + X_2 + ... + X_n$ sera la quantité d'argent dont vous disposerez à cet instant . Puisque les v.a. X_n sont centrées, (M_n) est une martingale ,(c'est à peu de choses près un cas particulier de l'exemple **3.2.3.**1) . En réalité , dans un casino , les règles du jeu sont faites de telle façon que la loi des v.a. X_n ont une espérance légèrement négative , et que l'évolution de votre fortune est une surmartingale .
Si vous êtes un peu joueur , 2 choses vous paraitront insupportables : ne pas pouvoir gagner d'argent si le coup est face d'une part , limiter vos mises a 1F quand vous vous sentez en veine d'autre part ; on peut supprimer ces inconvénients en introduisant une suite (U_n) de nombres réels destinée à figurer vos mises (une mise négative signifiant que vous avez parié sur face) . Comment déterminerez vous votre mise ? Si vous êtes comme tout le monde, vous serez influencé par l'évolution du jeu (à tort comme on va le voir dans quelques lignes mais laissons là cette question pour le moment) et votre mise à l'instant n sera fonction des observations que vous aurez faites jusqu'à cet instant .

En termes mathématiques , cela se traduit par le fait que votre mise U_n est aléatoire et que la suite (U_n) est adaptée . Malheureusement , les directeurs de casino ne sont pas disposés à admettre n'importe quelle suite adaptée : il suffirait , si tel était le cas , de miser $U_n = 10^6(1_{\{X_n=1\}} - 1_{\{X_n=-1\}})$ pour empocher un million de francs de façon certaine à chaque coup ; on vous imposera donc de décider de votre mise U_n au n-ieme coup en disposant uniquement des informations fournies par le jeu jusqu'à l'instant $n-1$ (cette restriction est opérée à la roulette à l'aide du fameux "rien ne va plus") ; on astreint donc (U_n) à être non seulement adaptée mais ***prévisible*** .

Ayant misé de cette facon, votre fortune au temps n est devenue , si l'on convient que $U_0 = 1$,

$$M_0 + U_1X_1 + U_2X_2 + ... + U_nX_n = U_OM_0 + U_1\Delta M_1 + U_2\Delta M_2 + ... + U_n\Delta M_n.$$

Vous avez reconnu la transformée de la martingale (M_n) par le processus previsible (U_n) .
Il existe des joueurs convaincus de l'existence de stratégies gagnantes que l'on peut construire après avoir passé des nuits à noter et étudier des séries d'un grand nombre de coups ; une telle stratégie est appelée dans le langage courant une "martingale" , qu'il faut se garder de confondre avec la notion de martingale des mathématiciens . Le théorème qui suit est désespérant : il montre qu'aucune stratégie concernant vos mises ne peut transformer la martingale de base (M_n) en une sousmartingale stricte (ce qui serait bien agréable , puisque ses espérances formeraient une suite strictement croissante) .

Théorème

Sur un espace filtré $(\Omega\,,(\mathcal{F}_n)\,,\mathbf{P})$ soient (M_n) une martingale (resp. une sur ou une sous-martingale) et (U_n) une suite prévisible (resp. une suite prévisible positive) ; $(U * M)_n$ est alors une martingale (resp. une sur ou une sous-martingale).

Démonstration : Bornons nous au cas des martingales et posons $M'_n = (U * M)_n$;on a $\Delta M'_n = U_n \Delta M_n$ d'où , en utilisant la prévisibilité de (U_n) ,

$$\mathbf{E}[\Delta M'_n | \mathcal{F}_{n-1}] = U_n \mathbf{E}[\Delta M_n | \mathcal{F}_{n-1}] = 0$$

3.3.2 Sur la propriété de représentation prévisible

Soient $(0 = M_0, M_1, ..., M_N)$ une martingale , $(\mathcal{F}_n)$ sa filtration naturelle ; on dira que (M_n) possède la ***propriété de représentation prévisible*** si, quelque soit la $(\mathcal{F}_n)$-martingale (Z_n) , il existe une constante c et un processus prévisible U_n tels que $(Z_n) = c + (U * M)_n \quad \forall n.$

Nous retrouverons cette propriété dans les prochains chapitres sous le nom de *marchés complets* ; pour l'instant , nous allons simplement nous demander s'il existe de telles martingales ; les résultats qui vont suivre montrent , que dans notre cadre discret il en existe , certes , mais très peu , (contrairement aux modèles à temps continu où beaucoup de martingales fondamentales possèdent la p.r.p.).
Soit $(\epsilon_1, \epsilon_2, ..., \epsilon_N)$ une suite de v.a.r. indépendantes , équidistribuées , centrées ; posons $\epsilon_0 = 0$ et appelons $(\mathcal{F}_n)$ la filtration naturelle engendrée par cette suite ; considérons la martingale $M_n = (\epsilon_0 + \epsilon_1 + ... + \epsilon_n)$. Pour éviter les trivialités , nous supposerons ϵ_1 non dégénérée . Les filtrations naturelles engendrées par les suites (M_n) et (ϵ_n) sont identiques, comme on peut le vérifier à titre d'exercice.

Proposition : (M_n) possède la propriété de représentation prévisible si et seulement si ϵ_1 prend ses valeurs dans un ensemble réduit à 2 points .

Démonstration : Soient a et $b > 0$; supposons tout d'abord que ϵ_1 prenne ses valeurs dans $\{-a, b\}$ et montrons que (M_n) possède alors la p.r.p. ; posons

$$p = \mathbf{P}[\epsilon_1 = b]\ \ ,\ q = 1 - p = \mathbf{P}[\epsilon_1 = -a].$$

Puisque ϵ_1 est centrée, $0 = pb - (1-p)a$ d'où $p = \frac{a}{a+b}$, $q = \frac{b}{a+b}$, $\frac{p}{q} = \frac{a}{b}$.
Considérons maintenant une v.a.r. F_n $\mathcal{F}_n$-mesurable verifiant $\mathbf{E}[F_n|\mathcal{F}_{n-1}] = 0$; nous allons montrer qu'il existe γ_n $\mathcal{F}_{n-1}$-mesurable telle que $F_n = \gamma_n \epsilon_n$; d'après la proposition

2.3.2 , il existe une fonction numérique f_n sur $\mathbb{R}^{n+1}$ (c'est à dire une fonction réelle de $n+1$ variables réelles) telle que $F_n = f_n(\epsilon_0, \epsilon_1, \epsilon_2, ..., \epsilon_n)$. Comme ϵ_n est indépendante de $\mathcal{F}_{n-1}$, l'espérance conditionnelle $\mathbf{E}[F_n|\mathcal{F}_{n-1}]$ se calcule à l'aide de la formule (2.23) de sorte que l'on a l'égalité $pf_n(\epsilon_0, \epsilon_1, ..., \epsilon_{n-1}, b) + qf_n(\epsilon_0, \epsilon_1, ...\epsilon_{n-1}, -a) = 0$ d'où l'on tire $f_n(\epsilon_0, ..., \epsilon_{n-1}, -a) = -\frac{a}{b} f_n(\epsilon_0, ..., \epsilon_{n-1}, b)$.
On peut alors écrire

$$\begin{aligned} F_n &= f_n(\epsilon_0, \epsilon_1, ..., \epsilon_{n-1}, b) 1_{\{\epsilon_n = b\}} + f_n(\epsilon_0, \epsilon_1, ..., \epsilon_{n-1}, -a) 1_{\{\epsilon_n = -a\}} \\ &= \frac{f_n(\epsilon_0, \epsilon_1, ..., \epsilon_{n-1}, b)}{b} (b 1_{\{\epsilon_n = b\}} - a 1_{\{\epsilon = -a\}}) \\ &= \frac{f_n(\epsilon_0, \epsilon_1, ..., \epsilon_{n-1}, b)}{b} \epsilon_n = \gamma_n \epsilon_n \end{aligned}$$

quand on a posé $\gamma_n = \frac{f_n(\epsilon_0, \epsilon_1, ..., \epsilon_{n-1}, b)}{b}$; cela démontre notre assertion . Considérons maintenant une martingale (N_n) ; appliquant ce qui précède à $F_n = \Delta N_n$, il vient $N_n = N_0 + \sum_1^n \gamma_k \epsilon_k$. Il reste simplement à remarquer que N_0 , qui est $\mathcal{F}_0$-mesurable , est une constante réelle.
Passons à la réciproque ; posons $\mathbf{E}[\epsilon_1^2] = \mathbf{E}[\epsilon_k^2] = \sigma^2$. Comme la suite $(\epsilon_k^2 - \sigma^2; k = 1, 2, ..., N)$ est formée de v.a. indépendantes et centrées , $\sum_1^n (\epsilon_k^2 - \sigma^2)$ est une martingale nulle à l'origine ; par hypothèse , il existe une suite (γ_n) prévisible telle que

$$\sum_{k=1}^{k=n} (\epsilon_k^2 - \sigma^2) = \sum_{k=1}^{k=n} \gamma_k \epsilon_k \quad \forall n \in \{1, 2, ..., N\}$$

Il en résulte que $\forall n \geq 1$, $\epsilon_n^2 - \sigma^2 = \gamma_n \epsilon_n$; puisque σ^2 est > 0, cela entraîne d'abord $\epsilon_n(\omega) \neq 0 \quad \forall \omega$; d'autre part , puisque $\frac{\epsilon_n^2 - \sigma^2}{\epsilon_n}$ est $\mathcal{F}_{n-1}$ -mesurable , on a

$$\frac{\epsilon_n^2 - \sigma^2}{\epsilon_n} = \mathbf{E}[\frac{\epsilon_n^2 - \sigma^2}{\epsilon_n} | \mathcal{F}_{n-1}] = \mathbf{E}[\frac{\epsilon_n^2 - \sigma^2}{\epsilon_n}] = \mathbf{E}[\frac{\epsilon_1^2 - \sigma^2}{\epsilon_1}]$$

Appelant C la constante figurant au dernier membre de cette chaine d'égalités , on voit que $\epsilon_n^2 - C\epsilon_n - \sigma^2 = 0 \ \ \forall n \geq 1$. Il en résulte que ϵ_n prend ses valeurs dans l'ensemble à 2 points constitué par les racines du trinôme $x^2 - Cx - \sigma^2$.

3.4 Temps d'arrêt

Commençons par introduire deux définitions importantes :

Définitions :

Soit $(X_n \ ; \ n \in \{0, 1, 2, ..., N\})$ une suite de v.a. et T une v.a. a valeurs dans $\{0, 1, 2, ..., N\}$, on notera X_T la *variable aléatoire* définie par

$$X_T(\omega) = X_{T(\omega)}(\omega)$$

ou encore

$$X_T = \sum_{n=0}^{n=N} X_n 1_{\{T=n\}}$$

et (X_n^T) la *suite de variables aléatoires* définie par

$$X_n^T = X_n 1_{\{n<T\}} + X_T 1_{\{n\geq T\}}$$

ou encore

$$X_n^T = X_{n\wedge T}$$

On dira que la suite X^T est la suite X *arrêtée a* T .

Parmi les stratégies qui s'offrent à notre joueur du paragraphe **3.3.1.** , il en est une , appelée stratégie d'arrêt qui va nous intéresser maintenant ; elle consiste à miser 1F j'usqu'à un certain temps T (aléatoire) puis, accès de sagesse après un gain ou bouffée dépressive après une perte, à quitter le jeu . Ce faisant, l'évolution (M_n) de la fortune du joueur est remplacée par la suite (M_n^T) ; cette dernière, c'est ce que nous allons voir maintenant, peut être considérée comme une transformée de (M_n) .

Proposition :

Soient $((M_n)\ ;\ 0 \leq n \leq N)$ une suite de v.a. et T une v.a à valeurs dans $\{0, 1, ..., N\}$;
on pose $U_n = 1_{\{n\leq T\}}$; la transformée de (M_n) par la suite (U_n) est (M_n^T)
Démonstration : Il suffit de vérifier, ce qui est facile au vu des définitions, que

$$\Delta(M^T)_n = U_n M_n \quad \forall n \geq 1$$

Nous allons maintenant nous demander àquelle condition (U_n) est prévisible. La réponse est fournie par les équivalences suivantes :

$$\forall n \quad \{n \leq T\} \in \mathcal{F}_{n-1} \Leftrightarrow \{n > T\} \in \mathcal{F}_{n-1} \Leftrightarrow \{T \leq n\} \in \mathcal{F}_n.$$

On remarquera d'autre part que

$$\{T \leq n\} \in \mathcal{F}_n \quad \forall n \quad \Leftrightarrow \quad \{T = n\} \in \mathcal{F}_n \quad \forall n \tag{3.6}$$

ce qui nous conduit a la définition suivante :

Définition :

On appellera temps d'arrêt de la filtration $(\mathcal{F}_n)$ *une variable aléatoire à valeurs dans* $\{0, 1, 2, ..., N\}$ *satisfaisant à l'une des conditions équivalentes* (3.6)
Ce que nous venons de voir rend immédiat le résultat suivant :

3.4.1 Le théorème d'arrêt

Soient $(\Omega, (\mathcal{F}_n), \mathbf{P})$ un espace filtré et T un temps d'arret ; si (M_n) est une $(\mathcal{F}_n)$-martingale (resp. une sur ou une sous-martingale), il en est de même de (M_n^T).

Voici d'autres formes utiles du théorème d'arrêt ; dans ce qui suit, (M_n) desigera une martingale, (U_n) une surmartingale, et T un temps d'arrêt.

1.
$$M_{T\wedge n} = \mathbf{E}[M_T|\mathcal{F}_n] \qquad\qquad M_0 = \mathbf{E}[M_T|\mathcal{F}_0]$$

2.
$$U_{T\wedge n} \geq \mathbf{E}[U_T|\mathcal{F}_n] \qquad\qquad \mathbf{E}[U_{T\wedge n}|\mathcal{F}_0] \geq \mathbf{E}[U_T|\mathcal{F}_0] \tag{3.7}$$

Montrons , par exemple , les inégalités de la deuxième ligne ; puisque (U_n^T) est une surmartingale , $U_n^T \geq \mathbf{E}[U_N^T|\mathcal{F}_n]$, ce qui montre la première inégalité, (et la seconde moyennant un conditionnement supplémentaire relativement a $\mathcal{F}_0$) .

Temps d'entrée :

Soit (X_n)une suite adaptée à une filtration $(\mathcal{F}_n)$ de v.a. à valeurs dans E et soit Γ un sous ensemble de E ; on pose
$$T_\Gamma = \inf\{n \leq N | X_n \in \Gamma\}$$
(avec la convention $\inf \phi = N$)

Proposition :

T_Γ est alors un temps d'arrêt qu'on appelle *temps d'entrée de la suite* (X_n) *dans l'ensemble* Γ

Démonstration :
$\{T = 0\} = \{X_0 \in \Gamma\} \in \mathcal{F}_0$, $\qquad \{T = N\} = \{X_0 \notin \Gamma; X_1 \notin \Gamma; ...; X_{N-1} \notin \Gamma\} \in \mathcal{F}_N$
et pour $0 < n < N$, $\qquad \{T = n\} = \{X_0 \notin \Gamma; ...; X_{n-1} \notin \Gamma; X_n \in \Gamma\} \in \mathcal{F}_n$

Il y a beaucoup d'autres exemples de temps d'arrêt, les plus simples d'entre eux étant les temps d'arrêt constants. On pourra d'autre part vérifier que la borne inférieure, la borne supérieure et la somme de 2 temps d'arrêt sont des temps d'arrêt ; ainsi, $T \wedge 1$ et $T + 1$ sont des temps d'arrêt ; on réfléchira une minute au fait que $T - 1$ n'est pas nécessairement un temps d'arrêt : un joueur peut décider de s'arrêter de jouer après le premier coup qui lui a fait perdre plus de 1000 F mais il lui est malheureusement impossible de le faire à l'instant immédiatement précédent.

3.4.2 Une réciproque au théorème d'arrêt

Proposition :

Soit $((X_n)\ ;\ 1 \leq n \leq N)$ une suite vectorielle adaptée telle que pour toute suite prévisible vectorielle (U_n), on ait $\mathbf{E}[(U * X)_N] = \mathbf{E}[< U_0, X_0 >]$; (X_n) est alors une martingale.

Démonstration : On se ramène aisément au cas unidimensionnel. Si T un temps d'arrêt, la suite $U_n = 1_{\{n\leq T\}}$ est prévisible, et $(U * X)_N = X_T$; on a donc

$$\mathbf{E}[X_T] = \mathbf{E}[X_0] \qquad \forall T$$

Si A est un événement de $\mathcal{F}_n$, on voit facilement que la variable aléatoire $T = n1_A + (n+1)1_{A^c}$ est un temps d'arrêt, pour lequel $X_T = X_n 1_A + X_{n+1}1_{A^c}$; on peut donc écrire

$$\mathbf{E}[X_0] = \mathbf{E}[X_n 1_A + X_{n+1}(1 - 1_A)] = \mathbf{E}[(X_n - X_{n+1})1_A] + \mathbf{E}[X_0] \, .$$

On a donc $\mathbf{E}[X_n 1_A] = \mathbf{E}[X_{n+1}1_A] \quad \forall A \in \mathcal{F}_n$; la caractérisation (**2.3.3**) entraîne alors l'égalité $X_n = \mathbf{E}[X_{n+1}|\mathcal{F}_n]$, ce qui termine la démonstration.

3.5 Un problème d'arrêt optimal

3.5.1 Enveloppe de Snell d'une suite adaptée

Soit $(Z_n \quad 0 \leq n \leq N)$ une suite de v.a.r. adaptée à une filtration $(\mathcal{F}_n)$. Définisssons par récurrence descendante une suite $(U_n \quad 0 \leq n \leq N)$ par les égalités

$$U_N = Z_N \qquad\qquad U_n = \sup(Z_n, \mathbf{E}[U_{n+1}|\mathcal{F}_n]) \quad (0 \leq n < N). \tag{3.8}$$

Il est clair que (U_n) est une suite adaptée ; d'autre part, les inégalités $U_n \geq Z_n$ et $U_n \geq \mathbf{E}[U_{n+1}|\mathcal{F}_n]$ résultant de (24) , montrent que (U_n) est une surmartingale majorant (Z_n) .

1. **Proposition :** (U_n) , qui est appelée *enveloppe de Snell* de (Z_n), est la plus petite surmartingale majorant (Z_n)

   Démonstration : elle se mène par récurrence descendante ; soit (V_n) une surmartingale majorant (Z_n) ; on a $V_N \geq Z_N = U_N$; d'autre part l'hypothèse de récurrence que l'on écrira $V_{n+1} \geq U_{n+1}$ entraine la suite d'inégalités

   $$V_n \geq \mathbf{E}[V_{n+1}|\mathcal{F}_n] \geq \mathbf{E}[U_{n+1}|\mathcal{F}_n] \geq U_n$$

   d'où le résultat .

2. **Proposition :** Posons $T_0 = \inf\{n \geq 0 \mid U_n = Z_n\}$; T_0 est un temps d'arrêt et $(U_n^{T_0})$ est une martingale .

   Démonstration : T_0 est un temps d'arrêt car c'est le temps d'entrée d'une suite adaptée (que nous laissons au lecteur le soin d'expliciter) , dans un sous ensemble de $\mathbb{R}$. D'autre part , $(U_n^{T_0})$ est la suite transformée de la suite (U_n) par la suite prévisible (ϕ_n) définie par $\phi_n = 1_{\{n\leq T_0\}}$. On a donc $\Delta U_n^{T_0} = \phi_n \Delta U_n$, puis $\mathbf{E}[\Delta U_n^{T_0}|\mathcal{F}_{n-1}] = \mathbf{E}[1_{\{n\leq T_0\}}\Delta U_n|\mathcal{F}_{n-1}]$; mais sur l'événement $\{n \leq T_0\} = \{n-1 < T_0\}$, U_{n-1} est strictement supérieur à Z_{n-1} par définition de T_0, de sorte que $U_{n-1} = \mathbf{E}[U_n|\mathcal{F}_{n-1}]$; on a donc

   $$\begin{aligned} \mathbf{E}[\Delta U_n^{T_0}|\mathcal{F}_{n-1}] &= \mathbf{E}[1_{\{n\leq T_0\}}(U_n - \mathbf{E}[U_n|\mathcal{F}_{n-1}])|\mathcal{F}_{n-1}] = 1_{\{n\leq T_0\}}\mathbf{E}[(U_n - \mathbf{E}[U_n|\mathcal{F}_{n-1}])|\mathcal{F}_{n-1}] \\ &= 0 \end{aligned}$$

   ce qui montre la proposition .

3.5.2 Temps d'arrêt optimaux

Reprenons les notations du paragraphe précédent .

1. **Définition :** On dit qu'un temps d'arrêt T est optimal si $\mathbf{E}[Z_T|\mathcal{F}_0] \geq \mathbf{E}[Z_S|\mathcal{F}_0]$ quelque soit le temps d'arrêt S .

   Le plus souvent , l'algèbre $\mathcal{F}_0$ est triviale , si bien que l'inégalité précédente se réduit à $\mathbf{E}[Z_T] \geq \mathbf{E}[Z_S]$; quand la suite (Z_n) représente l'évolution de la fortune d'un joueur , la signification de l'optimalité est tout à fait claire .

2. **Proposition :**Un temps d'arrêt T est optimal si et seulement s'il vérifie les 2 conditions :

   $$a) \quad Z_T = U_T \qquad\qquad b) \quad (U_n^T) \text{ est une martingale}$$

   Démonstration : Soit T un temps d'arrêt verifiant $a)$ *et* $b)$; on peut écrire , quelque soit le temps d'arrêt S , (cf (2.7))

   $$\mathbf{E}[Z_T|\mathcal{F}_0] = \mathbf{E}[U_T|\mathcal{F}_0] = U_0 \geq \mathbf{E}[U_S|\mathcal{F}_0] \geq \mathbf{E}[Z_S|\mathcal{F}_0] \ ,$$

   ce qui montre que T est optimal . Remarquons au passage que T_0 est optimal d'après la proposition 2 du paragraphe précédent ; on a donc $\mathbf{E}[Z_T|\mathcal{F}_0] = \mathbf{E}[Z_{T_0}|\mathcal{F}_0]$.
   Réciproquement ,si T un temps d'arrêt optimal , on peut écrire

   $$\mathbf{E}[U_T|\mathcal{F}_0] \geq \mathbf{E}[Z_T|\mathcal{F}_0] = \mathbf{E}[Z_{T_0}|\mathcal{F}_0] = U_0 \geq \mathbf{E}[U_T|\mathcal{F}_0]$$

   On a donc $\mathbf{E}[Z_T] = \mathbf{E}[U_T]$ et $Z_T \leq U_T$, ce qui entraine $Z_T = U_T$.
   Passons au point (b) : on a

   $$\mathbf{E}[U_T|\mathcal{F}_0] = \mathbf{E}[Z_T|\mathcal{F}_0] = U_0 \geq \mathbf{E}[U_{T\wedge n}|\mathcal{F}_0] \geq \mathbf{E}[U_T|\mathcal{F}_0] = \mathbf{E}[U_T|\mathcal{F}_n|\mathcal{F}_0] \ .$$

   Il en résulte que les v.a.r. $U_{T\wedge n}$ et $\mathbf{E}[U_T|\mathcal{F}_n]$ ont même espérance ; comme elles sont comparables , elles sont égales , ce qui termine la démonstration .

3. **Corollaire :** T_0 est optimal et tout temps d'arrêt optimal est $\geq T_0$.

   Nous avons en effet déjà noté que T_0 était optimal ; le reste provient de ce que , de par sa définition , T_0 est la plus petite v.a. vérifiant (a) .

4. **Remarque :** Désignons par $\mathcal{T}_n$ la famille des temps d'arrêt $\geq n$ et posons

   $$D_n = inf\{p \geq n \ |X_p = U_p\};$$

   on a alors

   $$\mathbf{E}[Z_{D_n}|\mathcal{F}_n] \geq \mathbf{E}[Z_T|\mathcal{F}_n] \quad \forall T \in \mathcal{T}_n.$$

   Nous laissons le lecteur réfléchir à une démonstration .

3.5.3 Compléments

Soient (Z_n) une suite adaptée et (U_n) son enveloppe de Snell ; $\mathcal{T}_n$ et D_n auront la signification donnée dans la remarque ci-dessus . Voici 2 autres façons de calculer l'enveloppe de Snell de (Z_n) .

Proposition

On a les égalités suivantes :

$$(a) \quad U_n = \bigvee_{T \in \mathcal{T}_n} \mathbf{E}[Z_T|\mathcal{F}_n] \quad ; \quad (b) \quad U_n = \mathbf{E}[U_{D_n}|\mathcal{F}_n] = \mathbf{E}[Z_{D_n}|\mathcal{F}_n]$$

Démonstration : Appelons X_n la v.a. figurant au second membre de l' égalité (a).

- Soient T_1 *et* $T_2 \in \mathcal{T}_n$; montrons qu'il existe un troisième temps d'arrêt $T \in \mathcal{T}_n$ tel que
$$\mathbf{E}[Z_T|\mathcal{F}_n] = \mathbf{E}[Z_{T_1}|\mathcal{F}_n] \vee \mathbf{E}[Z_{T_2}|\mathcal{F}_n]$$
Posons pour cela $A = \{\mathbf{E}[Z_{T_1}|\mathcal{F}_n] \geq \mathbf{E}[Z_{T_2}|\mathcal{F}_n]\}$ *et* $T = T_1 1_A + T_2 1_{A^c}$; nous laissons au lecteur le soin de montrer que T vérifie la propriété cherchée .
- Comme l'ensemble $\mathcal{T}_n$ est fini , on en déduit aisément par récurrence qu'il existe un t.a. $R_n \in \mathcal{T}_n$ vérifiant $X_n = \mathbf{E}[Z_{R_n}|\mathcal{F}_n]$.
- Puisque la contante $n \in \mathcal{T}_n$, $X_n \geq Z_n$. D'autre part , on a $\mathbf{E}[X_{n+1}|\mathcal{F}_n] = \mathbf{E}[Z_{R_{n+1}}|\mathcal{F}_n] \leq X_n$, car $R_{n+1} \in \mathcal{T}_n$. (X_n) est ainsi une surmartingale majorant (Z_n) et est de ce fait $\geq (U_n)$.
- Pour achever de montrer que $(X_n) = (U_n)$, notons que , quelque soit $T \in \mathcal{T}_n$, $\mathbf{E}[Z_T|\mathcal{F}_n] \leq \mathbf{E}[U_T|\mathcal{F}_n] \leq U_n$; un passage à la borne supérieure quand T décrit $\mathcal{T}_n$ prouve alors l'inégalité $X_n \leq U_n$.
- Passons aux egalités (b) ; la remarque faite au paragraphe précédent montre que $X_n = \mathbf{E}[Z_{D_n}|\mathcal{F}_n]$; de plus , de par la définition de D_n , $Z_{D_n} = X_{D_n}$, ce qui termine la démonstration .

3.5.4 Enveloppes de Snell et chaines de Markov

Soit $((X_n) \; ; \; 0 \leq n \leq N)$ une chaine de Markov à valeurs dans un epace d'états fini E et admettant P pour matrice de transition (cf **1.4.7**) . Soit $f : \mathbb{N} \times \mathbb{R} \mapsto \mathbb{R}$ une fonction numérique sur E dépendant du temps ; nous allons nous intéresser à la suite adaptée réelle (Z_n) définie par $Z_n = f(n, X_n)$ et à son enveloppe de Snell (U_n) . Nous allons voir que celle ci est aussi une fonction du couple (n, X_n) ; plus précisément :

Proposition

Définissons par récurrence descendante une fonction $u : \mathbb{N} \times E \mapsto \mathbb{R}$ à l'aide des égalités

$$\begin{aligned} u(N,i) &= f(N,i) \quad \forall i \in E \\ u(n,i) &= f(n,i) \vee Pu(n+1,.)(i) \quad (= \sum_{j \in E} p_{ij} u(n+1,j)) \quad \forall i \in E \quad \forall n < N \end{aligned}$$

(U_n) est alors égale à $(u(n, X_n))$.

Démonstration : Elle se mène par récurrence descendante ; si nous supposons vraie l'égalité

$$U_{n+1} = u(n+1, X_{n+1}) \; ,$$

on aura , en vertu de la propriété de Markov , $\mathbf{E}[U_{n+1}|\mathcal{F}_n] = Pu(n+1,.)(X_n)$, d'où

$$U_n = f(n, X_n) \vee Pu(n+1,.)(X_n) = u(n, X_n)$$

3.6 Décomposition de Doob d'une surmartingale

3.6.1 Théorème

Soit (U_n) une surmartingale ; il existe une martingale (M_n) et une suite croissante prévisible $((A_n)\ ,\ n \in \{0, 1, ..., N\}))$ telles que $A_0 = 0$ et

$$(U_n) = (M_n) - (A_n) \tag{3.9}$$

Une telle décomposition (appelée *décomposition de Doob*) est unique .

Démonstration : Pour établir l'unicité , montrons que la décomposition (2.9) détermine complètement (A_n) , (et par conséquent (M_n)) ; à l'origine , tout est facile : $A_0 = 0$ et $M_0 = U_0$, si bien qu'il suffit de déterminer les accroissements de (A_n) ; or , en conditionnant les 2 membres des égalités $\Delta U_n = \Delta M_n - \Delta A_n$ par $\mathcal{F}_{n-1}$, on obtient pour $n \geq 1$

$$\Delta A_n = -\mathbf{E}[\Delta U_n|\mathcal{F}_{n-1}] \tag{3.10}$$

d'où l'unicité ; l'existence se prouve en définissant (A_n) par la formule (2.10) ; la suite ainsi construite est manifestement prévisible ; de plus , ΔA_n est ≥ 0 puisque (U_n) est une surmartingale ; (A_n) est donc une suite croissante . Il nous reste à montrer que la suite (M_n) définie par $M_n = U_n + A_n$ est une martingale ; écrivons pour cela
$\Delta M_n = \Delta U_n - \mathbf{E}[\Delta U_n|\mathcal{F}_{n-1}]$ et conditionnons les 2 membres : il vient $\mathbf{E}[\Delta M_n|\mathcal{F}_{n-1}] = 0$, C.Q.F.D.

3.6.2 Un deuxième temps d'arrêt optimal

Reprenons les notations de **2.5** et effectuons la décomposition de Doob , $U_n = M_n - A_n$ de l'enveloppe de Snell de (Z_n) ; posons ensuite

$$T_1 = \inf\{n \mid A_{n+1>0}\}1_{\{A_N>0\}} + N1_{\{A_N=0\}}$$

On vérifie aisément , grâce à la prévisibilité de (A_n) , que T_1 est un temps d'arrêt ; de plus ,

Proposition :

T_1 est optimal ; il n'existe pas d'autre temps d'arrêt optimal $\geq T_1$.

Démonstration :Nous allons utiliser la caractérisation **2.5.2**.2 des temps d'arrêt optimaux .

1. $(U_n^{T_1})$ *est une martingale* : par définition de T_1 , on a les implications

   $$n \leq T_1 \Longrightarrow A_n = 0 \Longrightarrow U_n = M_n \ .$$

   Il en résulte que $U_{n\wedge T_1} = M_{n\wedge T_1} \quad \forall n$, ce qui s'écrit encore $(U_n^{T_1}) = (M_n^{T_1})$, d'où le résultat .

2. $U_{T_1} = Z_{T_1}$: On a $U_{T_1} = \sum_0^N U_n 1_{\{T_1=n\}} = \sum_0^{N-1} U_n 1_{\{T_1=n\}} + Z_N 1_{\{T_1=N\}}$. Plaçons nous , pour $n \leq N-1$, sur l'événement $\{T_1 = n\}$, sur lequel $A_n = 0$ (et donc $U_n = M_n$) , $A_{n+1} > 0$ et où l'on peut donc écrire

$$U_n = Z_n \vee \mathbf{E}[U_{n+1}|\mathcal{F}_n] = Z_n \vee (M_n - A_{n+1}) = Z_n \vee (U_n - A_{n+1}) .$$

Mais comme $U_n - A_{n+1} < U_n$, on a nécessairement $U_n = Z_n$; il en résulte que

$$U_{T_1} = \sum_0^N Z_n 1_{\{T_1=n\}} = Z_{T_1}.$$

3. Il nous reste à montrer le caractère maximal de T_1 : Si S un temps d'arrêt tel que $S \geq T_1$ *et* $\mathbf{P}[S > T_1] > 0$, on aura $\mathbf{E}[U_S] = \mathbf{E}[M_S - A_S] < \mathbf{E}[M_S] = \mathbf{E}[M_0] = \mathbf{E}[U_0]$, si bien que (U_n^S) ne peut pas être une martingale ; S n'est donc pas optimal , C.Q.F.D.

Chapter 4

Mathématiques des marchés financiers

4.1 Dictionnaire

- Nous appellerons *marché financier* un espace filtré fini $(\Omega, (\mathcal{F}_n), \mathbf{P} \ ; \ n = 0, 1, 2, \cdots, N)$ satisfaisant aux hypothèses suivantes :

$$\mathcal{F}_0 = \{\Omega, \phi\} \quad ; \quad \mathcal{F}_N = \mathcal{F} = \mathcal{P}(\Omega) \quad ; \quad \mathbf{P}(\omega) > 0 \quad \forall \omega \in \Omega$$

  N sera appelé *l'horizon du marché* ; ce sera le plus souvent la *date d'exercice* des options .
- *Actifs financiers* : On supposera qu'il existe sur ce marché $d+1$ *actifs financiers* notés $0, 1, 2, \cdots, d$; l'actif 0 jouera un rôle particulier ; il représentera en effet un *actif sans risque* (par exemple , un livret de caisse d'épargne ou un bond du trésor a taux fixe) ; les d actifs suivants, qu'on peut se figurer comme des actions, seront appelés *actifs risqués.*
- *Les prix ou cours des actifs* à l'instant n seront représentés par $d+1$ variables aléatoires réelles positives (le cours d'une action n'est jamais négatif) $S_n^0, S_n^1, \cdots, S_n^d$ adaptées à $(\mathcal{F}_n)$; le vecteur aléatoire à valeurs dans $\mathbb{R}_+^{d+1}$ ainsi défini sera noté S_n .
- *Le taux d'intérêt* sera un réel $r > 0$ gouvernant le prix de l'actif sans risque : on supposera dans toute la suite $S_0^0 = 1$; on posera donc $S_n^0 = (1+r)^n$.
- *Actualisation* : La valeur de la monnaie se déprécie au cours du temps ; pour pouvoir comparer le prix d' actifs observés à des dates différentes , on utilise souvent la valeur qu'avait la monnaie à l'instant 0 ; un prix ainsi calculé sera appelé *prix actualisé* . En supposant que le taux de dépréciation monétaire est égal au taux d'intérêt r , on introduira un *coefficient d'actualisation* $\beta_n = \frac{1}{S_n^0} = \frac{1}{(1+r)^n}$. Si P_n est le prix d'un actif observé à l'instant n son *prix actualisé* sera , par définition , $\tilde{P}_n = \beta_n P_n$, (le prix actualisé , calculé dans une monnaie plus forte , est inférieur au prix courant) . Voici une autre façon de se représenter le coefficient d'actualisation : β_n est la quantité d'argent qu'il faut investir a l'instant 0 dans l'actif sans risque pour disposer de 1F a l'instant n ; on notera que $\tilde{S}_n^0 = 1 \quad \forall n.$
- *Stratégies de gestion et portefeuilles* : Imaginons un investisseur possédant un portefeuille à l'instant n comportant une quantité ϕ_n^k de l'actif k ; au vu des informations dont il dispose à cet instant , (étude de l'evolution des cours) , il va décider de modifier sa composition . Mais il nous faut préciser ce que l'on entend exactement par là. Supposons que S_n représente le cours des actifs à la clôture de la séance boursière de la date n . La composition du portefeuille ϕ_n à

cet instant a été décidée alors que ce cours n'etait pas connu : la seule information disponible était l'historique des cours jusqu'a la date $n-1$. On se retrouve donc dans la même situation que notre vieil ami rencontré naguère (cf **3.3.1**) dans un casino à qui l'on avait imposé que la suite formée par ses mises successives soit prévisible . Cela conduit à la définition suivante :
Nous appellerons *stratégie de gestion* ou encore *portefeuille* une suite *prévisible* $(\phi_n \ , \ n \in \{1, ..., N\})$ de variables aléatoires à valeurs dans $\mathbb{R}^{d+1}$. On complétera la définition de ϕ à la date 0 en posant $\phi_0 = \phi_1$
On remarquera que , (contrairement aux cours des actifs), ces v.a.r. peuvent etre négatives ; cela tient à ce que plusieurs techniques boursières (ventes à découvert, prises en pension d'actifs), permettent d'avoir des dettes libellées en titres au lieu de l'être en numéraire comme le sont, je l'imagine, les vôtres.

– *Valeur d'un portefeuille* : Dénotant par $<,>$ le produit scalaire usuel sur $\mathbb{R}^{d+1}$, la valeur à l'instant n d'un portefeuille associé a la stratégie (ϕ_n) sera, par définition,
$$V_n(\phi) =< \phi_n, S_n >= \sum_{k=0}^{k=d} \phi_n^k S_n^k \tag{4.1}$$
Cette définition se passe de commentaires ; on notera toutefois que $(V_n(\phi))$ constitue une suite adaptée de v.a.r. .
La *valeur actualisée* de ce portefeuille sera donc
$$\tilde{V_n}(\phi) = \beta_n V_n(\phi) = \beta_n < \phi_n, S_n >=< \phi_n, \tilde{S_n} > .$$

4.2 Sratégies autofinancées

On appellera *stratégie autofinancée* une suite prévisible vectorielle (ϕ_n) vérifiant
$$< \phi_n, S_n >=< \phi_{n+1}, S_n > \quad \forall n < N \tag{4.2}$$
Cela correspond au comportement d'un investisseur réorganisant son portefeuille entre la date n et la date $n+1$, (c'est à dire remplacant ϕ_n par ϕ_{n+1}), et faisant en sorte que ce réaménagement conserve la valeur totale du portefeuille. Il s'interdit donc tout prélèvement destiné à sa consommation ainsi que tout apport d'argent frais.

4.2.1 Proposition

La stratégie (ϕ_n) est autofinancée si et seulement si, quelque soit $n \geq 1$,
$$V_n(\phi) =< \phi_0, S_0 > + \sum_{k=1}^{k=n} < \phi_k, \Delta S_k > . \tag{4.3}$$
On notera que cette égalité est équivalente à l'égalité "actualisée"
$$\tilde{V_n}(\phi) =< \phi_0, S_0 > + \sum_{k=1}^{k=n} < \phi_k, \Delta \tilde{S}_k > . \tag{4.4}$$
Démonstration : Pour $n \geq 1$, on a $< \phi_n, \Delta S_n >=< \phi_n, S_n > - < \phi_n, S_{n-1} >$ de sorte que (4.2) est équivalente aux égalités $< \phi_n, \Delta S_n >= \Delta V_n(\phi)$, c'est à dire à (4.3) .

Remarque : On peut donner une forme plus condensée a (4.3) et (4.4) ; il suffit pour cela d'étendre la notion de suite transformée au cas vectoriel : revenons à la définition **3.3.1.** en supposant maintenant que (M_n) et (U_n) sont à valeurs dans un espace Euclidien ; l'égalité

$$(U * M)_n = U_0 M_0 + \sum_{k=1}^{k=n} < U_k, \Delta M_k >$$

définit une suite *réelle* caractérisée par les relations

$$(U * M)_0 =< U_0, M_0 > \qquad\qquad \Delta(U * M)_n =< U_n, \Delta M_n >$$

Avec ces notations , (4.3) et (4.4) s'écrivent respectivement

$$V_n(\phi) = (\phi * S)_n \qquad\qquad \tilde{V}_n(\phi) = (\phi * \tilde{S})_n \tag{4.5}$$

La proposition suivante va nous montrer qu'il suffit de connaitre la gestion des actifs à risque d'un investisseur et la valeur initiale de son portefeuille pour en déduire la totalité de sa stratégie de gestion.

4.2.2 Proposition :

Soient V_0 une v.a.r. $\mathcal{F}_0$-mesurable (i.e. une constante) et $((\phi_n^1, \phi_n^2, ..., \phi_n^d) \ ; n \in \{0, 1, ..., N\})$ une suite prévisible à valeurs dans $\mathbb{R}^d$; il existe stratégie autofinancée $\phi = (\phi_n^0, \phi_n^1, \phi_n^2, ..., \phi_n^d)$ unique telle que $V_0(\phi) = V_0$.

Démonstration : Si $u = (u^0, u^1, ..., u^d)$ *et* $v = (v^0, v^1, ..., v^d) \in \mathbb{R}^{d+1}$, nous introduirons la notation

$$[u, v] = \sum_{k=1}^{k=d} u^k v^k =< u, v > -u^0 v^0. \tag{4.6}$$

Si ϕ est une stratégie de gestion satisfaisant aux conditions de l'énoncé, les relations (4.2) d'éfinissant l'autofinancement sont équivalente aux égalités

$$< \Delta\phi_n, \tilde{S}_{n-1} >= 0 \qquad \forall n \geq 1$$

qui s'écrivent encore

$$\Delta\phi_n^0 = -[\Delta\phi_n, \tilde{S}_{n-1}] \tag{4.7}$$

On aura en outre

$$V_0 = V_0(\phi) = \phi_0^0 + [\phi_0, S_0] \tag{4.8}$$

ou encore,

$$\phi_0^0 = V_0 - [\phi_0, S_0] \tag{4.9}$$

Les égalités (4.7) et (4.9) montrent que la suite (ϕ_n^0) est complètement déterminée par les suites $(\phi_n^k \ , \ k \geq 1)$, ce qui montre l'unicité de ϕ. Pour montrer son existence, revenons aux égalités (4.7) et (4.9) qui nous permettent de définir une suite (ϕ_n^0) et une stratégie de gestion $\phi = (\phi_n^0, \phi_n^1, ..., \phi_n^d)$. Il est clair que ϕ est autofinancée ; il reste à montrer que ϕ_n^0 est prévisible, ce qui résulte facilement de (4.7).

Corollaire :

L'ensemble $\mathcal{V} = \{V_N(\phi)\} = \{\tilde{V}_N(\phi)\}$, ϕ parcourant l'ensemble des stratégies autofinancées est un sous espace vectoriel de l'espace vectoriel $\mathcal{E}$ des v.a.r sur Ω.

On a en effet

$$\tilde{V}_N(\phi) = \{c + \sum_{n=1}^{n=N} < \psi_n, \Delta\tilde{S}_n >; c \in \mathbb{R} \ , \ \psi \in \mathcal{P}^d\}$$

quand $\mathcal{P}^d$ désigne l'espace vectoriel des suites de v.a. prévisibles $\{(\psi_n^1), (\psi_n^2), ..., (\psi_n^d)\}$ à valeurs dans $\mathbb{R}^d$

4.2.3 Stratégies d'arbitrage

On dira qu'une stratégie ϕ autofinancée est une *stratégie d'arbitrage* ou plus brièvement un *arbitrage* si elle vérifie les conditions suivantes

$$V_0(\phi) = 0 \ , \ V_N(\phi) \geq 0 \ , \ \mathbf{P}[V_N(\phi) > 0] > 0 \tag{4.10}$$

Une telle stratégie permet donc à un investisseur totalement démuni au temps 0 de n'avoir à coup sûr aucune dette à l'instant N tout en gagnant de l'argent avec une probabilité stictement positive. Dans un marché financier réel, cette possibilité est en principe interdite par les autorités de régulation (elles répriment en particulier les "délits d'initié"). L'axiome de viabilité que nous allons maintenant introduire est donc d'une grande importance. C'est à cette condition que notre espace probabilisé muni d'une suite de v.a. positives pourra servir de modèle à un marché financier.

4.3 Marchés viables

4.3.1 Définition :

On dira qu'un marché financier est *viable* s'il n'existe pas dans ce marché de stratégie d'arbitrage .

L'idée qu'un marché, (on n'ose dire une société), fonctionne efficacement s'il n'existe pas de stratégie d'arbitrage joue un rôle fondamental. Il y a pour cela des raisons techniques : si le cours d'une action a une évolution qui rend possible une stratégie d'arbitrage, elle va faire l'objet d'achats massifs, ce qui aura pour effet de faire monter son cours faisant ainsi s'évanouir les espoirs de gains sans risque qu'elle procurait. On peut donc admettre "qu'en régime stationnaire" , tout marché est viable . Notons que ce retour à la viabilite est d'autant plus rapide que la spéculation est plus importante ; celle ci, au lieu d'apparaitre comme une hydre malfaisante, (sous la terreur, la spéculation sur les grains conduisait les "affameurs" à la lanterne), devient un mal nécessaire permettant la régulation du marché. Allons un peu plus loin : ce qui est interdit dans ce qu'on appelle souvent une "écononomie casino", ce n'est pas le fait de "s'enrichir en dormant" qui indignait naguère un président de la république , c'est celui de s'enrichir sans risque, ce que nous appelons ici un arbitrage . Dans la nouvelle morale, fille de la nouvelle économie, la vertu la plus célébrée par les médias n'est d'ailleurs plus l'ardeur au travail mais le goût du risque .

Exercice :

Les 3 assertions suivantes sont équivalentes

1. Il existe une stratégie autofinancée ϕ telle que

$$V_0(\phi) \leq 0 \quad ; \quad V_N(\phi) \geq 0 \quad ; \quad \mathbf{P}[V_N(\phi) > 0] > 0. \tag{4.11}$$

2. Il exite un entier $p \in \{1, 2, \ldots, N\}$ et un vecteur aléatoire $\mathcal{F}_{p-1}$ – mesurable $\eta = (\eta^1, \eta^2, \ldots, \eta^d)$ à valeurs dans $\mathbb{R}^d$ tels que

$$[\eta, \Delta\tilde{S}_p \geq 0] \quad ; \quad \mathbf{P}\big[[\eta, \Delta\tilde{S}_p > 0]\big] > 0. \tag{4.12}$$

3. Il existe une stratégie d'arbitrage. (Autrement dit le marché n'est pas viable.)
4. Il existe une stratégie d'arbitrage ϕ vérifiant $V_n(\phi) \geq 0 \quad \forall n \leq N.$[1]
   Démonstration : Notons d'abord que 4. $\Longrightarrow$ 3. $\Longrightarrow$ 1. ; on procède ensuite comme suit :
   (a) 1. $\Longrightarrow$ 2. : Soit $\phi = (\phi^0, \psi)$ un portefeuille autofinancé vérifiant (4.11) ; posons

$$u(n) = \mathbf{P}[V_n(\phi) \geq 0] \quad ; \quad v(n) = \mathbf{P}[V_n(\phi) > 0] \quad ; \quad p = \inf\{n \mid u(n) = 1 \text{ et } v(n) > 0\}$$

   Il résulte des hypothèses faites sur ϕ que $1 \leq p \leq N$; d'autre part, $\forall n < p$ (et en particulier pour $n = p - 1$), on a soit $v(n) = 0$, soit $u(n) < 1$.

   i. Si $v(p-1) = 0$, nous poserons $\eta = \psi_p$; on aura alors, puisque ϕ est autofinancée,

$$[\eta, \Delta\tilde{S}_p] =< \phi_p, \Delta\tilde{S}_p >= \tilde{V}_p(\phi) - \tilde{V}_{p-1}(\phi) \geq \tilde{V}_p(\phi)$$

   ce qui entraîne (4.12)

   ii. Supposons maintenant $u(p-1) < 1$; nous poserons alors $\eta = \psi_p 1_{\{V_{p-1}(\phi) < 0\}}$. On peut alors écrire

$$[\eta, \Delta\tilde{S}_p] = 1_{\{V_{p-1}<0\}} < \phi_p, \Delta\tilde{S}_p >= 1_{\{V_{p-1}<0\}}[\tilde{V}_p(\phi) - \tilde{V}_{p-1}(\phi)] \geq 1_{\{V_{p-1}<0\}}\big(-\tilde{V}_{p-1}(\phi)\big)$$

   ce qui entraîne également (4.12)

   (b) 2. $\Longrightarrow$ 4. : Soit η un vecteur aléatoire vérifiant (4.12) ; définissons la suite prévisible de vecteurs aléatoires (ψ_n) à valeurs dans $\mathbb{R}^d$ par

$$\psi_p = \eta \quad \text{et} \quad \psi_n = 0 \quad \text{pour} \quad n \neq p$$

   et appelons (ϕ_n) le portefeuille autofinancé de valeur initiale nulle associé à (ψ_n) par la proposition **4.2.2** . On a

$$\tilde{V}_m(\phi) = \sum_{1 \leq n \leq m} < \phi_n, \Delta\tilde{S}_n >= \sum_{1 \leq n \leq m} [\psi_n, \Delta\tilde{S}_n],$$

   si bien que $\tilde{V}_m(\phi)$ est nul pour $m < p$ et vaut $[\eta, \Delta\tilde{S}_p]$ pour $m \geq p$. Le résultat en découle immédiatement.

[1] Une telle stratégie est parfois appelée *stratégie admissible.*

4.3.2 Une version du théorème de Hahn-Banach

Pour aborder l'étude des relations entre la viabilité d'un marché et la théorie des martingales, nous aurons besoin d'un résultat d'analyse, connu sous le nom de *théorème de Hahn-Banach.* Nous nous bornerons à en démontrer une version affaiblie qui sera suffisante pour la suite ; dans ce paragraphe, $(E, <.,.>)$ désignera un espace euclidien quelconque .

Théorème

Soient V un sous espace vectoriel et K une partie convexe compacte de E ne rencontrant pas V ; il existe alors un vecteur a de E vérifiant

$$< a, x >> 0 \quad \forall x \in K \qquad\qquad < a, y >= 0 \quad \forall y \in V$$

Démonstration : Notons tout d'abord que V, sous espace d'un espace vectoriel de dimension finie, est fermé . Posons $C = K+V = \{x+y \mid x \in K \ ; \ y \in V\}$; nous ferons la démonstration en 2 étapes.

1. C est fermé, convexe et ne contient pas 0 ; nous nous bornerons à montrer le premier point, les autres étant faciles ; soit $z_n = x_n + y_n \quad (x_n \in K \ ; \ y_n \in V)$ une suite convergente de points de C ; il s'agit de montrer que la limite z de la suite (z_n) est dans C ; puisque K est compact, il existe une sous suite (x_{n_k}) de (x_n) qui converge vers un vecteur x de K ; $y_{n_k} = z_{n_k} - x_{n_k}$ converge alors vers $z - x$, lequel appartient a V puisque V est fermé ; il ne reste qu'à écrire $z = x+(z-x)$ pour prouver que $z \in C$.

2. D'après un résultat classique sur les espaces euclidens, il existe dans C un point a unique de norme minimale. Soit x un vecteur quelconque de C et $\theta \in]0,1]$; puisque C est convexe, $a + \theta(x-a) \in C$, de sorte que $\|a\|^2 \leq \|a + \theta(x-a)\|^2$; développant le second membre , on obtient les inégalités

   $$2 < a, x-a > +\theta\|x-a\|^2 \geq 0 \quad \forall\theta \in]0\ 1] \ .$$

   Faisant tendre θ vers 0 en décroissant, il vient $< a, x >\geq \|a\|^2 > 0 \quad \forall x \in C$.
   Il nous reste à montrer que la forme lineaire $< a,. >$ est nulle sur V ;
   soient $x \in K$, $y \in V$, $\lambda \in \mathbb{R}$; $x + \lambda y$ est alors dans C, de sorte que
   $< a, x + \lambda y >\geq \|a\|^2$; la fonction affine $\lambda \mapsto < a, x > +\lambda < a, y >$, définie sur $\mathbb{R}$, ne peut être uniformément minorée que si elle est constante ; on a donc $< a, y >= 0$, ce qui termine la démonstration.

4.3.3 Marchés viables et martingales

Nous nous proposons dans ce paragraphe de montrer le resultat suivant :

Théorème :

Un marché financier est viable si et seulement s'il existe une probabilité $\mathbf{P^*}$ admettant Ω comme support sous laquelle les cours actualisés des actifs $((\tilde{S}_j^k) \quad 0 \leq j \leq N)$ sont des martingales.

Démonstration :

a) Supposons qu'il existe une probabilité $\mathbf{P^*}$ de support Ω sous laquelle les cours actualisés des actifs $(\tilde{S}_j^k \quad 0 \leq j \leq N)$ sont des martingales. Soit $\phi = (\phi_j^k)$ une stratégie autofinancée telle que $V_0(\phi) = 0$; puisque ϕ est autofinancée, on a, quelque soit $n \geq 1$,

$$\tilde{V}_n(\phi) = \sum_{i=1}^{i=n} < \phi_i \ , \ \Delta\tilde{S}_i >= \sum_{k=0}^{k=d} (\phi^k * \tilde{S}^k)_n .$$

Sous $\mathbf{P^*}$, $(\tilde{V}_n(\phi))$ est une martingale nulle à l'origine ; elle vérifie donc $\mathbf{E^*}[\tilde{V}_N(\phi)] = 0$; il est alors clair que $\tilde{V}_N(\phi)$ ne peut pas être une v.a. ≥ 0 non nulle.

b) Réciproquement, supposons le marché viable ; nous allons appliquer le théorème de Hahn-Banach dans la situation suivante :

– E sera l'espace euclidien des v.a.r. sur Ω muni du produit scalaire

$$< X, Y >= \sum_{\omega \in \Omega} X(\omega)Y(\omega) \ .$$

– $\mathbf{V}$ désignera le sous ensemble de E formé par les v.a.r. $\tilde{V}_N(\phi)$ quand ϕ décrit l'ensemble $\mathbf{S}_0$ des statégies autofinancées telles que $V_0(\phi) = 0$. Il est facile de voir que $\mathbf{V}$ est un sous espace vectoriel de E.
– On notera $\mathbf{K}$ le sous ensemble de E formé des v.a.r. X positives vérifiant $\sum_{\omega \in \Omega} X(\omega) = 1$; nous laissons au lecteur le soin de vérifier que $\mathbf{K}$ est convexe et compact. On remarquera qu'une v.a. X de $\mathbf{K}$ vérifie $[X > 0] \neq \phi$; puisqu'il n'existe pas de stratégie d'arbitrage, $\mathbf{V}$ et $\mathbf{K}$ sont disjoints.

Appliquons alors le théorème de Hahn-Banach à $\mathbf{V}$ et $\mathbf{K}$: il existe une v.a.r. α telle que

$$\sum_{\omega \in \Omega} \alpha(\omega)X(\omega) > 0 \quad \forall X \in \mathbf{K} \quad ; \quad \sum_{\omega \in \Omega} \alpha(\omega)\tilde{V}_N(\phi)(\omega) = 0 \quad \forall \phi \in \mathbf{S}_0 \qquad (4.13)$$

On notera que la première inégalité entraine $\alpha(\omega) > 0 \ \ \forall \omega$ (considérer les v.a. X qui sont les indicateurs des points). Posons alors

$$\| \alpha \| = \sum_{\omega \in \Omega} \alpha(\omega) \qquad ; \qquad \mathbf{P^*} = \frac{\alpha}{\| \alpha \|} \ .$$

$\mathbf{P^*}$ est une probabilité chargeant tous les points de Ω ; avec cette notation la deuxième égalité de (4.13) s'écrit

$$\mathbf{E^*}[\tilde{V}_N(\phi)] = 0 \quad \forall \phi \in \mathbf{S}_0 \ . \qquad (4.14)$$

Nous allons montrer que $(\tilde{S}_n)$ est une martingale sous **P*** en utilisant la caractérisation **3.4.2** ; à cet effet, donnons nous une suite vectorielle prévisible

$$(\psi_n) = (\psi_n^k) \ , \ k = 1, 2, \ldots, d \ , \ \ n = 0, 1, 2, \ldots, N$$

à valeurs dans $\mathbb{R}^d$ et désignons par (ϕ_n), la stratégie de $\mathbf{S}_0$ associée à ψ en application du théorème **4.2.2** (avec $V_0 = 0$). Notons (σ_n) la suite vectorielle, à valeurs dans $\mathbb{R}^d$ des actifs risqués actualisés et [,] le produit scalaire usuel dans $\mathbb{R}^d$. Un calcul simple conduit à l'égalité

$$(\psi * \sigma)_N = [\psi_0 \ , \ \sigma_0] + \tilde{V}_N(\phi)$$

Prenons l'espérance des 2 membres sous **P***; il vient, compte tenu de (4.14), $\mathbf{E^*}[(\psi * \sigma)_N] = \mathbf{E^*}[\psi_0 \ , \ \sigma_0]$ ce qui montre, (Cf.**3.4.2.**), que (σ_n) est une **P***-martingale . Il reste à remarquer que l'actif sans risque actualisé $(\tilde{S}_n^0)$ est une martingale triviale.

4.4 Options européennes

Rappelons d'abord brièvement les mécanismes boursiers (déjà décrits dans l'introduction) à l'origine des 2 options européennes les plus importantes : les *options d'achat* (en franglais *call*) et les *options de vente* (ou *put*).

Le possesseur d'un call européen de *prix d'exercice* $K > 0$ et de *date d'exercice* N émis par une banque sur l'actif à risque *sous-jacent* l se trouve investi du droit d'acheter à la date N l'actif l au prix K. Examinons de quelle façon il peut monnayer ce droit à la date d'exercice.

- Si $S_N^l > K$, Il achète l'actif l à la banque au prix K (on dit alors qu'il *exerce son option*) puis s'empresse de revendre ce même actif sur le marché au prix S_N^l ; la valeur de l'option à la date N est le bénéfice que lui procure cette opération, à savoir $S_N^l - K$.
- Si, au contraire, $S_N^l < K$, le droit procuré par la détention du call n'a plus aucun intérêt puisqu'on peut se procurer sur le marché l'actif sous-jacent à un prix $< K$. Dans ce cas, la valeur de l'option d'achat à la date N est nulle.

En résumé, la valeur ou le prix du call *à sa date d'exercice* N sera $(S_N^l - K)^+$.

Une *option de vente* européenne (ou *put* européen) de prix d'exercice K et de date d'exercice N sur l'actif sous -jacent l permet à son détenteur de *vendre* l á la date N au prix K. Par un raisonnement en tout point similaire au précédent, on conclut que le prix de ce put à la date N est $(K - S_N^l)^+$.
Par abut de langage on identifie souvent une option européenne à sa valeur à la date d'exercice ce qui explique la définition suivante :

4.4.1 Définition :

Nous appellerons *option européenne d'échéance* (ou encore de *date d'exercice*) N une variable aléatoire $h \geq 0$; la variable aléatoire $\tilde{h} := \beta_N h$ sera appelée *option actualisée* associée à h.

Exemples

Soit (U_n) la suite des cours d'un actif risqué u

1. Une *option d'achat* h (ou *call*) européenne de *prix d'éxercice* K et de *date d'exercice* N sur l'actif *sous-jacent* u sera définie par $h = (U_N - K)^+$.
2. On d'éfinit de la même façon une *option de vente* h (ou *put*) européenne par $h = (K - U_N)^+$.
3. On peut remplacer la v.a. U_N par la valeur moyenne $U_I = \frac{1}{|I|}\sum_{n\in I} U_n$ de l'actif (U_n) sur un intervalle de temps $I \subset \{0, 1, 2, \ldots, N\}$ prédéterminé ; les options construites de cette façon prennent le nom d'*options asiatiques*. Par exemple un put (européen) asiatique h de prix d'exercice K sur l'intervalle I sera défini par $h = (K - U_I)^+$.
4. Soient a et b deux réels > 0 avec $a < U_0 < b$; l'option européenne $h = U_N \vee a \wedge b$ est émise par certaines institutions financières sous le nom de *placement garanti* ; un investisseur trouvant trop risqué l'achat de l'actif u préfèrera celui de h. Il sera ainsi certain de limiter ses pertes au niveau a quitte à accepter de limiter ses gains au niveau b.

4.4.2 Les prix excluant l'arbitrage (*arbitrage-free prices ; p.e.a.*) d'une option européenne dans un marché viable

En pratique, les options européennes sont mises sur le marché et se négocient à tout instant à un prix qui équilibre l'offre et la demande. On peut donc les considérer comme des actifs à risque supplémentaires du marché financier. Nous allons considérer un prix comme raisonnable s'il n'introduit pas d'opportunité d'arbitrage, ce qui nous conduit à la définition suivante :

Définition

Soit h une option européenne dans un marché financier viable ; s'il existe une suite adaptée (H_n) de v.a. ≥ 0 telle que

– $H_N = h$

– le marché financier $(\Omega, (S_n), (H_n), \mathbf{P})$ soit encore viable,

on dira que H_0 est un *prix excluant l'arbitrage* de l'option h .
H_n sera appelée p.e.a. à l'instant n de l'option h.

Avant de discuter l'existence et l'unicité d'une telle suite, nous allons voir, sur l'exemple important qui suit, que cette exigence de viabilité introduit des contraintes fortes sur le prix des call et des put.

4.4.3 Relation de parité call-put dans un marché viable

Soient $(\Omega, (S_n), \mathbf{P})$ un marché financier viable et $l \in \{1, 2, \ldots, N\}$ un actif risqué. Une institution financière émet sur l'actif sous-jacent l un call européen $C = (S^l_N - K)^+$ ainsi qu'un put européen $P = (K - S^l_N)^+$ de même prix d'exercice $K > 0$. Soit (C_n) (resp.(P_n)) une suite de p.e.a. de l'option d'achat C (resp. de l'option de vente P)

Proposition :

On a les égalités

$$C_n - P_n = S_n^l - \frac{K}{(1+r)^{N-n}} \qquad \forall n \leq N \tag{4.15}$$

Démonstration : Notre marché comporte maintenant les $d+3$ actifs $\{0, 1, \ldots, d, c, p\}$ admettant pour cours à l'instant n $\bar{S_n} = S_n, C_n, P_n$ et est supposé viable. Nous allons montrer que, si l'égalité (4.15) n'est pas satisfaite, il existe une stratégie d'arbitrage ; pour ne pas trop alourdir ces notes, nous nous bornerons au cas $n = 0$ en laissant au lecteur, que je sais friand d'exercices, le soin d'en déduire une démonstration générale.

1. Supposons $C_0 - P_0 > S_0^l - \frac{K}{(1+r)^N}$.

   Soit $((\psi_n^k)\ ;\ 0 \leq n \leq N\ ;\ k \in \{1, 2, \ldots d, c, p\})$ la suite vectorielle prévisible à valeurs dans $\mathbb{R}^{d+2}$ définie par

   $$\psi_n^l = 1\ ;\ \psi_n^c = -1\ ;\ \psi_n^p = 1\ \ \forall n \geq 0\ \ ;\ \ \psi_n^k = 0\ \ \forall n \geq 0\ \ \forall k \notin \{l, c, p\}$$

   D'après la proposition **4.2.2**, il existe une statégie autofinancée ϕ coïncidant avec ψ sur les actifs risqués et telle que $V_0(\phi) = 0$.
   On a donc

   $$V_n(\phi) = S_n^l - C_n + P_n + \phi_n^0 S_n^0 \quad \forall n \geq 0 \tag{4.16}$$

   et en particulier,

   $$S_0^l - C_0 + P_0 + \phi_0^0 = 0 \tag{4.17}$$

   Puisque ϕ est autofinancée, on a d'autre part

   $$V_n(\phi) =< \phi_{n+1}, \bar{S_n} >= S_n^l - C_n + P_n + \phi_{n+1}^0 S_n^0 \tag{4.18}$$

   La comparaison de (4.16) et (4.18) montre que ϕ_n^0 ne dépend pas de n ; il résulte alors de (4.17) que

   $$\begin{aligned} \phi_n^0 &= \phi_0^0 = C_0 - P_0 - S_0^l & (4.19) \\ V_n(\phi) &= (C_0 - P_0 - S_0^l)(1+r)^n + S_n^l - C_n + P_n & (4.20) \\ V_N(\phi) &= (C_0 - P_0 - S_0^l)(1+r)^N + K & (4.21) \end{aligned}$$

   La stratégie ϕ constitue donc un arbitrage ce qui interdit au marché d'être viable.

   Dans une salle de marché , la manière pratique de mettre en oeuvre cette stratégie serait la suivante : à l'instant 0 vous achetez une action et un put et vous vendez un call ; cela vous procure une quantité (algébrique) d'argent liquide égale à $C_0 - P_0 - S_0^l$ que vous placez au taux r si elle est > 0, et que vous empruntez au même taux si elle est < 0 ; vous ne touchez à rien jusqu'à la date N ; votre quantité de numéraire est alors $(C_0 - P_0 - S_0^l)(1+r)^N$; deux cas peuvent maintenant se présenter :

   (a) $S_N^l > K$; l'acheteur du call l'exerce, vous lui livrez l'action que vous possédez au prix K, et votre put ne vaut rien. La valeur de votre portefeuille est donc bien celle que nous avons calculée.

(b) $S_N^l < K$; l'acheteur du call a perdu son argent, vous vendez votre action au prix K, ce que vous autorise la possession d'un put ; la valeur de votre portefeuille est encore la même .

Vous avez donc réussi à obtenir un "free lunch".

2. Si au contraire $C_0 - P_0 < S_0^l - \frac{K}{(1+r)^N}$, la stratégie consistant à acheter le call, à vendre le put et emprunter l'action conduit au même résultat (un arbitrage) ; le lecteur pourra avec profit écrire complètement le formalisme mathématique correspondant.
3. Vous disposez maintenant des éléments suffisants pour imaginer une stratégie prévisible permettant d'établir l'égalité (4.15), qui, vous l'avez noté, est une égalité entre variables aléatoires.

La fabrication d'un portefeuille sans risque à l'aide d'actifs (très) risqués

La relation de parité peut s'écrire

$$S_n^l - C_n + P_n = \frac{K}{(1+r)^N}(1+r)^n \tag{4.22}$$

ce qui conduit au résultat un peu surprenant suivant :
Un portefeuille constitué par
- l'achat de l'actif l et d'un put sur cet actif de prix d'exercice K
- la vente d'un call de même actif sous-jacent au même prix d'exercice

aura le même comportement que l'actif sans risque obtenu en plaçant la somme $\tilde{K} = \frac{K}{(1+r)^N}$ à la caisse d'épargne.

4.5 Options simulables (ou réplicables) dans un marché viable

Dans tout ce paragraphe, nous supposons donné un marché financier viable.
On dira qu'une option européenne h est *simulable* (on dit aussi *réplicable*) s'il existe une sratégie autofinancée ϕ telle que

$$h = V_N(\phi) =< \phi_N, S_N >$$

Puisque ϕ est autofinancée, l'égalité précédente peut encore s'écrire

$$h = V_0(\phi) + \sum_{k=1}^{k=N} < \phi_k, \Delta S_k > \quad \text{ou} \quad \tilde{h} = V_0(\phi) + \sum_{k=1}^{N} < \phi_k, \Delta \tilde{S}_k > \tag{4.23}$$

Définitions

Nous appellerons $\mathcal{P}$ l'ensemble des probabilités admettant Ω pour support sous lesquelles $(\tilde{S}_n)$ est une martingale vectorielle; nous avons vu en 4.3.3 que, dans un marché viable, $\mathcal{P}$ n'est pas vide. La stratégie ϕ est appel'ee *stratégie (ou portefeuille) de couverture de l'option h*

Proposition

Soient h une option européenne, ϕ un portefeuille simulant h et $\mathbf{P^*} \in \mathcal{P}$; on a l'égalité

$$\tilde{V}_n(\phi) = \mathbf{E^*}[\tilde{h}|\mathcal{F}_n] \quad \forall n \geq 0 \tag{4.24}$$

Démonstration : Sous $\mathbf{P^*}$, $(\tilde{V}_n(\phi))$ est la trasformée de $(\tilde{S}_n)$ par (ϕ_n) ; c'est donc une martingale de variable terminale $\tilde{h}$.

Remarque :

Si le marché est viable, toute stratégie autofinancée simulant h est nécessairement admissible, comme le prouve le petit raisonnement suivant : soit ϕ autofinancée telle que $V_N(\phi) = h \geq 0$; si $\mathbf{P^*} \in \mathcal{P}$, nous venons de voir que $\tilde{V}_n(\phi) = \beta_N \mathbf{E^*}[h|\mathcal{F}_n] \geq 0$. Il en résulte que $V_n(\phi) \geq 0 \quad \forall n$.

Le prix d'une option réplicable.

En l'absence d'hypothèse supplémentaire, (voir le paragraphe suivant), $\mathbf{P^*}$ n'est pas unique, pas plus que la stratégie ϕ de couverture de h. Un examen de (4.24) conduit néammoins à une remarque importante : le premier membre de cette égalité ne dépend pas de $\mathbf{P^*} \in \mathcal{P}$ tandis que le second membre ne dépend pas de la stratégie de couverture ϕ. La valeur commune aux 2 membres de (4.24) ne dépend donc que de h et de n, ce qui justifie les notations suivantes :

$$\pi_n(h) = V_n(\phi) = \frac{1}{(1+r)^{N-n}} \mathbf{E^*}[h|\mathcal{F}_n] \ \ ; \ \ \pi_0(h) = V_0(\phi) = \frac{1}{(1+r)^N} \mathbf{E^*}[h] \tag{4.25}$$

La *variable aléatoire* $\mathcal{F}_n$-mesurable $\pi_n(h)$ s'appellera *prix de l'option* h *à la date* n.

La *constante réelle positive* $\pi_0(h)$, (que l'on notera plus simplement $\pi(h)$) s'appellera *prix de l'option* h. C'est à ce prix qu'un mathématicien employé par une banque conseillera de vendre l'option h à la date 0 ; en voici une première raison :

L'émetteur de l'option h disposant au temps 0 de la somme $V_0(\phi)$ est assuré pouvoir faire face à ses obligations, (c'est à dire débourser la somme h quand l'option sera exercée par son acheteur à la date d'exercice) ; il lui suffit pour cela d' investir cet argent sur le marché en suivant la stratégie ϕ. Il est donc naturel qu' il vende cette option au prix $\pi_0(h)$ pour apaiser ses angoisses.
La proposition suivante, nous en fournit une seconde, faisant intervenir la nature profonde d'un marché financier (absence d'arbitrage).

Proposition

Soit h une option européenne réplicable ; $\pi(h)$ est l'unique prix excluant l'arbitrage de cette option.
Démonstration :

1. Montrons d'abord que $\pi(h)$ est un p.e.a. de l'option h ; revenons aux notations de 4.4.2 et posons $H_n = V_n(\phi)$. Soit d'autre part $\mathbf{P^*} \in \mathcal{P}$; comme $(\tilde{S}_n)$ et $(\tilde{H}_n)$ sont des $\mathbf{P^*}$-martingales, il en est de même de $(\tilde{S}_n, \tilde{H}_n))$. Il en résulte que le marché financier étendu $\Big(\Omega, (\bar{S}_n) := (S_n, H_n), \mathbf{P}\Big)$ est viable.

2. Passons à l'unicité; désignons par (h_n) une suite adaptée de v.a.r. ≥ 0 vérifiant $h_N = h$ et posons $\pi = h_0$. Nous allons montrer que, si $\pi \neq \pi(h)$, on peut construire une statégie d'arbitrage dans le marché financier étendu $\left(\Omega, \bar{S}_n := (S_n, h_n), \mathbf{P}\right)$. (On considère maintenant que l'option, introduite dans le marché, constitue un actif à risque supplémentaire noté $d+1$ dont le cours à la date n est h_n).
   Supposons par exemple $\pi > \pi(h)$; posons alors $\bar{\psi}_n = (\phi_n^1, \phi_n^2, \ldots, \phi_n^d, -1)$ et désignons par $(\bar{\phi}_n)$ le portefeuille autofinancé coïncidant avec $\bar{\psi}$ sur les actifs à risque du marché étendu et vérifiant $V_0(\bar{\phi}) = 0$ (Cf. 4.2.2). On a

$$\begin{aligned} V_n(\bar{\phi}) &= \bar{\phi}_n^0 S_n^0 + \phi_n^1 S_n^1 + \ldots + \phi_n^d S_n^d - h_n && (4.26) \\ &= (\bar{\phi}_n^0 - \phi_n^0) S_n^0 + V_n(\phi) - h_n && (4.27) \end{aligned}$$

   (4.27) écrite en $n = 0$ conduit à l'égalité

$$\bar{\phi}_0^0 - \phi_0^0 = \pi - \pi(h) \qquad (4.28)$$

   Comme les stratégies $\bar{\phi}$ et ϕ sont autofinancées, (4.26) s'écrit encore

$$\begin{aligned} V_n(\bar{\phi}) &= \bar{\phi}_{n+1}^0 S_n^0 + \phi_{n+1}^1 S_n^1 + \ldots + \phi_{n+1}^d S_n^d - h_n && (4.29) \\ &= (\bar{\phi}_{n+1}^0 - \phi_{n+1}^0) S_n^0 + V_n(\phi) - h_n && (4.30) \end{aligned}$$

   Comparant (4.27) et (4.30), on obtient les relations

$$\bar{\phi}_{n+1}^0 - \bar{\phi}_n^0 = \phi_{n+1}^0 - \phi_n^0$$

   Compte tenu de (4.28), on a donc

$$\begin{aligned} \bar{\phi}_n^0 &= \pi - \pi(h) + \phi_n^0 \\ V_n(\bar{\phi}) &= (\pi - \pi(h))(1+r)^n + V_n(\phi) - h_n \\ V_N(\bar{\phi}) &= (\pi - \pi(h))(1+r)^N \ > 0 \end{aligned}$$

   La stratégie $\bar{\phi}$ constitue donc un arbitrage qui peut être mis en pratique de la façon suivante : le spéculateur attiré par la surévaluation de l'option h l'emprunte à la date 0, la vend immédiatement sur le marché au prix π, et prélève sur cette somme la quantité $\pi(h)$ qui lui permet d'acheter le portefeuille de couverture ϕ de l'option h ; la somme restante, $\pi - \pi(h)$, est placée à la caisse d'épargne. Après avoir laissé au lecteur le soin d' imaginer une stratégie d'arbitrage dans le cas où $\pi < \pi(h)$, on conclut qu'on a nécessirement $\pi = \pi(h)$.

4.5.1 Les prix d'une option excluant l'arbitrage dans le cas général

Plaçons[2] nous dans un marché viable et donnons nous une option européenne h non nécessairement simulable. La proposition suivante va en particulier montrer l'existence de p.e.a. pour cette option .

[2] on pourra omettre ce paragraphe en première lecture.

Proposition :

Soit $\Pi(h)$ le sous ensemble de $\mathbb{R}_+$ contitué par les prix excluant l'arbitrage de l'option h ; on a

$$\Pi(h) = \{ \ \mathbf{E^*}[\tilde{h}] \ ; \ \mathbf{P^*} \in \mathcal{P} \ \} \tag{4.31}$$

Démonstration

1. Soit $\pi \in \Pi(h)$; il existe alors une suite adaptée (h_n) vérifiant $h_N = h$ et $h_0 = \pi$ telle que le marché $\big(\Omega, (S_n, h_n), \mathbf{P}\big)$ soit viable. Il existe donc (Cf. 4.3.3) une probabilité $\bar{\mathbf{P}}$ chargeant tous les points de Ω sous laquelle $(\tilde{S}_n)$ et $(\tilde{h}_n)$ sont des martingales ; il en résulte en particulier que $\bar{\mathbf{P}} \in \mathcal{P}$. De plus, $\pi = h_0 = \bar{\mathbf{E}}[\tilde{h}]$, ce qui montre que π appartient au membre de droite de (4.31).
2. Inversement, soient $\mathbf{P^*} \in \mathcal{P}$ et $\pi = \mathbf{E^*}[\tilde{h}]$; posons $\tilde{h}_n = \mathbf{E^*}[\tilde{h}|\mathcal{F}_n] = \beta_n h_n$. Comme $(\tilde{S}_n)$ et $(\tilde{h}_n)$ sont des $\mathbf{P^*}$-martingales, le marché étendu $\big(\Omega, (S_n, h_n), \mathbf{P}\big)$ est viable ; d'autre part, on voit immédiatement que $h_0 = \pi$ et $h_N = h$; il en résulte que $\pi \in \Pi(h)$.

Complément

On peut aller plus loin dans la description de $\Pi(h)$ dans un marché viable :

- Ou bien h est réplicable et, comme nous venons de le voir, $\Pi(h)$ comporte un seul point.
- Dans le cas contraire, $\Pi(h)$ est un intervalle ouvert de $\mathbb{R}$. (Cf. Follmer et Schied : Stochastic Finance)

4.6 Marchés complets

4.6.1 Définition :

On dit qu'un marché financier est complet si toute option européenne est simulable.

Contrairement à l'axiome de viabilité qui paraît inhérent à la nature d'un marché, cette hypothèse apparaît comme une commodité mathématique ; on peut même se demander si le fait d'exiger que toute option puisse être simulée par une stratégie convenable est bien réaliste ; on notera tout de même que l'on ne demande pas pour l'instant d'exhiber cette stratégie : savoir qu'elle existe est suffisant pour obtenir des résultats. D'autre part, le modèle de Black and Scholes qui est le plus couramment utilisé fournit un exemple important de marché complet ; il en est de même du modèle discret que nous allons étudier au chapitre suivant.

4.6.2 Martingales et marchés complets

Théorème

Un marché viable est complet si et seulement s'il existe une probabilité *unique* sous laquelle les prix actualisés des actifs sont des martingales

Démonstration :

a) Supposons le marché viable et complet et soit h une v.a.r. ≥ 0 ; il existe une stratégie admissible ϕ telle que $h = V_N(\phi)$; comme ϕ est autofinancée , on peut écrire (cf. la remarque **4.2.1**)

$$\beta_N h = \tilde{V}_N(\phi) = V_0(\phi) + \sum_1^N < \phi_j, \Delta\tilde{S}_j > .$$

Si $\mathbf{P}_1$et $\mathbf{P}_2$ sont deux probabilités sous lesquelles $(\tilde{S}_n)$ est une martingale, on aura $\beta_N \mathbf{E}_1[h] = \beta_N \mathbf{E}_2[h]$, ce qui entraine évidemment $\mathbf{P}_1 = \mathbf{P}_2$

b) Inversement , supposons le marché viable et non complet ; il exite alors une v.a. $h \geq 0$ non simulable ; reprenant les notations du corollaire **4.2.2**, cela signifie que $\mathcal{V}$ est un sous espace vectoriel strict de $\mathcal{E}$ que nous munirons maitenant du produit scalaire $< X, Y >= \mathbf{E^*}[XY]$; d'après ce que l'on sait des espaces euclidiens, il existe dans $\mathcal{E}$ un vecteur X non nul (i.e. une v.a.r. non nulle) orthogonal à $\mathcal{V}$; cela entraine en particulier, puisque X est orthogonale aux constantes, $\mathbf{E^*}[X] = 0$. Définissons alors $\mathbf{P}_*$ par l'égalité

$$\mathbf{P}_* = \frac{k + X}{k}\mathbf{P^*} = (1 + \frac{X}{k})\mathbf{P^*}$$

où k est un réel > 0 assez grand pour que $k + X > 0$; il est alors facile de voir que $\mathbf{P}_*$ est une probabilité différente de $\mathbf{P^*}$; de plus, quelque soit la suite prévisible ψ à valeurs dans $\mathbb{R}^d$,

$$\mathbf{E}_*[\sum_1^N < \psi_n, \Delta\tilde{S}_n >] = \mathbf{E^*}[X \sum_1^N < \psi_n, \Delta\tilde{S}_n >] = 0.$$

Il résulte alors du lemme **3.3.4** que $(\tilde{S}_n)$ est une martingale sous $\mathbf{P}_*$.

4.7 Prix et couverture d'une option européenne dans un marché complet

Nous en arrivons à la question centrale : à quel prix doit on vendre une option européenne h ? La première estimation qui vient à l'esprit est la valeur moyenne de l'option à sa date d'exercice , à savoir $\mathbf{E}[h]$; c'est oublier que la vente de l'option procure à son émetteur , dès l'instant 0 , une somme d'argent qu'il peut faire fructifier à sa guise ; il peut, par exemple, placer cet argent à la caisse d'épargne et, dans cette hypothèse, se permettre de vendre moins cher, au prix

$$\beta_N \mathbf{E}[h] = \mathbf{E}[\tilde{h}] = \frac{1}{(1+r)^N}\mathbf{E}[h] \tag{4.32}$$

Mais, dans un marché complet, il peut faire mieux : il a (au moins théoriquement) la possibilité de simuler h s'il consent à investir dans des actifs à risque ; il dispose ainsi d'une stratégie ϕ (appelée *stratégie de couverture* de l'option) lui permettant de constituer un portefeuille dont la valeur à l'instant N est exactement h. Il est donc naturel de définir le prix V_0 de l'option h à l'instant 0 et son prix V_n à l'instant n par

$$V_0 = V_0(\phi) \quad ; \quad V_n = V_n(\phi).$$

On notera que V_0 est une constante , (c'est le prix conseillé à l'émetteur de l'option) , tandis que V_n est une variable aléatoire $\mathcal{F}_n$-mesurable , c'est à dire une fonction de l'évolution des cours observés jusqu'à la date n , (dans les cas usuels ce sera une fonction du dernier cours S_n) , ce qui parait raisonnable .

Dans la réalité, notre option h sera mise sur le marché tout comme une action, et deviendra ainsi un actif à risque supplémentaire ; elle aura un prix W_n à l'instant n (différent de V_n) déterminé par la loi de l'offre et de la demande . W_n est la somme d'argent qu'il vous faudra débourser s'il vous prend la fantaisie d'acheter cette option à la date n ; que peut on penser s'il existe un écart important entre $V_n$et W_n ? Trois choses différentes suivant les tempéraments :

- que l'on a perdu son temps en lisant ces notes qui exposent une théorie totalement étrangère aux réalités du marché .
- que la théorie probabiliste est valide mais que les statisticiens ont mal fait leur travail : la loi de probabilité $\mathbf{P}$ gouvernant le cours des actifs est incorrecte .
- que le marché a tort parce qu'il ne saurait avoir raison contre un raisonnement mathématique ; si vous observez à la date n un prix W_n très inférieur à V_n , vous investissez alors votre fortune et celle de toute votre famille dans l'option h .

Bien entendu, si les opérateurs utilisent ces modèles , c'est parce que les prix qu'ils permettent de calculer ne sont pas trop éloignés des prix réels constatés sur les marchés financiers ; mais il est de mauvais esprits pour prétendre que n'importe quelle formule magique fournie par une cartomancienne ou un astrologue aboutirait au même résultat , pourvu que tous les opérateurs soient convaincus de sa pertinence : si la totalité d'entre eux sont persuadés qu'il existe un "juste prix" pour une option, c'est évidemment à ce prix là que s'effectueront les transactions. Mais revenons aux mathématiques ; on a le résultat suivant :

Théorème

$$V_0 = \frac{1}{(1+r)^N}\mathbf{E^*}[h] \quad ; \quad V_n = \frac{1}{(1+r)^{N-n}}\mathbf{E^*}[h|\mathcal{F}_n] \tag{4.33}$$

On pourra constater que la faculté donnée au vendeur de l'option de se couvrir par des actifs à risque se traduit en langage mathématique par la substitution de $\mathbf{P^*}$ à $\mathbf{P}$; (comparer (4.13) et (4.14)) . Le prix d'une option se calculant à l'aide de $\mathbf{P^*}$, on peut oublier la probabilité $\mathbf{P}$, et même dans certains cas , ne pas la connaitre avec précision .Nous allons en voir un exemple au chapitre suivant.

Démonstration : Puisque $(\tilde{V}_n(\phi))$ est une martingale sous $\mathbf{P^*}$, on a

$$V_0(\phi) = \mathbf{E^*}[\tilde{V}_N(\phi)] = \beta_N\mathbf{E^*}[h] \quad , \quad \beta_n V_n(\phi) = \beta_N\mathbf{E^*}[h|\mathcal{F}_n]$$

Les formules (3.12) en résultent immédiatement .

Remarque

Les formules précédentes permettent de donner une démonstration plus rapide (mais qui n'est valable que dans un marché complet) de la relation de parité call-put (36) ; on a en effet

$$C_n - P_n = \frac{1}{(1+r)^{N-n}}\mathbf{E^*}[S_N - K|\mathcal{F}_n] = (1+r)^n\mathbf{E^*}[\tilde{S}_N|\mathcal{F}_n] - \frac{K}{(1+r)^{N-n}} = S_n - \frac{K}{(1+r)^{N-n}}$$

Chapter 5

Le modèle de Cox, Ross et Rubinstein

5.1 Description du modèle .

Soient a et b deux réels vérifiant $-1 < a < b$; nous allons construire un marché financier fini servant de modèle mathématique à la situation boursière simpliste suivante : il n'existe qu'un actif à risque $((S_n) \quad ; 0 \leq n \leq N)$ et , comme d'habitude , un actif non risqué $S_n^0 = (1+r)^n$. Nos hypothèses sur le comportement de l'action S seront celles ci : entre les instants n et $n+1$, le cours S_n de l'actif risqué est multiplié de façon aléatoire soit par $1+a$, soit par $1+b$; autrement dit ,

$$S_{n+1} = T_{n+1} S_n$$

où (T_n) est une suite de v.a à valeurs dans l'ensemble à deux points $\{(1+a), (1+b)\}$. Cette hypothèse est très restrictive : imposer qu'un cours de bourse soit chaque jour , ou mutiplié par 2 , ou divisé par 3 , apparait comme hautement irréaliste ; elle conduit néammois à un modèle qui nous intéressera pour 3 raisons :

- il est mathématiquement très simple et permet des calculs explicites .
- comme nous le verrons ci-dessous , c'est un marché complet .
- il est l'analogue discret du modèle à temps continu de Black and Scholes .

En revanche , on n'impose aucune contrainte ni sur les lois des v.a. T_n (qui peuvent dépendre de n) , ni sur leurs interactions (en particulier , on ne suppose pas leur indépendance) .

5.1.1 L'espace des épreuves

On pose $\Omega = \{1+a, 1+b\}^N$; une épreuve ω sera donc une suite $\{t_1, t_2, ..., t_N\}$ de N réels valant $1+a$ ou $1+b$. Les fonctions coordonnées T_i , définies par $T_i(\omega) = t_i$, seront les N variables aléatoires centrales de ce modèle . On notera $(\mathcal{F}_n)$ la filtration naturelle engendrée par la suite (T_n) , complétée à l'origine par $\mathcal{F}_0 = \{\Omega, \phi\}$.
Soit $s_0 \in \mathbb{R}$; on définit une suite réelle $((S_n) \ ; \ 0 \leq n \leq N)$ en posant

$$S_0 = s_0 \ , \ S_1 = s_0 T_1, ..., S_n = S_{n-1} T_n = s_0 T_1 T_2 ..... T_n$$

Nous laisserons au lecteur le soin de montrer que $(\mathcal{F}_n)$ est aussi la filtration naturelle engendrée par la suite (S_n) . Rappelons que nous disposons d'une deuxième suite S^0 , déterministe , définie par

$$S_n^0 = (1+r)^n = \frac{1}{\beta^n}$$

On munit Ω d'une probabilité $\mathbf{P}$ soumise à une seule restriction : charger tous les points de Ω .

5.2 Viabilité et complétude de ce marché

Soit $\mathbf{P}$ une probabilité quelconque sur Ω ; examinons à quelles condiditions la suite $(\tilde{S}_n)$ est une martingale sous $\mathbf{P}$. On a

$$\mathbf{E}[\tilde{S}_{n+1}|\mathcal{F}_n] = \beta_{n+1}\mathbf{E}[S_{n+1}|\mathcal{F}_n] = \beta_{n+1}S_n\mathbf{E}[T_{n+1}|\mathcal{F}_n] = \beta\tilde{S}_n\mathbf{E}[T_{n+1}|\mathcal{F}_n] = \frac{1}{1+r}\tilde{S}_n\mathbf{E}[T_{n+1}|\mathcal{F}_n],$$

de sorte qu'une condition nécessaire et suffisante pour qu'il en soit ainsi est : $\mathbf{E}[T_{n+1}|\mathcal{F}_n] = 1+r$; Il en résulte (prendre l'espérance des 2 membres) , que $\mathbf{E}[T_n] = 1+r$. Puisque

$$1+a < (1+a)\mathbf{P}[T_n = 1+a] + (1+b)\mathbf{P}[T_n = 1+b] = 1+r < 1+b,$$

on a nécessairement $a < r < b$. Mais, si le marché est viable, on peut appliquer ce qui vient d'être dit à la probabilité $\mathbf{P^*}$; ainsi :

Une condition nécessaire pour que le marché soit viable est : $r \in]a,b[$; nous supposerons cette condition vérifiée dans toute la suite.

Il n'était pas nécessaire de recourir aux résultats du paragraphe **3.3.4** pour obtenir ce résultat : si l'on suppose, par exemple, $r > b$, il est facile de gagner de l'argent à coup sûr, sans apport initial et sans risque : à l'instant 0, on emprunte une action que l'on vend immédiatement pour placer l'argent obtenu à la caisse d'épargne (c'est ce qu'on appelle une "vente à découvert" : on vend une action qu'on ne possède pas), puis, à l'instant N, on rachète l'action sur le marché pour rembourser sa dette. Nous laisserons le lecteur écrire la statégie ϕ correspondante et vérifier que l'on est en présence d'un arbitrage ; que faire si $r < a$?
Pour aller plus loin, nous aurons besoin du lemme suivant :

5.2.1 Lemme :

Soit T une variable aléatoire réelle à valeurs dans l'esemble à deux points $\{\alpha, \beta\}$ $(\alpha < \beta)$ et soit $\mathcal{G}$ une sous algèbre de $\mathcal{F}$; on suppose qu'il existe une constante m telle que $\mathbf{E}[T|\mathcal{G}] = m$; T est alors indépendante de $\mathcal{G}$ et sa loi est donnée par les égalités

$$\mathbf{P}[T = \alpha] = \frac{\beta - m}{\beta - \alpha} \qquad ; \qquad \mathbf{P}[T = \beta] = \frac{m - \alpha}{\beta - \alpha}$$

Démonstration On montre d'abord comme précédemment que $\alpha < m < \beta$.
Posons $A = \mathbf{P}[T = \alpha|\mathcal{G}]$, $B = \mathbf{P}[T = \beta|\mathcal{G}]$; en conditionnant par rapport à $\mathcal{G}$ les égalités

$$1_{[T=\alpha]} + 1_{[T=\beta]} = 1 \ , \ T = \alpha 1_{[T=\alpha]} + \beta 1_{[T=\beta]}$$

on obtient

$$A + B = 1 \quad , \quad \mathbf{E}[T|\mathcal{G}] = \alpha A + \beta B.$$

Il en résulte que les variables aléatoires A et B sont solutions du système de 2 équations à 2 inconnues :

$$A + B = 1 \qquad , \qquad \alpha A + \beta B = m \quad ,$$

dont l'unique solution est

$$A = \frac{\beta - m}{\beta - \alpha} \qquad , \qquad B = \frac{m - \alpha}{\beta - \alpha} \ .$$

Il est alors important de noter que les variables aléatoires A et B sont des constantes ; cela signifie en effet que la loi conditionnelle de T quand $\mathcal{G}$ ne dépend pas de ω. La v.a. T est donc indépendante de l'algèbre $\mathcal{G}$; l'assertion concernant sa loi est à peu près évidente .

Notations

Dans la suite de ce chapitre , on adoptera les notations suivantes :

- On posera $p = \frac{b-r}{b-a}$, $q = \frac{r-a}{b-a}$
- μ désignera la loi de probabilité portée par $\{1+a, 1+b\}$ définie par $\mu(1+a) = p$, $\mu(1+b) = q$
- On notera enfin $\mathbf{P^*}$ la probabilité $\mu^N = \mu \times \mu \times ... \times \mu$; rappelons (cf.**1.3**) que $\mathbf{P^*}$ est l'unique probabilité sur Ω sous laquelle les fonctions coordonnées $T_1, T_2, ..., T_N$ constituent une famille de v.a. indépendantes équidistribuées de loi μ.

Proposition :

$\mathbf{P^*}$ est l'unique probabilité sous laquelle la suite $((\tilde{S}_n) \ ; \ 0 \leq n \leq N)$ est une martingale. Le marché de Cox, Ross et Rubinstein est donc viable et complet.

Démonstration :

1. Vérifions tout d'abord que, sous $\mathbf{P^*}$, $(\tilde{S}_n)$ est une martingale ; comme T_{n+1} est indépendante de $\mathcal{F}_n$, on a
$$\mathbf{E^*}[T_{n+1}|\mathcal{F}_n] = \mathbf{E^*}[T_{n+1}] = \frac{1}{b-a}[(1+a)(b-r) + (1+b)(r-a)] = 1 + r,$$
ce qui, nous l'avons vu au début de ce paragraphe, entraine le résultat.
2. Soit $\mathbf{\hat{P}}$ une autre probabilité sur Ω induisant la même propriété sur la suite $(\tilde{S}_n)$; on aura alors $\mathbf{\hat{E}}[T_{n+1}|\mathcal{F}_n] = 1 + r$; il résulte alors du lemme **4.3.1** que T_{n+1} est indépendante de $\mathcal{F}_n$ (c'est à dire de $(T_1, T_2, ..., T_n)$) et a pour loi μ. On en déduit aisément que les v.a. coordonnées T_n sont, sous $\mathbf{\hat{P}}$, indépendantes et de loi μ ; on a donc $\mathbf{\hat{P}}$=$\mathbf{P^*}$, c.q.f.d.

5.2.2 Prix d'une option européenne dans le marché de Cox, Ross et Rubinstein

Le théorème qui suit est une conséquence immédiate de l'étude faite au chapitre précédent et des égalités (38)

Théorème Soit h une option européenne et soit $\mathbf{P^*}$ la probabilité sur Ω définie en **4.3.2** ; le prix V_n de cette option à la date n est donnée par la formule $V_n = \frac{1}{(1+r)^{N-n}}\mathbf{E^*}[h|\mathcal{F}_n]$; en particulier , son prix à la date 0 est $\frac{1}{(1+r)^N}\mathbf{E^*}[h]$.

Beaucoup d'entre vous trouveront cette formule trop théorique donc peu digne d'attention ; les 2 conséquences suivantes peuvent peut être les faire changer d'avis :

- Soient $\mathbf{P}_1$ et $\mathbf{P}_2$ 2 probabilités sur Ω vérifiant respectivement $\mathbf{P}_1[T_n = 1+b] = 0,99 \ \ \forall n$ (marché très haussier) et $\mathbf{P}_2[T_n = 1+a] = 0,99 \ \ \forall n$ (krach permanent) ; comme $\mathbf{P^*}$ est la même dans les 2 cas , il en sera de même du prix d'une option : contrairement à une intuition trompeuse , le prix d'une option ne dépend pas de la tendance du marché .On observe le même phénomène dans le modèle de Black and Scholes . On peut avancer l'explication suivante : dans l'hypothèse haussière , par exemple , le vendeur d'une option , disposant d'argent frais à l'instant 0 a la possibilité le faire fructifier en achetant des actifs risqués ; il peut donc se permettre de vendre son option à un prix plus bas que celui auquel on pourrait s'attendre .
- La probabilité $\mathbf{P^*}$ est plus simple que la probabilité $\mathbf{P}$ gouvernant le marché ; le fait que le prix d'une option se calcule sous $\mathbf{P^*}$ est une bénédiction : faire des calculs avec des v.a. indépendantes équidistribuées met tout probabiliste dans l'état euphorique que nous allons éprouver à la lecture du paragraphe suivant.

5.3 Calculs explicites

Nous allons maintenant considérer un call et un put sur l'actif risqué de prix d'exercice K dont les prix à la date n seront notés C_n et P_n ;

Lemme

Soit $T_1, T_2, ..., T_l$ une suite de l variables aléatoires indépendantes équidistribuées de loi commune μ , et soit ψ une fonction numérique symétrique sur $\mathbb{R}^l$;on a , si l'on note $s^{(j)}$ l'élément $(s, s, ..., s)$ de $\mathbb{R}^j$

$$\mathbf{E}[\psi(T_1, T_2, ..., T_l)] = \sum_{j=0}^{j=l} \binom{l}{j} p^j (1-p)^{l-j} \psi((1+a)^{(j)}, (1+b)^{(l-j)}) \qquad (39)$$

Démonstration : Posons $J(\omega) = \sharp\{i \ |T_i(\omega) = 1+a\}$; la démonstration est très simple si l'on veut bien réfléchir aux 2 points suivants :

- la loi de J est binomiale ; on peut en effet écrire $J = \sum_{i=1}^{l} U_i$ quand on a posé $U_i = 1_{\{T_i = 1+a\}}$. Il reste alors à remarquer que $(U_i \ ; \ 0 \leq i \leq l)$ constitue une suite indépendante de v.a. de Bernoulli telles que $\mathbf{P}[U_i = 1] = p$, $\mathbf{P}[U_i = 0] = 1-p$
- compte tenu de la symétrie supposée sur ψ , la v.a $\psi(T_1, T_2, ...T_l)$ est une fonction de J ; plus précisément
$$\psi(T_1, T_2, ..., T_l) = \psi((1+a)^{(J)}, (1+b)^{(l-J)})$$

On a donc

$$\begin{aligned}
\mathbf{E}[\psi(T_1, T_2, ..., T_l)] &= \sum_{j=0}^{j=l} \psi((1+a)^{(j)}, (1+b)^{(l-j)}) \mathbf{P}[J=j] \\
&= \sum_{j=0}^{j=l} \begin{pmatrix} l \\ j \end{pmatrix} p^j (1-p)^{l-j} \psi((1+a)^{(j)}, (1+b)^{(l-j)})
\end{aligned}$$

5.3.1 Théorème :

Le prix d'un call européen est donné par la formule

$$C_n = \frac{1}{(1+r)^{N-n}} \sum_{j=0}^{j=N-n} \begin{pmatrix} N-n \\ j \end{pmatrix} p^j (1-p)^{N-n-j} [S_n (1+a)^j (1+b)^{N-n-j} - K]^+ \tag{5.1}$$

Démonstration :

$$(1+r)^{N-n} C_n = \mathbf{E^*}[(S_N - K)^+ | \mathcal{F}_n] = \mathbf{E^*}[(S_n T_{n+1} .... T_N - K)^+ | \mathcal{F}_n]$$

Posons alors

$$c(n, s) = \frac{1}{(1+r)^{N-n}} \mathbf{E^*}[(s T_{n+1} T_{n+2} .... T_N - K)^+]$$

de sorte que ,(cf (13)) , $C_n = c(n, S_n)$. Le lemme précédent appliqué à la fonction symétrique

$$\psi_s(t_{n+1}, t_{n+2}, ..., t_N) = (s t_{n+1} t_{n+2} .... t_N - K)^+$$

conduit à l'égalité

$$c(n, s) = \frac{1}{(1+r)^{N-n}} \sum_{j=0}^{j=N-n} \begin{pmatrix} N-n \\ j \end{pmatrix} p^j (1-p)^{N-n-j} [s(1+a)^j (1+b)^{N-n-j} - K]^+.$$

On en déduit immédiatement la formule (4.1) . Il est inutile de recommencer un calcul analogue pour obtenir le prix du put : il est plus simple de se servir de la formule de parité .

5.3.2 Couverture du call

Rappelons que l'émetteur de notre call n'a pu le vendre à un prix raisonnable que parceque nous lui avons dit qu'il existait , (puisque le marché est complet), une stratégie de couverture de cette option . Mais si le mathématicien se satisfait aisément d'un théorème d'existence , la banque qui l'emploie exigera qu'il lui fournisse explicitement le moyen de se couvrir ; c'est ce que nous allons faire maintenant . Soit $(\Phi_n) = (\phi_n^0, \phi_n)$ une statégie prévisible simulant l'option d'achat ; on a

$$C_n = c(n, S_n) = V_n(\Phi) = \phi_n^0 (1+r)^n + \phi_n S_n,$$

ce que l'on écrira $c(n, T_n S_{n-1}) = \phi_n^0 (1+r)^n + \phi_n T_n S_{n-1}$; rappelons nous que T_n ne prend que les 2 valeurs $1+a$ et $1+b$; on a donc

$$\left\{ \begin{array}{l} c(n,(1+a)S_{n-1})1_{[T_n=1+a]} = [\phi_n^0(1+r)^n + (1+a)\phi_n S_{n-1}]1_{[T_n=1+a]} \\ c(n,(1+b)S_{n-1})1_{[T_n=1+b]} = [\phi_n^0(1+r)^n + (1+b)\phi_n S_{n-1}]1_{[T_n=1+b]} \end{array} \right.$$

Plaçons nous sous **P*** et prenons les espérances conditionnelles des 2 égalités précédentes par rapport à l'algèbre $\mathcal{F}_{n-1}$; compte tenu de ce que

- ϕ_n est $\mathcal{F}_{n-1}$-mesurable
- T_n est indépendante de $\mathcal{F}_{n-1}$

on obtient

$$\left\{ \begin{array}{r} c(n,(1+a)S_{n-1})p = [\phi_n^0(1+r)^n + (1+a)\phi_n S_{n-1}]p \\ c(n,(1+b)S_{n-1})(1-p) = [\phi_n^0(1+r)^n + (1+b)\phi_n S_{n-1}](1-p) \end{array} \right.$$

Il reste à résoudre ce système linéaire à 2 inconnues pour obtenir

$$\phi_n^0 = \frac{(1+b)c(n,(1+a)S_{n-1}) - (1+a)c(n,(1+b)S_{n-1})}{(b-a)(1+r)^n} \tag{5.2}$$

$$\phi_n = \frac{c(n,(1+b)S_{n-1}) - c(n,(1+a)S_{n-1})}{(b-a)S_{n-1}} \tag{5.3}$$

On notera que l'on obtient bien 2 suites prévisibles ; il y a plus : à l'instant n , il suffit de connaitre les cours de l'instant immédiatement précédent pour réorganiser son portefeuille . Nous laissons en exercice la couverture d'une option de vente .

Chapter 6

Les options américaines

Dans tout ce chapitre , nous nous placerons dans un marché complet .
Rappelons nous que le détenteur d'un call américain a la possibilité d'exercer son option à n'importe quel moment compris entre 0 (sa date d'émission) et N (sa date d'exercice) ; il est donc naturel de représenter ce call par la suite de v.a. $(S_n - K)^+$. Plus généralement ,

Définition

Nous appellerons option américaine toute suite adaptée de variables aléatoires positives .

6.1 Valeur d'une option américaine

Nous ferons appel aux mêmes idées que pour les options européennes : la valeur U_n d'une option (Z_k) à la date n sera la somme d'argent dont le vendeur doit disposer à cette date pour être en mesure d'honorer son contrat à toute date postérieure à n s'il recourt à une stratégie optimale ; raisonnons par récurrence descendante et supposons connue à la date n la somme d'argent U_n nécessaire à la couverture de l'option entre n et N ; U_{n-1} doit alors satisfaire les 2 conditions suivantes :

- permettre de couvrir un éventuel exercice de l'option à la date $n-1$, ce qui entraine $U_{n-1} \geq Z_{n-1}$.
- être assuré de disposer le lendemain de la somme U_n , ce qui peut être obtenu en simulant la v.a. U_n à la date n ; on doit donc avoir $U_{n-1} \geq V(\phi)_{n-1}$, où ϕ est un portefeuille vérifiant $V_n(\phi) = U_n$. Plaçons nous sous $\mathbf{P*}$; comme $(\tilde{V}_n(\phi))$ est une martingale , $\beta_{n-1} V_{n-1}(\phi) = \beta_n \mathbf{E*}[V_n(\phi)|\mathcal{F}_{n-1}]$, d'où l'on tire $V_{n-1}(\phi) = \frac{1}{1+r}\mathbf{E*}[U_n|\mathcal{F}_{n-1}]$

Cela nous conduit à la définition suivante :

6.1.1 Proposition et définition

1. Soit (Z_k) une option américaine ; définissons par récurrence descendante la suite (U_n) par les égalités
$$U_N = Z_N \qquad , \qquad U_{n-1} = Z_{n-1} \vee \frac{1}{1+r}\mathbf{E*}[U_n|\mathcal{F}_{n-1}]$$
On dira que $U_n$est la valeur de l'option (Z_k) à la date n
2. $(\tilde{U}_n)$ est , sous $\mathbf{P*}$, l'enveloppe de Snell de la suite $(\tilde{Z}_n)$

On vérifie en effet facilement que $\tilde{U}_{n-1} = \tilde{Z}_{n-1} \vee \mathbf{E^*}[\tilde{U}_n|\mathcal{F}_{n-1}]$.

Exercice

On pose

$$V_n = \bigvee_{0 \le p \le N-n} \frac{1}{(1+r)^p} \mathbf{E^*}[Z_{n+p}|\mathcal{F}_n]$$

1. Montrer que $V_n \le U_n \quad \forall n$; (on pourra comparer V_n avec la v.a. X_n de lexercice **2.5.1**)
2. Réfléchir aux raisons pour lesquelles , malgré les apparences , la somme d'argent V_0 est en pratique insuffisante pour couvrir l'option américaine à tout instant .

6.2 Comparaison des options américaines et européennes

Une option américaine offre à son acheteur plus de possibilités que son homologue européenne , il parait donc vraisemblable que son prix sera plus élevé ; c'est ce que confirme la proposition ci-dessous dans laquelle nous adopterons les notations suivantes :

- $(Z_n \quad 0 \le n \le N)$ désignera une option américaine
- C_n sera le prix de cette option à la date n
- c_n désignera le prix à la date n de l'option européenne Z_N

6.2.1 Proposition

1. $C_n \ge c_n \qquad \forall n$
2. Si $(c_n) \ge (Z_n)$, alors $(C_n) = (c_n)$

   Démonstration :

   1. Comme (C_n) est une surmartingale et (c_n) une martingale , on a $C_n \ge \mathbf{E^*}[C_N|\mathcal{F}_n] = \mathbf{E^*}[c_N|\mathcal{F}_n] = c_n$
   2. (c_n) est une (sur)martingale majorant (Z_n) ; le caractère minimal d'une enveloppe de Snell entaine alors $C_n \le c_n$

6.2.2 Call américains et call européens

La proposition précédente a pour conséquence le résultat surprenant suivant :

Proposition

Soient (S_n) un actif risqué et K un réel ≥ 0 ; les prix du call américain $((S_n - K)^+ \quad 0 \le n \le N)$ et du call européen $(S_N - K)^+$ sont égaux .

Démonstration : Posant $Z_n = (S_n - K)^+$, et appelant c_n le prix à la date n du call européen $(S_N - K)^+$, il suffit de montrer que $c_n \ge Z_n \quad \forall n$. Or ,
$\tilde{c}_n = \mathbf{E^*}[(\tilde{S}_N - \beta_N K)^+|\mathcal{F}_n] \ge \mathbf{E^*}[\tilde{S}_N - \beta_N K|\mathcal{F}_n] = \tilde{S}_n - \beta_N K$.
Mais comme $\tilde{c}_n \ge 0$, on a aussi $\tilde{c}_n \ge (\tilde{S}_n - \beta_N K)^+ \ge (\tilde{S}_n - K)^+$, d'où le résultat .

Le résutat analogue n'est pas vrai pour les options de vente .

6.3 Prix et couverture d'un put américain dans le modèle de C.R.R.

6.3.1 Proposition

Soit $(n,x) \mapsto u(n,x) \quad (n \in \{0,1,2,...,N\} \quad ; \quad x \in \mathbb{R}_+)$ la suite de fonctions definie par récurrence descendante par les égalités

$$u(N,x) = (K-x)^+ \;\; ; \;\; u(n-1,x) = (K-x)^+ \vee \frac{pu(n,(1+a)x) + (1-p)u(n,(1+b)x)}{1+r} \quad .$$

Le prix U_n du put américain de prix d'exercice K sur l'actif risqué S est alors $U_n = u(n, S_n)$ à la date n

Démonstration : On a , d'après **5.2.2** , $U_{n-1} = Z_{n-1} \vee \frac{1}{1+r}\mathbf{E^*}[U_n|\mathcal{F}_{n-1}]$; raisonnons par récurrence descendante , et supposons que U_n puisse s'écrire $u(n, S_n)$; l'égalité précédente devient alors

$$U_{n-1} = (K - S_{n-1})^+ \vee \frac{1}{1+r}\mathbf{E^*}[u(n, S_{n-1}T_n)|\mathcal{F}_{n-1}]$$

Il ne reste plus qu'à appliquer la formule (13) pour obtenir le résultat .

6.3.2 Exercice

On posera $u = u(o,.)$

1. Montrer par récurrence que ,quelque soit n , $u(n,.)$ est borne supérieure d'une suite de fonctions affines ; en d'éduire que u est affine par morceaux , continue , décroissante et convexe . Montrer en outre que u est > 0 sur le segment $[0, \frac{K}{(1+a)^N}[$ et nulle sur son complémentaire .
2. On pose $A = \{x|u(x) = (K-x)^+\}$; montrer qu'il existe un réel $x^* \geq 0$ tel que $A = [0, x^*]$
3. Montrer que pour le détenteur d'un put américain sur l'actif S dont le cours à la date 0 est s_0 , il exite un seuil $\sigma \geq 0$ tel que
   – si $s_0 \leq \sigma$, il est indifférent d'exercer ou de vendre son option à l'instant 0 ,
   – si $s_0 > \sigma$ il est plus rémunérateur de vendre l'option plustôt que de l'exercer .

6.3.3 Stratégie de couverture

Puisque le marché est complet , il existe une stratégie ϕ prévisible et autofinancée vérifiant

$$U_n = V_n(\phi) = \phi_n^0 S_n^0 + \phi_n S_n.$$

Un raisonnement en tous points analogue à **4.3.2.** conduit alors à la formule donnant la quantité d'actif risqué à détenir à l'instant n pour couvrir notre option de vente :

$$\phi_n = \frac{u(n\;,(1+b)S_{n-1}) - u(n\;,(1+a)S_{n-1})}{(b-a)S_{n-1}}$$

Chapter 7

La formule de Black and Scholes

7.1 Taux d'intérêt instantanés

Soit $s : \mathbb{R}^+ \mapsto \mathbb{R}_+^*$ une fonction strictement positive et dérivable supposée représenter l'évolution des cours d'un actif sans risque (déterministe) au cours du temps (continu) . Si $t < u$, la quantité

$$r_{tu} = \frac{1}{s(t)} \frac{s(u) - s(t)}{u - t}$$

est le taux d'intérêt entre les instants t et u . Le taux d'intérêt instantané $r(t)$ à l'instant t sera , par définition ,

$$r_t = \lim_{u \downarrow t} r_{tu} = \frac{s'(t)}{s(t)} \tag{7.1}$$

La connaissance de la fonction $t \mapsto r_t$ et de la valeur à l'origine s_0 de notre actif sans risque permet de reconstituer la fonction s ; celle-ci est en effet l'unique solution de l'équation différentielle

$$\left\{ \begin{array}{l} ds(t) = s(t)r(t)dt \\ s(0) = s_0 \end{array} \right.$$

On a donc

$$s(t) = s_0 \exp(\int_0^t r(v)dv) \;\; ; \;\; s(u) = s(t) \exp(\int_t^u r(v)dv) \tag{7.2}$$

Dans le cas où le taux instantané est une constante r , les égalités précédentes deviennent

$$s(t) = s_0 e^{rt} \quad , \quad s(u) = s(t) e^{r(u-t)} \tag{7.3}$$

7.2 Une suite de processus de Cox , Ross et Rubinstein

Dans tout ce paragrahe , nous supposerons données 3 constantes strictement positives t, R, σ . Pour chaque entier $N > 0$ nous allons construire un marché financier de C.R.R. de la façon suivante :

- Les valeurs des 2 actifs seront observés aux $N + 1$ instants $0, t/N, 2t/N, \cdots , Nt/N = t$
- Le taux d'intérêt r_N entre les 2 dates nt/N et $(n + 1)t/N$ sera égal à $\frac{rT}{N}$.

– On définit a_N et b_N par les égalités

$$1+a_N=(1+r_N)e^{-\sigma/\sqrt{N}} \quad , \quad 1+b_N=(1+r_N)e^{\sigma/\sqrt{N}}$$

– On notera $(T_1^N,T_2^N,\cdots,T_N^N)$ une suite de N v.a. indépendantes équidistribuées à valeurs dans $\{1+a_N,1+b_N\}$; on supposera en outre $\mathbf{E}^*{}_N[T_1^N]=1+r_N$ ce qui ,nous l'avons vu , détermine la loi des v.a. T_n^N .
– Le cours de l'actif risqué ,noté S_n^N est donné par la formule $S_n^N=T_1^NT_2^N\cdots T_n^N$, (on suppose $S_0^N=1$)

Le pas des subdivisions que nous avons construites sur $[0\ t]$ tend vers 0 quand $N\to\infty$. Il parait naturel de penser que notre suite de processus disccrets va converger (en un sens que nous ne préciserons pas ici) vers un marché financier à temps continu d'horizon t . Puisque

$$\lim_{N\to\infty}(1+r_N)^N=\lim_{N\to\infty}(1+\frac{Rt}{N})^N=e^{Rt} \ ,$$

R apparait comme le taux d'intérêt instantané du processus limite .
Pour avoir une interprétation analogue de σ , nous aurons besoin du lemme suivant

7.2.1 Lemme

Pour tout $N\geq 1$, on se donne une suite $(X_n^N\ ;\ 1\leq n\leq N)$ de N variables aléatoires indépendantes et équidistribuées à valeurs dans l'ensemble à 2 points $\{-\frac{\sigma}{\sqrt{N}}\ ,+\frac{\sigma}{\sqrt{N}}\}$; on suppose en outre que l'espérance μ_N de X_1^N est telle que

$$\lim_{N\to\infty}N\mu_N=\mu \tag{7.4}$$

La suite (Y_N) définie par $Y_N=X_1^N+X_2^N+\cdots+X_N^N$ converge alors en loi vers une gaussienne de moyenne μ et de variance σ^2 .

Démonstration : Si ϕ_N est la fonction caractéristique de Y_N , on a

$$\phi_N(\lambda)=\mathbf{E}[e^{i\lambda Y_N}]=(\mathbf{E}[e^{i\lambda X_1^N}])^N$$

Utilisant alors le théorème de Paul Lévy relatif à la convergence en loi , on voit qu'il nous faut montrer que

$$\lim_{N\to\infty}\phi_N(\lambda)=\exp(i\lambda\mu-\lambda^2\sigma^2/2) \tag{7.5}$$

La formule de Taylor à l'ordre 2 avec reste de Lagrange permet d'écrire

$$e^{i\lambda x}=1+i\lambda x-(\lambda x)^2/2+R(x)$$

avec $|R(x)|=\frac{1}{6}|\lambda|^3|x|^3$; on a donc

$$e^{i\lambda X_1^N}=1+i\lambda X_1^N-\frac{1}{2}\lambda^2(X_1^N)^2+R(X_1^N) \tag{7.6}$$

avec $|R(X_1^N)|\leq\frac{|\lambda\sigma|^3}{6N\sqrt{N}}$. Prenant alors l'espérance des 2 membres de (6.6) , on obtient

$$\mathbf{E}[e^{i\lambda X_1^N}]=1+i\lambda\mu_N-\frac{\lambda^2\sigma^2}{2N}+\mathbf{E}[R(X_1^N)] \tag{7.7}$$

Mais comme

– $|\mathbf{E}[R(X_1^N)]| \leq \mathbf{E}[|R(X_1^N)|] \leq \frac{|\lambda|^3\sigma^3}{6N\sqrt{N}} = o(1/N)$
– $\mu_N = \frac{\mu}{N} + o(1/N)$,

on peut écrire

$$\mathbf{E}[e^{i\lambda X_1^N}] = 1 + i\lambda\frac{\mu}{N} - \frac{\lambda^2\sigma^2}{2N} + o(1/N) \qquad (7.8)$$

Revenons alors à l'étude de ϕ_N : on a

$$\begin{aligned} \phi_N(\lambda) &= \exp[N\log(1 + i\lambda\frac{\mu}{N} - \frac{\lambda^2\sigma^2}{2N} + o(1/N))] \\ &= \exp[i\lambda\mu - \frac{\lambda^2\sigma^2}{2} + No(1/N) + \epsilon(1/N)] \end{aligned}$$

(ϵ désignant une fonction tendant vers 0 à l'origine , et le symbole log la détermination principale du logarithme complexe) . Il ne reste plus quà faire tendre N vers l'infini pour obtenir (6.5)

7.2.2 Proposition

Posons $Y_N = \log \tilde{S}_N^N$; la suite (Y_N) converge en loi vers une gaussienne de moyenne $-\sigma^2/2$ et de variance σ^2 qui , rappelons le , a pour densité de probabilité la fonction ν définie sur $\mathbb{R}$ par

$$\nu(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(y+\frac{\sigma^2}{2})^2}{2\sigma^2}) \qquad (7.9)$$

Démonstration : On a $\log \tilde{S}_N^N = X_1^N + X_2^N + \cdots X_N^N$ quand on a posé $X_n^N = \log \frac{T_n^N}{1+r_N}$
Vérifions que les suites $(X_n^N\ ; 1 \leq n \leq N)$ satisfont aux hypothèses du lemme **6.2.1** :
– X_n^N prend ses valeurs dans l'ensemble à 2 points

$$\{\log\frac{1+a_N}{1+r_N}\ ,\ \log\frac{1+b_N}{1+r_N}\} = \{-\frac{\sigma}{\sqrt{N}}\ ,\ +\frac{\sigma}{\sqrt{N}}\}$$

– les hypothèses d'indépendance sont clairement vérifiées
– la vérification de (6.4) nécessite le calcul suivant :

Soient f et $u : \mathbb{R}_+ \mapsto \mathbb{R}$ définies par

$$f(x) = \log\frac{x}{1+r_N} \quad ; \quad u(x) = \frac{(x-(1+a_N))f(1+b_N) + (1+b_N-x)f(1+a_N)}{b_N - a_N}$$

Comme u est égale à f aux 2 points $1+a_N$ et $1+b_N$, $\mathbf{E*}_N[f(T_1^N)] = \mathbf{E*}_N[u(T_1^N)]$; d'autre part , puisque u est affine ,

$$\mathbf{E*}_N[u(T_1^N)] = u(\mathbf{E*}_N[T_1^N]) = u(1+r_N)$$

On a donc , après avoir noté que $f(1+b_N) = -f(1+a_N) = \sigma/\sqrt{N}$,

$$\mu_N = \mathbf{E*}_N[X_1^N] = \mathbf{E*}_N[f(T_1^N)] = u(1+r_N) = \frac{\sigma(2r_N - (a_N+b_N))}{(b_N-a_N)\sqrt{N}}$$

Rappelons alors que

$$\begin{aligned} 1 + a_N &= (1+r_N)\exp(-\frac{\sigma}{\sqrt{N}}) = (1+r_N)[1 - \frac{\sigma}{\sqrt{N}} + \frac{\sigma^2}{2N} + o(1/N)] \\ 1 + b_N &= (1+r_N)\exp(\frac{\sigma}{\sqrt{N}}) = (1+r_N)[1 + \frac{\sigma}{\sqrt{N}} + \frac{\sigma^2}{2N} + o(1/N)] \end{aligned}$$

Il en résulte que

$$b_N - a_N = \frac{2\sigma}{\sqrt{N}} + o(1/N) \ \ ; \ \ 2r_N - (a_N + b_N) = -\frac{\sigma^2}{N}(1+r_N) + o(1/N)$$

On a donc

$$\lim_{N\to\infty} N\mu_N = \lim_{N\to\infty} \frac{-\sigma^2(1+r_N) + No(1/N)}{2(1+r_N) + o(1/N)} = -\sigma^2/2$$

ce qui , compte tenu du lemme **6.2.1** , démontre la proposition .

7.3 Les formules de Black and Scholes

Nous allons obtenir les prix des options d'achat et de vente "à temps continu" en passant à la limite dans les formules correspondantes du modèle de **C.R.R.** . Pour une fois , nous nous intéresserons d'abord , (il y a pour cela des raisons techniques que nous découvrirons plus bas) , aux options de vente .
Plaçons sous la probabilité $\mathbf{P}_N^*$ décrite dans le paragrahe précédent ; le prix P_0^N à l'instant 0 du put de prix d'exercice K est donné par les égalités

$$P_0^N = \frac{1}{(1+r_N)^N}\mathbf{E}_N^*[(K - S_N^N)^+] = \mathbf{E}_N^*[(\frac{K}{(1+r_N)^N} - \tilde{S}_N^N)^+] = \mathbf{E}_N^*[(\frac{K}{(1+r_N)^N} - e^{Y_N})^+]$$

Il s'agit maintenant d''étudier la limite de cette quantité quand $N \to +\infty$. Soit ψ la fonction définie par $\psi(y) = (Ke^{-Rt} - e^y)^+$; on a

$$\begin{aligned} |P_0^N - \mathbf{E}_N^*[\psi(Y_N)]| &= |\mathbf{E}_N^*[(\frac{K}{(1+r_N)^N} - e^{Y_N})^+ - (Ke^{-Rt} - e^{Y_N})^+]| \\ &\leq \mathbf{E}_N^*[|(\frac{K}{(1+r_N)^N} - e^{Y_N})^+ - (Ke^{-Rt} - e^{Y_N})^+|] \\ &\leq K|\frac{1}{(1+r_N)^N} - e^{-Rt}| \to 0 \end{aligned}$$

(La dernière ligne provenant de l'inégalité entre nombres réels : $|a^+ - b^+| \leq |a - b|$) ; il en résulte que

$$\lim_{N\to\infty} P_0^N = \lim_{N\to\infty} \mathbf{E}_N^*[\psi(Y_N)]$$

Mais , d'après la définition même de la convergence en loi , on a , puisqe ψ est continue et bornée , (ce qui ne serait pas le cas pour le calcul analogue relatif au call) ,

$$\lim_{N\to\infty} P_0^N = \int_{\mathbb{R}} \psi(y)\nu(y)dy = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}[Ke^{-Rt} - \exp(-\frac{\sigma^2}{2} + \sigma y)]^+ e^{-y^2/2}dy \qquad (7.10)$$

7.3.1 Introduction d'une valeur initiale

Nous avons jusqu'à présent supposé que la valeur initiale S_0^N del'actif risqué était égale à 1 ; reprenons maintenant les calculs au début du paragraphe **6.2** en supposant que S_0^N est une constante déterministe S_0 (ne dépendant pas de N) .
Appelant S' et P' les valeurs respectives de l'actif sous jacent et du put dans cette nouvelle situation , on a ${S'}_N^N = S_0 S_N^N$ de sorte que

$$ {P'}_0^N = \frac{1}{(1+r_N)^N}\mathbf{E}_N^*[(K - {S'}_N^N)^+] = \frac{S_0}{(1+r_N)^N}\mathbf{E}_N^*[(\frac{K}{S_0} - S_N^N)^+] $$

La limite P de cette expression quand N tend vers l'infini sera le prix du put dans le modèle à temps continu ; il s'écrit , en vertu de (6.10)

$$ P = \frac{S_0}{\sqrt{2\pi}}\int_{\mathbb{R}}[\frac{K}{S_0}e^{-Rt} - exp(-\frac{\sigma^2}{2} + \sigma y)]^+ e^{-\frac{y^2}{2}}dy = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}[Ke^{-Rt} - S_0 exp(-\frac{\sigma^2}{2} + \sigma y)]^+ e^{-\frac{y^2}{2}}dy $$

On exprime d'habitude ce résultat de façon légèrement différente : posons

$$ F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{-y^2}{2}}dy \;\; , \;\; d_1 = \frac{1}{\sigma}(\frac{\sigma^2}{2} + Log\frac{S_0}{K} + Rt) \;\; , \;\; d_2 = d_1 - \sigma. \tag{7.11} $$

On voit facilement que

$$ P = Ke^{-Rt}F(-d_2) - S_0 F(-d_1). \tag{7.12} $$

On procède de la même façon pour définir "à temps continu" la valeur du call de prix d'exercice K et de date d'exercice t : C sera , par définition ,la limite des quantités ${C'}_0^N$ quand $N \to \infty$. La formule de parité écrite pour les processus de **C.R.R.** conduit aux égalités

$$ {C'}_0^N - {P'}_0^N = S_0 - \frac{K}{(1+r_N)^N} $$

D'où , en passant à la limite , $C - P = S_0 - Ke^{-Rt}$; on a donc $C = S_0[1 - F(-d1)] - Ke^{-Rt}[1 - F(-d_2)]$, ce qui peut encore s'écrire

$$ C = S_0 F(d_1) - Ke^{-Rt}F(d_2) \tag{7.13} $$

Les égalités (6.11) et (6.13) constituent les "formules de **Black and Scholes .**"