

# **MODELISATION FINANCIERE**

# **ET**

# **GESTION DE PORTEFEUILLE**

*Louis ESCH*

**Master complémentaire en gestion des risques financiers**  
**Option « Finance »**

## RAPPELS

1. Mathématique	4
1.1. Calcul matriciel	4
1.2. Fonctions d'une variable	10
1.3. Fonctions de plusieurs variables	13
2. Probabilités	16
2.1. Variables aléatoires	16
2.2. Aléas multivariés	23
2.3. Processus stochastiques	26
3. Statistique	30
3.1. Echantillonnage et estimation	30
3.2. Régressions	33
4. Tableur	35
4.1. Notion de formule matricielle	35
4.2. Quelques fonctions matricielles	36
4.3. Quelques fonctions statistiques	37
Table de la distribution normale	41

## INTRODUCTION

Deux éléments fondamentaux	42
Deux principes de base	43
Notion de modèle d'évaluation	44

## ACTIONS

1. Fondements	47
1.1. Rendement et risque	47
1.2. Efficience des marchés	58
1.3. Modèles d'évaluation d'une action	62
2. Diversification et gestion de portefeuille	66
2.1. Principes de diversification	66
2.2. Diversification et taille du portefeuille	70
2.3. Modèle de Markowitz et algorithme de la ligne critique	72
2.4. Modèle à indice simple de Sharpe	89
2.5. Modèle avec titre non risqué	96
2.6. Méthode d'Elton, Gruber et Padberg	102
2.7. Théorie de l'utilité et sélection du portefeuille optimal	108
2.8. Le modèle de marché	114
3. Modèle d'équilibre des actifs financiers et applications	117
3.1. Capital asset pricing model	117
3.2. Arbitrage pricing theory	122
3.3. Evaluation des performances	125
3.4. Stratégies de gestion d'un portefeuille d'actions	129
4. Modèles d'évolution des actions	132
4.1. Modèles déterministes	135
4.2. Modèles stochastiques	137

## **OBLIGATIONS**

1. Caractéristiques et évaluation	142
1.1. Définitions	142
1.2. Rendement d'une obligation	143
1.3. Evaluation d'une obligation	147
2. Obligations et risque financier	148
2.1. Sources de risque	148
2.2. Duration	150
2.3. Convexité	157
3. Structure déterministe des taux d'intérêt	159
3.1. Courbes de taux	159
3.2. Structure statique de taux d'intérêt	160
3.3. Structure dynamique de taux d'intérêt	164
3.4. Modélisation déterministe et modélisation aléatoire	166
4. Stratégies de gestion d'un portefeuille obligataire	167
4.1. Stratégie passive : immunisation	167
4.2. Stratégie active	170
5. Modèles stochastiques d'évolution des obligations	172
5.1. Modèles d'arbitrage à une variable d'état	173
5.2. Modèle de Vasicek	176
5.3. Modèle de Cox, Ingersoll et Ross	180

## **OPTIONS**

1. Définitions	184
1.1. Caractéristiques	184
1.2. Utilisation	186
2. Valeur d'une option	188
2.1. Valeur intrinsèque et valeur-temps	188
2.2. Volatilité	191
2.3. Paramètres de sensibilité	192
2.4. Propriétés générales	195
3. Modèles d'évaluation	198
3.1. Modèle binomial pour les options sur action	200
3.2. Modèle de Black et Scholes pour les options sur action	208
3.3. Autres modèles d'évaluation	215

# CHAPITRE 1

## MATHEMATIQUE, PROBABILITES, STATISTIQUE, TABLEUR : RAPPELS

### 1.1. MATHEMATIQUE

#### 1.1.1. Calcul matriciel

##### *Matrices et vecteurs*

Nous appellerons *matrice* d'ordre  $n$  un ensemble de  $n^2$  nombres réels constituant un tableau carré formé de  $n$  lignes et de  $n$  colonnes<sup>1</sup>. Une matrice est généralement représentée par une lettre majuscule ( $A$  par exemple), ses éléments se représentant par la lettre minuscule correspondante ( $a$ ) affectée de deux indices repérant respectivement la ligne et la colonne d'appartenance de l'élément :  $a_{ij}$  est l'élément de la matrice  $A$  situé à l'intersection de la  $i$ -ème ligne et de la  $j$ -ème colonne de  $A$ . La matrice  $A$  peut donc s'écrire de manière générale comme suit :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}.$$

De même, un *vecteur* de dimension  $n$  est un ensemble de  $n$  nombres réels formant un tableau d'une colonne. Les éléments d'un vecteur sont les composantes de celui-ci et se repèrent par un seul indice :

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}.$$

---

<sup>1</sup> D'une manière plus générale, une matrice est en fait un tableau rectangulaire de format  $(m, n)$  :  $m$  lignes et  $n$  colonnes. Nous n'aurons néanmoins besoin ici que des matrices carrées.

### Matrices particulières

Les éléments diagonaux d'une matrice  $A$  sont les éléments  $a_{11}, a_{22}, \dots, a_{nn}$ . Ils sont situés sur la diagonale du tableau qui part du coin supérieur gauche ; celle-ci porte le nom de diagonale principale.

Une matrice est qualifiée de *symétrique* si ses éléments symétriques par rapport à la diagonale principale sont égaux :  $a_{ij} = a_{ji}$ . En voici un exemple :

$$A = \begin{pmatrix} 2 & -3 & 0 \\ -3 & 1 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix}.$$

Une matrice *triangulaire supérieure* est une matrice dont les éléments situés au-dessous de la diagonale principale sont nuls :  $a_{ij} = 0$  lorsque  $i < j$ . Par exemple,

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix};$$

On définit bien sûr la notion de matrice triangulaire inférieure de la même manière.

Enfin, une matrice *diagonale* est une matrice à la fois triangulaire supérieure et triangulaire inférieure. Ses seuls éléments éventuellement non nuls sont les éléments diagonaux :  $a_{ij} = 0$  lorsque  $i$  et  $j$  sont différents. D'une manière générale, une telle matrice peut se représenter par

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} = \text{diag}(a_1, a_2, \dots, a_n).$$

### Opérations

La somme de deux matrices ainsi que la multiplication d'une matrice par un scalaire sont des opérations très naturelles : on effectue l'opération en question pour chacun des éléments. Ainsi,

$$\begin{aligned} (A + B)_{ij} &= a_{ij} + b_{ij} \\ (\lambda A)_{ij} &= \lambda a_{ij} \end{aligned}$$

Les définitions en question sont encore valables pour les vecteurs :

$$(X + Y)_i = x_i + y_i$$

$$(\lambda X)_i = \lambda x_i$$

Le produit de deux matrices  $A$  et  $B$  est une matrice de même ordre que  $A$  et  $B$  dont l'élément  $(i, j)$  s'obtient en effectuant la somme des produits des éléments de la  $i$ -ème ligne de  $A$  par les éléments correspondants de la  $j$ -ème colonne de  $B$  :

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj}.$$

On a par exemple

$$\begin{pmatrix} 2 & 0 & -1 \\ 3 & -2 & 1 \\ -3 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 5 & -2 \\ 3 & -1 & 0 \\ 2 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 10 & -3 \\ -4 & 17 & -7 \\ 6 & -17 & 6 \end{pmatrix}.$$

Malgré sa définition apparemment complexe, le produit matriciel possède un certain nombre de propriétés classiques : il est associatif et distributif par rapport à l'addition. Il convient néanmoins de le manipuler avec prudence car une de ces propriétés classiques lui manque : il n'est pas commutatif : en général,  $AB \neq BA$  !

On définit par ce même procédé « lignes par colonnes » le produit d'une matrice par un vecteur :

$$(AX)_i = \sum_{k=1}^n a_{ik}x_k.$$

La transposée d'une matrice  $A$  est la matrice  $A^t$  obtenue en permutant les éléments symétriques par rapport à la diagonale principale ou, ce qui revient au même, en permutant le rôle des lignes et des colonnes de la matrice  $A$  :

$$(A^t)_{ij} = a_{ji}.$$

Ainsi, une matrice est symétrique si et seulement si elle est égale à sa transposée. De plus, cette opération, appliquée à un vecteur, donne comme résultat le vecteur-ligne correspondant.

L'inverse de la matrice  $A$  est une matrice  $A^{-1}$ , lorsqu'elle existe, telle que

$$AA^{-1} = A^{-1}A = \text{diag}(1, \dots, 1).$$

Par exemple, on vérifie aisément que

$$\begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & -3 \\ 0 & 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 0 & 1 \\ -2 & -1 & 1 \end{pmatrix}.$$

Définissons enfin la trace d'une matrice. Il s'agit de la somme des éléments diagonaux de celle-ci :

$$\text{tr}(A) = a_{11} + a_{22} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}.$$

### **Forme quadratique et classe d'une matrice symétrique**

Une *forme quadratique* est une fonction polynomiale de  $n$  variables ne contenant que des termes du second degré :

$$Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j.$$

Si on forme la matrice  $A$  des coefficients  $a_{ij}$  ( $i, j = 1, \dots, n$ ) et le vecteur  $X$  des variables  $x_i$  ( $i = 1, \dots, n$ ), on peut donner une expression matricielle pour la forme quadratique :

$$Q(X) = X^t A X.$$

En effet, en développant le membre de droite, on a

$$\begin{aligned} X^t A X &= \sum_{i=1}^n x_i (AX)_i \\ &= \sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} x_j \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \end{aligned}$$

A une matrice  $A$ , on peut donc toujours associer une forme quadratique et réciproquement. Cependant, la matrice n'est pas unique. En effet, à la forme quadratique

$$Q(x_1, x_2) = 3x_1^2 - 4x_1 x_2,$$

on peut associer les matrices

$$A = \begin{pmatrix} 3 & -2 \\ -2 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 0 \\ -4 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 3 & -6 \\ 2 & 0 \end{pmatrix}$$

ainsi qu'une infinité d'autres. Parmi toutes ces matrices, une seule est symétrique ( $A$  dans l'exemple ci-dessus). Il y a donc bijection entre l'ensemble des formes quadratiques et l'ensemble des matrices symétriques.

On définit alors la *classe* d'une matrice symétrique à partir du signe de la forme quadratique associée. Ainsi, la matrice symétrique non nulle  $A$  est dite définie positive (d.p.) si

$$X'AX > 0 \text{ pour tout } X \neq 0$$

et elle est dite semi-positive (s.p.) lorsque

$$\begin{cases} X'AX \geq 0 \text{ pour tout } X \neq 0 \\ \text{il existe } Y \neq 0 \text{ tel que } Y'AY = 0 \end{cases}$$

On définit une matrice définie négative (d.n.) et semi-négative (s.n.) par les inégalités inverses et on qualifie d'indéfinie une matrice symétrique pour laquelle

$$\text{il existe } X \text{ et } Y \neq 0 \text{ tels que } X'AX > 0 \text{ et } Y'AY < 0.$$

Ainsi, la matrice symétrique

$$A = \begin{pmatrix} 5 & -3 & -4 \\ -3 & 10 & 2 \\ -4 & 2 & 8 \end{pmatrix},$$

est d.p. car la forme quadratique associée peut s'écrire

$$\begin{aligned} Q(x, y, z) &= 5x^2 + 10y^2 + 8z^2 - 6xy - 8xz + 4yz \\ &= (x - 3y)^2 + (2x - 2z)^2 + (y + 2z)^2 \end{aligned}$$

Celle-ci n'est jamais négative et elle s'annule uniquement lorsque

$$\begin{cases} x - 3y = 0 \\ 2x - 2z = 0 \\ y + 2z = 0 \end{cases}$$

c'est-à-dire lorsque  $x = y = z = 0$ .

### ***Système d'équations linéaires***

Un système de  $n$  équations linéaires à  $n$  inconnues est un ensemble de relations du type

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$$

où les  $a_{ij}$ , les  $x_j$  et les  $b_i$  sont respectivement les coefficients, les inconnues et les seconds membres. Ils s'écrivent de manière naturelle sous forme matricielle et vectorielle :  $A$ ,  $X$  et  $B$ . Avec ces notations, le système s'écrit de manière plus condensée mais équivalente



$$AX = B.$$

Par exemple, le système d'équations

$$\begin{cases} 2x + 3y = 4 \\ 4x - y = -2 \end{cases}$$

s'écrit de manière équivalente

$$\begin{pmatrix} 2 & 3 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}.$$

Si l'inverse de la matrice  $A$  existe, il est facile de voir que le système admet une et une solution, donnée par  $X = A^{-1}B$ .

### 1.1.2. Fonctions d'une variable

#### *Dérivée*

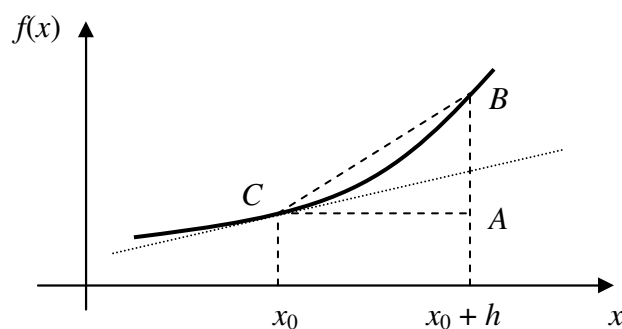
La *dérivée*<sup>2</sup> de la fonction  $f$  au point  $x_0$  se définit par

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

si cette limite existe et est finie.

Si la fonction  $f$  est dérivable en tous les points d'un intervalle ouvert  $]a; b[$ , elle constitue une nouvelle fonction définie dans cet intervalle : la fonction dérivée, notée  $f'$ .

Pour une petite valeur de  $h$ , le numérateur de la définition représente l'accroissement (ou la diminution) de valeur de la fonction lorsque la variable  $x$  passe de la valeur  $x_0$  à la valeur voisine  $(x_0 + h)$ , c'est-à-dire la longueur du segment  $AB$  de la figure.



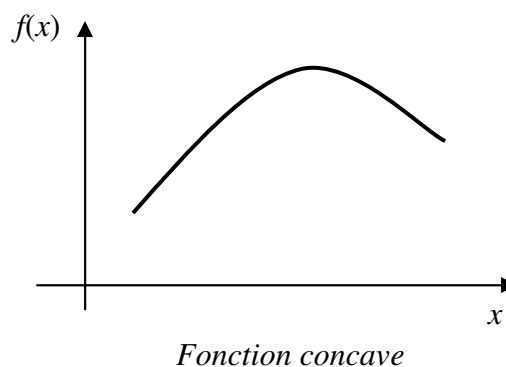
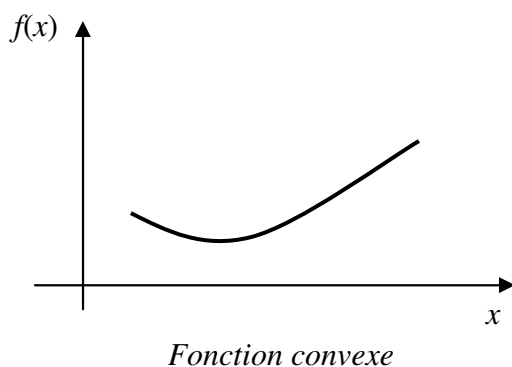
*Interprétation géométrique de la dérivée*

<sup>2</sup> On dit aussi *dérivée première*.

Le dénominateur de la même expression,  $h$ , est égal quant à lui à la longueur du segment  $AC$ . Le rapport est donc égal à la pente de la droite  $BC$ . Lorsque  $h$  tend vers 0, cette droite  $BC$  se déplace vers la tangente au graphe de la fonction au point  $C$ .

L'interprétation géométrique de la dérivée est donc la suivante :  $f'(x_0)$  représente la pente de la tangente au graphe de  $f$  au point d'abscisse  $x_0$ . En particulier, le signe de la dérivée caractérise le type de variation de la fonction : à une dérivée positive (resp. négative) correspond une fonction *croissante* (resp. *décroissante*). La dérivée mesure donc la vitesse à laquelle la fonction croît (ou décroît) au voisinage d'un point.

La dérivée de la dérivée, appelée dérivée seconde et notée  $f''$ , sera donc positive lorsque la fonction  $f'$  est croissante, c'est-à-dire lorsque la pente de la tangente au graphe de  $f$  augmente lorsque la variable  $x$  augmente : on dit que la fonction est *convexe*. A l'opposé, une fonction dont la dérivée seconde est négative est *concave*.



Rappelons enfin les règles élémentaires de calcul des dérivées. Celles relatives aux opérations entre fonctions tout d'abord :

$$\begin{aligned}(f + g)' &= f' + g' \\ (\lambda f)' &= \lambda f' \\ (fg)' &= f'g + fg' \\ \left(\frac{f}{g}\right)' &= \frac{f'g - fg'}{g^2}\end{aligned}$$

celle concernant les fonctions composées ensuite :

$$[g(f)]' = g'(f) \cdot f'$$

et enfin les formules donnant les dérivées de quelques fonctions élémentaires :

$$\begin{aligned}(x^m)' &= mx^{m-1} \\ (e^x)' &= e^x \\ (a^x)' &= a^x \ln a \\ (\ln x)' &= \frac{1}{x} \\ (\log_a x)' &= \frac{1}{x \ln a}\end{aligned}$$

## **Extrema**

Le point  $x_0$  est un *maximum* (resp. *minimum*) *local* de la fonction  $f$  si

$$f(x_0) \geq f(x) \quad (\text{resp. } f(x_0) \leq f(x))$$

pour tout  $x$  voisin de  $x_0$ .

La détermination des extrema d'une fonction dérivable se réalise grâce à deux conditions.

La condition (nécessaire) du premier ordre exprime que si  $x_0$  est un *extremum* de  $f$ , alors  $f'(x_0) = 0$ . En un tel point, appelé *point stationnaire*, la tangente au graphe de  $f$  est donc horizontale.

La condition (suffisante) du second ordre permet de « trier » les points stationnaires selon leur nature : si  $x_0$  est un point stationnaire de  $f$  et que  $f''(x_0) > 0$ , il s'agit d'un minimum ; à l'opposé, si  $f''(x_0) < 0$ , il s'agit d'un maximum.

## **Formule de Taylor**

Considérons une fonction  $f$  qu'on désire étudier au voisinage de  $x_0$  (disons en  $x_0 + h$ ). Une manière de procéder est de remplacer celle-ci par un polynôme – fonction facile à manipuler – de la variable  $h$  :

$$f(x_0 + h) = a_0 + a_1 h + a_2 h^2 + \dots$$

Pour avoir une bonne représentation de la fonction  $f$  par le polynôme, il faut que tous deux

- prennent la même valeur en  $h = 0$  ;
- aient la même pente (c'est-à-dire la même dérivée première) en  $h = 0$  ;
- aient la même convexité ou concavité (c'est-à-dire la même dérivée seconde) en  $h = 0$  ;

et ainsi de suite, le nombre de conditions à imposer devant correspondre au nombre de coefficients à déterminer dans le polynôme. Il est facile de montrer que ces conditions conduisent à

$$\begin{aligned}
a_0 &= f(x_0) = \frac{f(x_0)}{0!} \\
a_1 &= f'(x_0) = \frac{f'(x_0)}{1!} \\
a_2 &= \frac{f''(x_0)}{2} = \frac{f''(x_0)}{2!} \\
&\dots \\
a_k &= \frac{f^{(k)}(x_0)}{k!} \\
&\dots
\end{aligned}$$

D'une manière générale, on peut donc écrire

$$f(x_0 + h) = \frac{f(x_0)}{0!} + \frac{f'(x_0)}{1!}h + \frac{f''(x_0)}{2!}h^2 + \dots + \frac{f^{(n)}(x_0)}{n!}h^n + R_n,$$

où  $R_n$ , appelé reste du développement, est l'écart entre la fonction  $f$  à étudier et le polynôme d'approximation. Ce reste est, sous certaines conditions de régularité, négligeable en ce sens que lorsque  $h$  tend 0, il tend vers 0 plus rapidement que  $h^n$ , ce qu'on note  $o(h^n)$ .

L'utilisation qui est faite de la formule de Taylor dans cet ouvrage ne nécessite pas la considération d'un polynôme de degré élevé et on écrira donc plus simplement

$$f(x_0 + h) \approx f(x_0) + \frac{f'(x_0)}{1!}h + \frac{f''(x_0)}{2!}h^2 + \frac{f'''(x_0)}{3!}h^3 + \dots$$

Pour certaines fonctions élémentaires, le développement de Taylor prend une forme particulière qu'il est utile de conserver en mémoire :

$$\begin{aligned}
e^x &\approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \\
(1+x)^m &\approx 1 + \frac{m}{1!}x + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \dots \\
\ln(1+x) &\approx x - \frac{x^2}{2} + \frac{x^3}{3} - \dots
\end{aligned}$$

### **Série géométrique**

Si, dans la formule de Taylor pour  $(1+x)^m$ , on remplace  $x$  par  $(-x)$  et  $m$  par  $(-1)$ , on obtient

$$\frac{1}{1-x} \approx 1 + x + x^2 + x^3 + \dots$$

Il n'est pas difficile de démontrer que lorsque  $|x| < 1$ , la suite

$$\begin{array}{c}
1 \\
1+x \\
1+x+x^2 \\
\vdots \\
1+x+x^2+\dots+x^n \\
\vdots
\end{array}$$

converge vers le nombre  $1/(1+x)$ . La limite d'une telle suite est donc une somme comprenant une infinité de termes et porte le nom de *série*. Celle qui nous occupe ici est la *série géométrique* :

$$1+x+x^2+\dots+x^n+\dots = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

### 1.1.3. Fonctions de plusieurs variables

#### *Dérivées partielles*

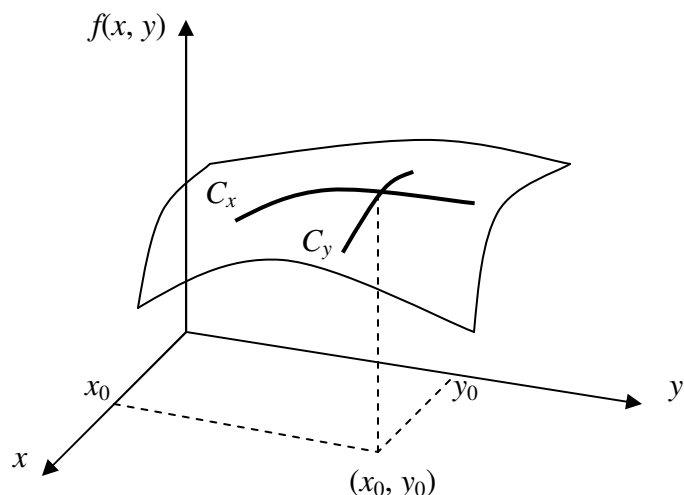
Pour une fonction  $f$  de  $n$  variables  $x_1, x_2, \dots, x_n$ , la notion de dérivée se définit de manière similaire, si ce n'est que l'accroissement  $h$  peut porter sur n'importe quelle variable. On aura donc  $n$  notions de dérivées, relatives à chacune des  $n$  variables, et elles portent le nom de *dérivées partielles*. Ainsi la dérivée partielle de  $f(x_1, x_2, \dots, x_n)$  par rapport à  $x_k$  au point  $(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$  se définit par

$$f'_{x_k}(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) = \lim_{h \rightarrow 0} \frac{f(x_1^{(0)}, x_2^{(0)}, \dots, x_k^{(0)} + h, \dots, x_n^{(0)}) - f(x_1^{(0)}, x_2^{(0)}, \dots, x_k^{(0)}, \dots, x_n^{(0)})}{h}$$

L'interprétation géométrique des dérivées partielles ne peut guère s'envisager que pour les fonctions de deux variables puisque le graphe d'une telle fonction est plongé dans l'espace à trois dimensions (une dimension pour chacune des deux variables et la troisième, l'ordonnée, pour les valeurs de la fonction). C'est ainsi que nous examinerons les dérivées partielles

$$\begin{aligned}
f'_x(x_0, y_0) &= \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \\
f'_y(x_0, y_0) &= \lim_{h \rightarrow 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}
\end{aligned}$$

Considérons donc le graphe de cette fonction  $f(x, y)$ . Il s'agit d'une surface de l'espace à trois dimensions.



### Interprétation géométrique des dérivées partielles

Considérons par ailleurs le plan vertical passant par le point  $(x_0, y_0)$  et parallèle à l'axe  $Ox$  ; son intersection avec le graphe de  $f$  est la courbe  $C_x$ . Le même raisonnement que celui tenu pour les fonctions d'une variable montre que la dérivée partielle  $f'_x(x_0, y_0)$  est égale à la pente de la tangente à cette courbe  $C_x$  au point d'abscisse  $(x_0, y_0)$  (c'est-à-dire la pente du graphe de  $f$  dans la direction de  $x$ ). De la même manière,  $f'_y(x_0, y_0)$  représente la pente de la tangente à  $C_y$  au point d'abscisse  $(x_0, y_0)$ .

### Extrema libres

Le point  $(x_1^{(0)}, \dots, x_n^{(0)})$  est un *maximum* (resp. *minimum*) local de la fonction  $f$  si

$$f(x_1^{(0)}, \dots, x_n^{(0)}) \geq f(x_1, \dots, x_n) \quad \left[ \text{resp. } f(x_1^{(0)}, \dots, x_n^{(0)}) \leq f(x_1, \dots, x_n) \right]$$

pour tout  $(x_1, \dots, x_n)$  voisin de  $(x_1^{(0)}, \dots, x_n^{(0)})$ .

Comme pour les fonctions d'une seule variables, la détermination des extrema d'une fonction dérivable se réalise grâce à deux conditions.

La condition (nécessaire) du premier ordre exprime que si  $x_0$  est un *extremum* de  $f$ , alors toutes les dérivées partielles de  $f$  sont nulles en  $x^{(0)}$  :

$$f'_{x_i}(x^{(0)}) = 0 \quad (i = 1, \dots, n).$$

En faisant référence à l'interprétation géométrique des dérivées partielles d'une fonction de deux variables, en un tel point  $(x_0, y_0)$ , appelé *point stationnaire*, les tangentes aux courbes  $C_x$  et  $C_y$  sont donc horizontales.

La condition (suffisante) du second ordre permet de « trier » les points stationnaires selon leur nature, mais nécessite tout d'abord de définir la *matrice hessienne* de la fonction  $f$  au point  $x$ , constituée des dérivées partielles secondes de  $f$  :

$$H(f(x_1, \dots, x_n)) = \begin{pmatrix} f''_{x_1 x_1}(x) & f''_{x_1 x_2}(x) & \cdots & f''_{x_1 x_n}(x) \\ f''_{x_2 x_1}(x) & f''_{x_2 x_2}(x) & \cdots & f''_{x_2 x_n}(x) \\ \vdots & \vdots & & \vdots \\ f''_{x_n x_1}(x) & f''_{x_n x_2}(x) & \cdots & f''_{x_n x_n}(x) \end{pmatrix}.$$

Si  $x^{(0)}$  est un point stationnaire de  $f$  et que  $H(f(x))$  est d.p. en  $x^{(0)}$  ou est s.p. dans un voisinage de  $x^{(0)}$ , il s'agit d'un minimum ; à l'opposé, si  $H(f(x))$  est d.n.  $x^{(0)}$  ou est s.n. dans un voisinage de  $x^{(0)}$ , il s'agit d'un maximum.

### Extrema liés

Il s'agit ici d'une notion similaire, mais où on restreint l'analyse du problème d'extrema aux seuls  $x$  qui obéissent à une ou plusieurs contraintes.

Le point  $(x_1^{(0)}, \dots, x_n^{(0)})$  est un *maximum* (resp. *minimum*) *local* de la fonction  $f$  sous les contraintes

$$\begin{cases} g_1(x) = 0 \\ \dots \\ g_r(x) = 0 \end{cases}$$

si  $x^{(0)}$  vérifie lui-même les contraintes et que

$$f(x_1^{(0)}, \dots, x_n^{(0)}) \geq f(x_1, \dots, x_n) \quad \left[ \text{resp. } f(x_1^{(0)}, \dots, x_n^{(0)}) \leq f(x_1, \dots, x_n) \right]$$

pour tout  $(x_1, \dots, x_n)$  voisin de

$$\begin{cases} \text{voisin de } (x_1^{(0)}, \dots, x_n^{(0)}) \\ \text{vérifiant les } r \text{ contraintes} \end{cases}$$

La résolution de ce problème passe par la considération du *lagrangien* du problème. Il s'agit d'une fonction des  $(n + r)$  variables  $(x_1, \dots, x_n; m_1, \dots, m_r)$ , les  $r$  dernières – appelées *multiplieurs de Lagrange* – correspondant chacune à une contrainte :

$$L(x_1, \dots, x_n; m_1, \dots, m_r) = f(x) + m_1 \cdot g_1(x) + \dots + m_r \cdot g_r(x).$$

Nous n'entrerons pas dans les détails techniques de la résolution de ce problème. Signalons néanmoins un résultat essentiel : si le point  $(x^{(0)}; m^{(0)})$  est tel que  $x^{(0)}$  vérifie les contraintes et que  $(x^{(0)}; m^{(0)})$  est un extremum (libre) du lagrangien, alors  $x^{(0)}$  est un extremum pour le problème d'extrema liés.

## 1.2. PROBABILITES

### 1.2.1. Variables aléatoires

#### *Variable aléatoire et loi de probabilité*

Considérons un phénomène fortuit, c'est-à-dire un phénomène pour lequel, à des conditions initiales données, correspondent plusieurs résultats possibles. On note généralement  $\Omega$  l'ensemble de ces résultats possibles, ces derniers étant représentés par  $\omega$ . Une grandeur numérique dépendant du résultat observé porte le nom de *variable aléatoire* (v.a.).

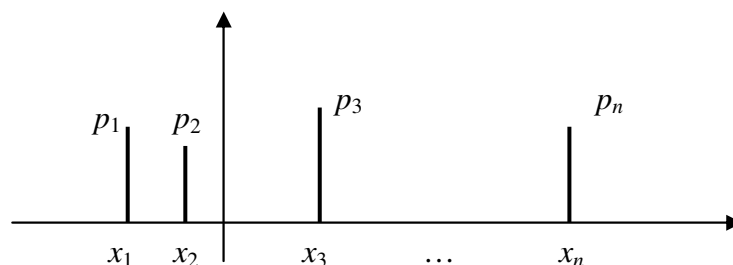
Par ailleurs, des probabilités sont associées aux différents résultats possibles ou aux événements définis dans le cadre de ce phénomène fortuit. Il est donc intéressant de connaître la probabilités des divers événements qui sont définis à partir de cette v.a. Il s'agit là de la notion de *loi de probabilité* de la v.a. Ainsi, si la v.a. se note  $X$ , la loi de probabilité de  $X$  se définit par l'ensemble des probabilités suivantes :  $\Pr[X \in A]$ , quel que soit le sous-ensemble  $A$  de  $\mathbf{R}$ .

#### *Variable aléatoire discrète*

Une *v.a. discrète* correspond à une situation où l'ensemble des valeurs possibles de cette variable est fini ou infini dénombrable. Dans ce cas, si on connaît la liste des différentes valeurs possibles (parfois appelées *atomes*) et les probabilités correspondantes

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$
$$\Pr[X = x_i] = p_i \quad i = 1, 2, \dots, n, \dots$$
$$\sum_i p_i = 1$$

Généralement, on représente graphiquement une telle v.a. par un diagramme en bâtonnets où les abscisses correspondent aux valeurs possibles et les hauteurs des bâtonnets aux probabilités correspondantes.



*Variable aléatoire discrète*



On peut déterminer aisément la loi de probabilité de  $X$  :

$$\Pr[X \in A] = \sum_{\{i: x_i \in A\}} p_i .$$

En particulier, une v.a. est qualifiée de *dégénérée* si elle ne peut prendre qu'une seule valeur  $x$  (on parle aussi de variable certaine) :

$$\Pr[X = x] = 1 .$$

### ***Variable aléatoire à densité***

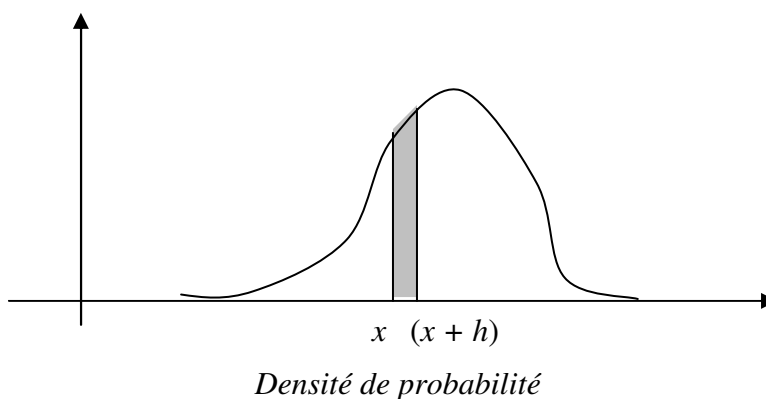
A l'opposé des v.a. discrètes, il arrive que l'ensemble des valeurs possibles d'une v.a. soit continu (un intervalle par exemple) et qu'aucune valeur individuelle n'ait de probabilité strictement positive :

$$\Pr[X = x] = 0 \quad \forall x .$$

Dans ce cas, la répartition des probabilités sur l'ensemble des valeurs possibles ne peut se mesurer que sur des ensembles du type intervalles et s'exprime par l'intermédiaire d'une fonction de *densité*  $f$  : pour  $h$  suffisamment petit, on a

$$\Pr[x < X \leq x + h] \approx hf(x) .$$

Cette définition est illustrée sur la figure ci-après.



La loi de probabilité s'obtient à partir de la densité par la relation suivante :

$$\Pr[X \in A] = \int_A f(x) dx$$

et on a bien entendu

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

### ***Valeurs typiques d'une variable aléatoire***

A côté de l'information complète que constitue la loi de probabilité d'une v.a. (ou sa densité si celle-ci existe), on peut être intéressé par un résumé des propriétés de cette loi de probabilité par un nombre limité d'indicateurs : les *valeurs typiques*. Les principales catégories de ces indicateurs sont

- les *indices de localisation* ou *valeurs centrales*, qui partagent la loi de probabilité en deux parties « similaires » ;
- les *indices de dispersion*, qui mesurent l'étalement de la loi de probabilité autour d'une valeur centrale ;
- les *indices de dissymétrie*, qui détectent la concentration de la loi de probabilité à gauche ou à droite d'une valeur centrale ;
- les *indices d'aplatissement*, qui détectent la présence de queues de distribution plus ou moins épaisses que celles de la loi normale<sup>3</sup> (voir plus loin).

Nous envisagerons ci-après seulement une valeur centrale (la moyenne) et deux indices de dispersion (la variance et l'écart type). Par ailleurs, nous ne donnons pas les définitions générales, mais seulement les formules de calcul pour les cas particuliers des v.a. discrètes et des v.a. à densité, ainsi que les propriétés les plus importantes

### ***Moyenne d'une variable aléatoire***

La *moyenne* – ou encore l'*espérance (mathématique)* – d'une v.a. est la plus importante des valeurs centrales. Elle se note indifféremment  $\mu$  ou  $E(X)$ , ou encore  $\mu_X$  si une confusion est possible. Elle se calcule, pour une v.a. discrète et pour une v.a. à densité, par les formules

$$\mu = \sum_i x_i p_i$$

$$\mu = \int_{-\infty}^{+\infty} x f(x) dx$$

qui montrent que  $\mu$  intègre les différentes valeurs possibles de la v.a.  $X$  en les pondérant par les probabilités associées à ces valeurs.

On montre que l'espérance est un opérateur linéaire :

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

et que si  $X$  et  $Y$  sont indépendants, alors

$$E(XY) = E(X) \cdot E(Y).$$

De plus, pour une v.a. discrète ou une v.a. à densité, l'espérance d'une fonction de la v.a. est donnée par

<sup>3</sup> Les distributions à queues plus épaisses que celles de la loi normale sont qualifiées de *leptocurtiques*.

$$E(g(X)) = \sum_i g(x_i) p_i$$

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

Rappelons enfin la *loi des grands nombres*<sup>4</sup> qui, pour une suite de v.a.  $X_1, X_2, \dots, X_n, \dots$  indépendantes, identiquement distribuées et de moyenne  $\mu$ , exprime que, quel que soit  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \Pr \left[ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| \leq \varepsilon \right] = 1.$$

Cette loi justifie que l'on prenne la moyenne d'un échantillon pour estimer la moyenne de la population dont il est extrait et, en particulier, que l'on évalue la probabilité d'un événement par la fréquence d'occurrence de cet événement lors d'un grand nombre de réalisations du phénomène fortuit.

### ***Variance et écart type d'une variable aléatoire***

Un des indices de dispersion les plus utilisés est la variance  $\sigma^2$  qui se définit par

$$\sigma^2 = \text{var}(X) = E[(X - \mu)^2].$$

En développant le membre de droite, on peut aussi écrire la variance

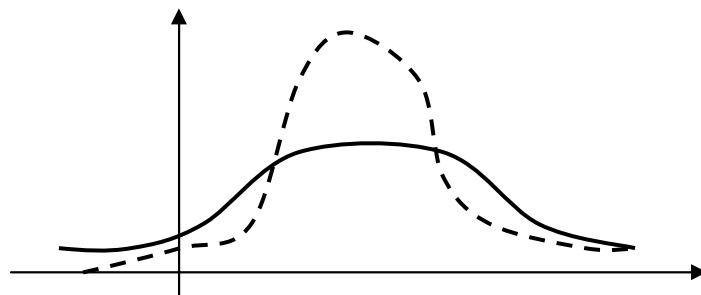
$$\sigma^2 = E(X^2) - \mu^2,$$

ce qui donne, pour une v.a. discrète et pour une v.a. à densité,

$$\sigma^2 = \sum_i (x_i - \mu)^2 p_i = \sum_i x_i^2 p_i - \mu^2$$

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 f(x) dx - \mu^2$$

Une illustration de la notion de dispersion se trouve à la figure ci-dessous.



*Dispersion d'une variable aléatoire*

<sup>4</sup> Nous présentons ici cette loi sous sa forme faible.

On peut démontrer que

$$\text{var}(aX + b) = a^2 \text{var}(X)$$

et que si  $X$  et  $Y$  sont indépendants, alors

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y).$$

De plus, la variance d'une v.a. est nulle si et seulement si celle-ci est *dégénérée*, c'est-à-dire qu'elle ne peut prendre qu'une seule valeur (et n'est donc pas aléatoire).

A côté de la variance, dont la dimension est le carré de la dimension de  $X$ , on utilise également l'écart type, qui en est simplement la racine carrée :

$$\sigma = \sqrt{\text{var}(X)}.$$

Les propriétés se déduisent de celles de la variance ; en particulier,

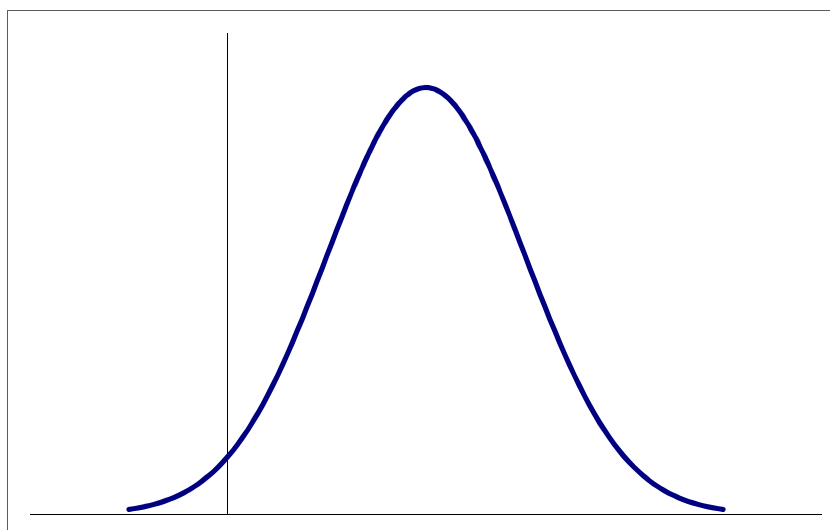
$$\sigma_{aX+b} = |a| \cdot \sigma_X.$$

### ***Distribution normale***

Rappelons qu'une v.a. *normale* de paramètres  $(\mu ; \sigma)$  est définie par sa densité :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],$$

dont le graphe est représenté ci-après.



*Densité normale*

Le graphe de la densité normale est symétrique par rapport à la droite verticale d'abscisse  $\mu$  et il présente deux points d'inflexion, en  $(\mu - \sigma)$  et  $(\mu + \sigma)$ .

Les valeurs typiques de cette distribution sont données par

$$E(X) = \mu$$

$$\text{var}(X) = \sigma^2$$

Si la v.a.  $X$  est distribuée suivant une loi normale de paramètres  $(\mu ; \sigma)$ , on peut démontrer que la v.a.  $(aX + b)$  est également distribuée suivant une loi normale. En particulier, la v.a.

$$\frac{X - \mu}{\sigma},$$

suit une loi normale de paramètres  $(0 ; 1)$ . Cette dernière porte le nom de *loi normale réduite*.

Du point de vue des valeurs numériques, les probabilités associées à une v.a.  $Y$  normale réduite se calculent à partir de la table située à la fin de ce chapitre, et qui recense les valeurs de

$$\Phi(t) = \Pr[Y \leq t] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx ;$$

On passe aux probabilités associées à une v.a.  $X$  normale de paramètres  $(\mu ; \sigma)$  à l'aide de la relation  $X = \mu + \sigma Y$ .

Il est par ailleurs intéressant de connaître approximativement la valeur de la probabilité  $\Pr[\mu - k\sigma \leq X \leq \mu + k\sigma]$  qu'une v.a.  $X$  normale de paramètres  $(\mu ; \sigma)$  ne s'écarte pas de sa moyenne  $\mu$  de plus de  $k$  fois son écart type  $\sigma$ :

k	Pr
2/3	0,5
1	0,68
2	0,95
3	0,997

L'importance de cette loi normale en théorie des probabilités et en statistique provient du célèbre *théorème de tendance normale*<sup>5</sup>. Celui-ci exprime que si les v.a.  $X_1, X_2, \dots, X_n, \dots$

- sont indépendantes ;
- sont de moyenne  $\mu_k$  et d'écart type  $\sigma_k$  fini ( $k = 1, \dots, n, \dots$ ) ;
- et n'ont aucune variance prépondérante par rapport à l'ensemble :

$$\lim_{n \rightarrow \infty} \frac{\sigma_k^2}{\sigma_1^2 + \dots + \sigma_n^2} = 0 \quad \forall k ;$$

<sup>5</sup> *Central limit theorem* en anglais ; presque toujours appelé *Théorème central limite* en fort laid français.

alors la distribution de la v.a.

$$\frac{(X_1 + \dots + X_n) - (\mu_1 + \dots + \mu_n)}{\sqrt{\sigma_1^2 + \dots + \sigma_n^2}},$$

tend, lorsque  $n$  tend vers l'infini, vers une loi normale réduite.

Beaucoup plus intuitivement, le théorème de tendance normale exprime que la somme d'un grand nombre d'effets indépendants, où aucun n'a une variabilité importante par rapport à l'ensemble, est distribuée suivant une loi normale et ce sans la moindre hypothèse sur la distribution des différents termes de la somme.

Signalons enfin que si les v.a.  $X_1, X_2, \dots, X_N$  sont indépendantes et distribuées normalement avec

$$E(X_k) = \mu_k \quad \text{var}(X_k) = \sigma_k^2 \quad k = 1, \dots, N$$

alors la v.a.

$$a_1 X_1 + \dots + a_N X_N = \sum_{k=1}^N a_k X_k,$$

suit une loi normale de paramètres

$$\left( \sum_{k=1}^N a_k \mu_k ; \sqrt{\sum_{k=1}^N a_k^2 \sigma_k^2} \right).$$

### ***Distribution de Poisson***

La v.a. *de Poisson* de paramètre  $\mu$  est une v.a.  $X$  discrète qui prend toutes les valeurs entières positives 0, 1, 2, ... avec, comme probabilités associées

$$\Pr[X = k] \approx e^{-\mu} \frac{\mu^k}{k!} \quad k \in \mathbb{N}.$$

Les valeurs typiques de cette distribution sont données par

$$\begin{aligned} E(X) &= \mu \\ \text{var}(X) &= \mu \end{aligned}$$

### ***Distribution binomiale***

Le *schéma de Bernoulli* est un modèle probabiliste s'appliquant à des situations très diverses. Il est caractérisé par

- un nombre fini d'essais indépendants ;

- lors de chaque essai, deux résultats seulement sont possibles : succès et échec ;
- lors de chaque essai, la probabilité d'occurrence d'un succès est la même.

Si on appelle  $n$  le nombre d'essais et  $p$  la probabilité de succès lors de chaque essai, on parle de schéma de Bernoulli de paramètres  $(n; p)$  et le nombre de succès sur les  $n$  essais est une v.a. *binomiale* de paramètres  $(n; p)$ , que l'on note  $B(n; p)$ . Cette variable aléatoire discrète prend les valeurs  $0, 1, 2, \dots, n$  avec, comme probabilités associées<sup>6</sup>

$$\Pr[B(n; p) = k] = \binom{n}{k} p^k (1-p)^{n-k} \quad k \in \{0, 1, \dots, n\}.$$

La somme de ces probabilités est bien égale à 1 en vertu de la formule du binôme de Newton.

Par ailleurs, les valeurs typiques de cette distribution sont données par

$$\begin{aligned} E(B(n; p)) &= np \\ \text{var}(B(n; p)) &= np(1-p) \end{aligned}$$

La distribution binomiale admet deux approximations intéressantes lorsque le paramètre  $n$  est « grand ». Ainsi, pour  $p$  très petit, on a l'approximation par une loi de Poisson de paramètre  $np$  :

$$\Pr[B(n; p) = k] \approx e^{-np} \frac{(np)^k}{k!}$$

et pour  $p$  pas trop proche de 0 ni de 1, la v.a. binomiale tend vers une loi normale de paramètres  $(np; \sqrt{np(1-p)})$  et, de manière plus précise,

$$\Pr[B(n; p) = k] \approx \Phi\left(\frac{k - \mu + \frac{1}{2}}{\sigma}\right) - \Phi\left(\frac{k - \mu - \frac{1}{2}}{\sigma}\right).$$

## 1.2.2. Aléas multivariés

### *Notion d'aléa multivarié*

On est souvent amené à considérer simultanément plusieurs v.a.  $X_1, X_2, \dots, X_N$  associées à un même phénomène fortuit<sup>7</sup>. Le  $N$ -uplet<sup>8</sup>

$$(X_1, X_2, \dots, X_N),$$

---

<sup>6</sup> Rappelons que  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ .

<sup>7</sup> Le return de différents actifs financiers par exemple.

<sup>8</sup> Qui peut être considéré comme un vecteur de dimension  $N$ .

porte le nom d'*aléa N-varié*.

Nous présenterons simplement ici

- les valeurs typiques associées à un *aléa bivarié*, c'est-à-dire un couple  $(X, Y)$  de v.a. ;
- la distribution multinormale.

### ***Covariance et coefficient de corrélation***

La *covariance* entre deux v.a.  $X$  et  $Y$  se définit par

$$\sigma_{XY} = \text{cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

et elle peut aussi se calculer par

$$\text{cov}(X, Y) = E(XY) - \mu_X \mu_Y.$$

Pour des v.a. discrètes et pour des v.a. à densité, on calcule la covariance par

$$\begin{aligned}\text{cov}(X, Y) &= \sum_i \sum_j (x_i - \mu_X)(y_j - \mu_Y) p_{ij} = \sum_i \sum_j x_i y_j p_{ij} - \mu_X \mu_Y \\ \text{cov}(X, Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf(x, y) dx dy - \mu_X \mu_Y\end{aligned}$$

L'interprétation de la covariance est la suivante : il s'agit d'une mesure du degré de liaison linéaire existant entre les deux v.a. Ainsi, une covariance positive correspond à des valeurs du produit

$$(X - \mu_X)(Y - \mu_Y),$$

majoritairement positives et les deux facteurs sont majoritairement de même signe : à des « grandes » valeurs de  $X$  (supérieures à  $\mu_X$ ) correspondent des « grandes » valeurs de  $Y$  (supérieures à  $\mu_Y$ ) et à des petites valeurs de  $X$  correspondent des petites valeurs de  $Y$ . Le même genre de raisonnement est valable pour une covariance négative.

On peut démontrer que

$$\begin{aligned}\text{cov}(aX + bY + c, Z) &= a \text{cov}(X, Z) + b \text{cov}(Y, Z) \\ \text{cov}(X, X) &= \text{var}(X) \\ E(XY) &= E(X) \cdot E(Y) + \text{cov}(X, Y) \\ \text{var}(X + Y) &= \text{var}(X) + \text{var}(Y) + 2 \text{cov}(X, Y)\end{aligned}$$

et que si  $X$  et  $Y$  sont indépendants, leur covariance est nulle. En effet, dans ce cas,



$$\begin{aligned}
\text{cov}(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\
&= E(X - \mu_X) \cdot E(Y - \mu_Y) \\
&= (E(X) - \mu_X) \cdot (E(Y) - \mu_Y) \\
&= 0
\end{aligned}$$

Un autre paramètre mesurant le degré de liaison linéaire entre les deux v.a est le *coefficient de corrélation* :

$$\rho_{XY} = \text{corr}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y}.$$

Les propriétés se déduisent de celles de la covariance. Par exemple,

$$\text{corr}(aX + b, cY + d) = \pm \text{corr}(X, Y),$$

suivant que  $ac$  est positif ou négatif. De plus,  $\rho_{XX} = 1$ .

L'intérêt du coefficient de corrélation par rapport à la covariance est qu'il s'agit d'un nombre sans dimension alors que l'unité de mesure de la covariance est égale au produit des unités des deux v.a. De plus, le coefficient de corrélation ne peut prendre que des valeurs comprises entre -1 et 1 et ces deux valeurs extrêmes correspondent à l'existence d'une relation linéaire (croissante ou décroissante selon que  $\rho = 1$  ou  $\rho = -1$ ) parfaite entre les deux v.a.

Deux v.a. dont le coefficient de corrélation (ou la covariance) est nul sont qualifiées de *non corrélées*. On a dit plus haut que des v.a. indépendantes sont non corrélées. La réciproque est fausse ! L'indépendance de deux v.a. exclut en effet l'existence de toute relation entre les variables alors que la non corrélation exclut simplement l'existence d'une relation linéaire.

### **Matrice des variances-covariances**

Pour un aléa  $N$ -varié  $(X_1, X_2, \dots, X_N)$ , on peut calculer les covariances deux à deux et ainsi construire la *matrice des variances-covariances* :

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{pmatrix}.$$

La formule  $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y)$  se généralise aisément en

$$\begin{aligned}
\text{var}(a_1 X_1 + a_2 X_2) &= a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + 2a_1 a_2 \sigma_{12} \\
&= \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}
\end{aligned}$$

ce qui montre que l'on peut exprimer la variance d'une combinaison linéaire de v.a. comme forme quadratique associée à la matrice des variances-covariances, calculée pour le vecteurs des coefficients de cette combinaison :

$$\text{var}\left(\sum_{i=1}^N a_i X_i\right) = \sum_{i=1}^N \sum_{j=1}^N a_i a_j \sigma_{ij} = a^t V a$$

La matrice des variances-covariances est une matrice s.p. ou d.p. puisque la forme quadratique associée est une variance. On peut même dire qu'elle est d.p. sauf lorsqu'il existe des coefficients  $a_1, a_2, \dots, a_N$  non tous nuls tels que la v.a.

$$a_1 X_1 + \dots + a_N X_N = \sum_{i=1}^N a_i X_i,$$

est dégénérée, auquel cas  $V$  est s.p. Cette dégénérescence peut survenir par exemple lorsque

- une des variables est dégénérée ;
- certaines variables sont parfaitement corrélées ;
- la matrice  $V$  est obtenue à partir d'observations dont le nombre est strictement inférieur au nombre de variables.

### ***Distribution multinormale***

Un aléa  $N$ -varié  $(X_1, X_2, \dots, X_N)$  est dit distribué suivant une *loi multinormale* de paramètres  $(\mu; V)$  s'il admet une densité multivariée donnée par

$$f(x_1, \dots, x_N) = \frac{1}{\sqrt{(2\pi)^N \text{dtm}(V)}} \exp\left[-(x - \mu)^t V^{-1} (x - \mu)\right],$$

où  $\mu$  et  $V$  représentent respectivement le vecteur des moyennes et la matrice des variances-covariances des v.a.  $X_k$  ( $k = 1, \dots, N$ ).

La propriété de la combinaison linéaire de v.a. normales indépendantes se généralise : pour un aléa multinormal  $(X_1, X_2, \dots, X_N)$  de paramètres  $(\mu; V)$ , la combinaison linéaire  $a_1 X_1 + \dots + a_N X_N$  suit une loi normale de paramètres  $(a^t \mu; a^t V a)$ .

## **1.2.3. Processus stochastiques**

### ***Généralités***

On appelle *processus stochastique* une variable aléatoire fonction de la variable temps :

$$\{X_t : t \in T\}.$$

Si l'ensemble  $T$  des époques est discret, le processus stochastique est simplement une suite de variables aléatoires. Cependant, dans un certain nombre d'applications financières comme le modèle de Black et Scholes, il sera nécessaire de considérer des processus stochastiques à temps continu.

Pour chaque résultat possible  $\omega \in \Omega$ , la fonction de  $X_t(\omega)$  de la variable  $t$  porte le nom de *trajectoire* du processus stochastique.

Un processus stochastique est dit à accroissements indépendants lorsque, quelles que soient les époques  $t_1 < t_2 < \dots < t_n < \dots$ , les v.a.

$$X_{t_1}, X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}}, \dots$$

sont indépendantes. De même, on dira d'un processus stochastique qu'il est à accroissements stationnaires lorsque, pour tout  $t$  et  $h$ , les v.a.

$$X_{t+h} - X_t \quad \text{et} \quad X_h,$$

sont identiquement distribuées.

### ***Processus poissonnien***

On considère un processus de survenances aléatoires d'un événement dans le temps (correspondant à l'ensemble  $[0 ; +\infty[$ ). L'intérêt principal ne correspond pas directement ici aux époques de survenances, mais aux nombres de survenances dans des intervalles donnés. On notera  $n(t_1 ; t_2)$  la v.a. représentant le nombre de survenances dans l'intervalle  $[t_1 ; t_2 [$ .

Un tel processus est un *processus de Poisson* s'il obéit aux hypothèses suivantes :

- les nombres de survenances dans des intervalles de temps disjoints sont indépendants ;
- la distribution du nombre de survenances dans un intervalle de temps ne dépend de ce dernier que par l'intermédiaire de sa durée :  $\Pr[n(t_1 ; t_2) = k]$  est une fonction de  $(t_2 - t_1)$ , qu'on notera désormais  $p_k(t_2 - t_1)$  ;
- il n'y a pas de survenance multiple : si  $h$  est petit,  $\Pr[n(0 ; h) \geq 2] = o(h)$ .
- il existe un *taux de survenance*  $\alpha$  tel que  $\Pr[n(0 ; h) = 1] = \alpha h + o(h)$ .

On peut démontrer que, sous ces hypothèses, la v.a. « nombre de survenances dans un intervalle de durée  $t$  » est distribuée suivant une loi de Poisson de paramètre  $\alpha$  :

$$p_k(t) = e^{-\alpha} \frac{(\alpha t)^k}{k!} \quad k = 0, 1, 2, \dots$$

Notons, pour simplifier,  $X_t = n(0 ; t)$ . Il s'agit d'un processus stochastique qui réalise le comptage des survenances au cours du temps. Une trajectoire d'un tel processus est donc une fonction en escalier, les abscisses des marches correspondant aux époques des survenances et les hauteurs des marches étant égales à 1. On peut démontrer que ce processus est à accroissements indépendants et stationnaires et que  $E(X_t) = \text{var}(X_t) = \alpha t$ .

On peut généraliser ce processus de la manière suivante.

On considère

- un processus de Poisson  $X_t$  tel qu'il vient d'être défini ; en notant  $T_k$  l'époque de la  $k$ -ème survenance, on a :  $X_t = \#\{k : T_k \leq t\}$  ;
- une suite  $Y_1, Y_2, \dots$  de v.a. indépendantes et identiquement distribuées, et indépendantes du processus de Poisson  $X_t$ .

Le processus

$$Z_t = \sum_{\{k: T_k \leq t\}} Y_k,$$

est appelé *processus de Poisson composé*.

Les trajectoires d'un tel processus sont des fonctions en escalier, les abscisses des marches correspondant aux époques des survenances du processus de Poisson sous-jacent et les hauteurs des marches étant des valeurs réalisées des v.a.  $Y_k$ . De plus, on a

$$\begin{aligned} E(Z_t) &= \alpha t \cdot \mu_Y \\ \text{var}(Z_t) &= \alpha t \cdot (\sigma_Y^2 + \mu_Y^2) \end{aligned}$$

### **Mouvement brownien**

Un *mouvement brownien standard* (on dit aussi *processus de Wiener*) est un processus stochastique  $w_t$  défini sur  $\mathbf{R}^+ = [0 ; +\infty[$  tel que

- $w_t$  est à accroissements indépendants et stationnaires ;
- $w_0 = 0$  ;
- $w_t$  est distribué suivant une loi normale de paramètres  $(0; \sqrt{t})$  :

$$f_{w_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}.$$

Ce processus s'introduit naturellement lorsque l'on désire réaliser un passage à la limite (pour un nombre de pas tendant vers l'infini) d'une promenade aléatoire<sup>9</sup> symétrique. Celle-ci correspond, pour chacune des époques

$$0, \frac{t}{n}, \frac{2t}{n}, \dots, \frac{(n-1)t}{n}, t,$$

à une évolution de  $\pm \Delta X$  avec probabilités  $\frac{1}{2}$ .

Il est possible de montrer que les trajectoires d'un mouvement brownien standard sont continues, mais ne sont différentiables en aucune époque.

---

<sup>9</sup> *Random walk* en anglais.

## Différentielle stochastique

Comme on vient de le voir, la différentielle d'un mouvement brownien standard ne peut pas se définir comme celle d'une fonction certaine. Le calcul stochastique d'Itô permet néanmoins de réaliser l'extension du concept de différentiation aux processus stochastiques. Le processus stochastique

$$X_t : \begin{cases} [0; T] \times \Omega \rightarrow \mathbf{R} \\ (t, \omega) \mapsto X_t(\omega) \end{cases}$$

admet une différentielle stochastique dans l'intervalle  $[0 ; T]$  et on note

$$dX_t = a_t dt + b_t dw_t ,$$

s'il existe deux processus stochastiques  $a_t(\omega)$  et  $b_t(\omega)$  vérifiant certaines conditions d'intégrabilité pour presque tout  $\omega$  et que, lorsque  $t_1, t_2 \in [0 ; T]$  avec  $t_1 < t_2$ , on a

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} a_t dt + \int_{t_1}^{t_2} b_t dw_t ,$$

où le second terme est une intégrale stochastique (par rapport au mouvement brownien standard  $w_t$ ) ; cette dernière est définie comme limite d'une suite d'intégrales de fonctions en escalier. La construction rigoureuse de ces outils mathématique est trop compliquée pour être abordée ici.

Donnons deux propriétés de cette différentielle stochastique.

La première permet de calculer la différentielle d'un produit de deux processus stochastiques : si on connaît la différentielle stochastique des processus  $X_t^{(1)}$  et  $X_t^{(2)}$

$$dX_t^{(i)} = a_t^{(i)} dt + b_t^{(i)} dw_t \quad i = 1, 2 ,$$

alors la différentielle stochastique du produit est donnée par

$$d(X_t^{(1)} X_t^{(2)}) = X_t^{(1)} dX_t^{(2)} + X_t^{(2)} dX_t^{(1)} + b_t^{(1)} b_t^{(2)} dt .$$

La seconde propriété est la célèbre *formule d'Itô*<sup>10</sup>. Elle donne la différentielle d'une fonction de deux variables : un processus stochastique dont on connaît la différentielle stochastique, et le temps. Si le processus  $X_t$  a pour différentielle stochastique

$$dX_t = a_t dt + b_t dw_t ,$$

et si  $f(x, t)$  est une fonction de classe  $C_2$ , alors le processus  $f(X_t, t)$  admet la différentielle stochastique

---

<sup>10</sup> Appelée également *lemme d'Itô*.

$$df(X_t, t) = \left[ f'_t(X_t, t) + f'_x(X_t, t)a_t + \frac{1}{2}f''_{xx}(X_t, t)b_t^2 \right] \cdot dt + f'_x(X_t, t)b_t \cdot dw_t.$$

Les règles classiques de différentiation des fonctions certaines, jointes à la relation donnant la différentielle stochastique de  $X_t$  auraient donné ce résultat sans son troisième terme. En fait, celui-ci est nécessaire pour rendre compte complètement des variations du premier ordre, les accroissements stochastiques du mouvement brownien standard étant de l'ordre de  $\sqrt{t}$ .

## 1.3. STATISTIQUE

### 1.3.1. Echantillonnage et estimation

#### *Principes*

En *statistique inférentielle*, on s'intéresse généralement à une population et à des variables mesurées sur les individus de cette population. Malheureusement, la population toute entière est souvent beaucoup trop volumineuse, voire imparfaitement connue, pour être manipulée directement. On se contente donc dans ce cas d'informations observées sur un sous-ensemble de la population : un *échantillon*. On tente ensuite, à partir de constatations relatives à cet échantillon, d'en déduire (d'en inférer) des conclusions relatives à la population.

L'opération consistant à extraire l'échantillon de la population porte le nom d'*échantillonnage*. C'est ici que la théorie des probabilités intervient, constituant le lien entre la population et l'échantillon. Celui-ci est qualifié de simplement aléatoire lorsque les individus sont extraits indépendamment les uns des autres de la population et ont tous la même probabilité d'être prélevés. En pratique, ce n'est pas nécessairement le cas et les procédures mises en place pour procéder à l'échantillonnage doivent imiter ce hasard le mieux possible.

De manière similaire, une v.a. dont on ne connaît pas la loi de probabilité<sup>11</sup> peut être considéré comme une population. Afin d'avoir des informations sur cette distribution, on observe les valeurs réalisées de cette v.a. dans le temps. Les nombres constituant cet historique d'observations forment en fait un échantillon. La représentativité de cet échantillon consiste en un bon ajustement de la loi de probabilité par l'histogramme observé.

#### *Distribution d'échantillonnage*

Supposons que l'on s'intéresse à un paramètre  $\theta$  de la population. Si l'on extrait un échantillon

<sup>11</sup> Ce sera par exemple le cas du return d'une action.

$$x_1, x_2, \dots, x_n,$$

de cette population, on peut calculer le paramètre  $\theta$  pour cet échantillon :

$$\theta(x_1, x_2, \dots, x_n).$$

L'échantillonnage étant précisément à l'origine de l'aspect fortuit du procédé, pour un autre échantillon

$$x'_1, x'_2, \dots, x'_n,$$

on aurait obtenu une autre valeur du paramètre :

$$\theta(x'_1, x'_2, \dots, x'_n).$$

On construit ainsi une v.a.  $\Theta$ , dont les différentes valeurs possibles sont les résultats du calcul de  $\theta$  pour l'ensemble des échantillons possibles. La loi de probabilité de cette v.a.  $\Theta$  porte le nom de *distribution d'échantillonnage*.

Afin d'illustrer cette notion, considérons la distribution d'échantillonnage de la moyenne d'une population et supposons que la variable considérée a pour moyenne et variance respectivement  $\mu$  et  $\sigma^2$ . A partir des différents échantillons, on peut calculer chaque fois une moyenne :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \bar{x}' = \frac{1}{n} \sum_{i=1}^n x'_i \quad \dots$$

On a ainsi défini une v.a.  $\bar{X}$  pour laquelle il est possible de démontrer que

$$E(\bar{X}) = \mu$$

$$\text{var}(\bar{X}) = \frac{\sigma^2}{n}$$

La première de ces deux relations justifie que l'on choisisse la moyenne d'échantillon comme *estimateur* de la moyenne de population : on dit qu'il s'agit d'un estimateur *sans biais*.

### Remarque

Si on examine de manière similaire la distribution d'échantillonnage de la variance, calculée à partir d'un échantillon par

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

la v.a. qui est associée,  $S^2$ , est telle que

$$E(S^2) = \frac{n-1}{n} \sigma^2 .$$

On n'a donc plus affaire à un estimateur sans biais, mais asymptotiquement sans biais (pour  $n$  tendant vers l'infini). C'est la raison pour laquelle on choisit parfois comme estimateur de la variance l'expression

$$\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 . \clubsuit$$

### ***Estimation***

Si le problème est d'estimer un paramètre  $\theta$  de la population, on doit construire un *estimateur*  $\Theta$  qui est une fonction des valeurs observées par le procédé d'échantillonnage. Il est donc intéressant que cet estimateur  $\Theta$  soit de bonne qualité pour l'évaluation du paramètre  $\theta$ . Ainsi, on exige souvent qu'il s'agisse d'un *estimateur sans biais* :

$$E(\Theta) = \theta .$$

Néanmoins, parmi tous les estimateurs sans biais, on désire que l'estimateur retenu possède d'autres propriétés, et notamment que sa dispersion autour de la valeur centrale  $\theta$  soit la plus petite possible : il faut que sa variance

$$\text{var}(\Theta) = E((\Theta - \theta)^2)$$

soit minimale<sup>12</sup>.

### ***Remarque***

A côté de cette estimation ponctuelle (pour un échantillon, on a une seule estimation), on calcule généralement une précision pour cette estimation par la détermination d'un intervalle  $[\Theta_1 ; \Theta_2]$  centré sur la valeur  $\Theta$  qui contiendra la vraie valeur du paramètre  $\theta$  à estimer avec une probabilité donnée :

$$\Pr[\Theta_1 \leq \theta \leq \Theta_2] = 1 - \alpha ,$$

avec  $\alpha = 0,05$  par exemple. Un tel intervalle s'appelle *intervalle de confiance* pour  $\theta$  et le nombre  $\alpha$  en est le *niveau d'incertitude*.

Cette estimation par intervalle de confiance n'est possible que si l'on connaît la distribution d'échantillonnage de  $\theta$ , par exemple parce que la population obéit à telle ou telle distribution connue ou si on peut lui appliquer certains résultats asymptotiques comme le théorème de tendance normale.♣

---

<sup>12</sup> Par exemple, la moyenne d'échantillon est l'estimateur sans biais de variance minimale pour la moyenne de population.



Dans le cas où l'on étudie une v.a.  $X$  à partir d'observations historiques  $x_t$  ( $t = 1, 2, \dots, T$ ), l'estimateur correspondant est qualifié d'*ergodique*. Les formules d'estimation qui seront utilisées par la suite sont les suivantes :

$$\text{estimateur de } E(X) : \bar{x} = \frac{1}{T} \sum_{t=1}^T x_t$$

$$\text{estimateur de } \text{var}(X) : s_x^2 = \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2 = \frac{1}{T} \sum_{t=1}^T x_t^2 - \bar{x}^2$$

$$\text{estimateur de } \text{cov}(X, Y) : s_{xy} = \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y}) = \frac{1}{T} \sum_{t=1}^T x_t y_t - \bar{x} \cdot \bar{y}$$

$$\text{estimateur de } \text{corr}(X, Y) : r_{xy} = \frac{s_{xy}}{s_x s_y}$$

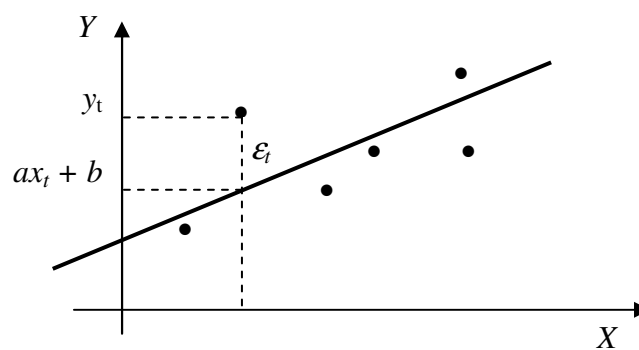
### 1.3.2. Régressions

#### *Régression simple*

Supposons qu'une variable  $Y$  dépende d'une autre variable  $X$  par l'intermédiaire d'une relation linéaire  $Y = aX + b$  et que l'on dispose d'un ensemble d'observations pour ce couple de variables  $(X, Y)$  :

$$(x_t, y_t) \quad t = 1, 2, \dots, T.$$

Si on représente les couples d'observations dans le plan  $(X, Y)$ , on constate qu'il existe des écarts entre ceux-ci et une droite. Ces écarts peuvent provenir, surtout dans le domaine économique, par la non prise en considération de certains facteurs explicatifs de la variable  $Y$ .



*Régression simple*

Il faut donc trouver la droite passant « au plus près » du nuage de points, c'est-à-dire la droite pour laquelle les

$$\varepsilon_t = y_t - (ax_t + b),$$

sont globalement les plus petits possible. Le critère qui est le plus couramment utilisé est de minimiser la somme des carrés de ces écarts (on parle de *méthode des moindres carrés*). Le problème se ramène donc à la recherche des paramètres  $a$  et  $b$  pour lesquels l'expression

$$\sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T [y_t - (ax_t + b)]^2 ,$$

est minimale. Il est facile de montrer que ces paramètres valent

$$\hat{a} = \frac{s_{xy}}{s_x^2} = \frac{\sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^T (x_t - \bar{x})^2}$$

$$\hat{b} = \bar{y} - \hat{a}\bar{x}$$

Ce sont des estimateurs sans biais des vrais paramètres inconnus  $a$  et  $b$  ; de plus, parmi tous les estimateurs sans biais qui s'expriment de manière linéaire en fonction des  $y_t$ , ce sont ceux qui ont la variance minimale<sup>13</sup>.

La droite obtenue par ce procédé porte le nom de *droite de régression*.

### **Régression multiple**

Le modèle de régression que nous venons de présenter peut se généraliser au cas où plusieurs variables explicatives interviennent :

$$Y = \alpha_0 + \alpha_1 X_1 + \dots + \alpha_k X_k .$$

Dans ce cas, si les observations  $x$  et  $y$ , ainsi que les paramètres  $\alpha$ , sont présentés sous forme matricielle :

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{T1} & \cdots & x_{Tk} \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix} \quad \alpha = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} ,$$

on peut montrer que le vecteur des estimations des paramètres est donné par

$$\hat{\alpha} = (X'X)^{-1}(X'Y) .$$

### **Régression non linéaire**

Il arrive que la relation qui permet d'expliquer  $Y$  par  $X_1, X_2, \dots, X_k$  ne soit pas linéaire :

<sup>13</sup> On dit en anglais qu'ils sont BLUE : « best linear unbiased estimators ».

$$Y = f(X_1, \dots, X_k).$$

Dans ce cas, parfois, la relation peut se linéariser par une transformation analytique simple : par exemple,  $Y = aX^b$  se transforme par passage aux logarithmes en

$$\begin{aligned}\ln Y &= \ln a + b \ln X \\ Y^* &= a^* + bX^*\end{aligned}$$

et on est ainsi ramené à un modèle de régression linéaire.

D'autres modèles ne peuvent pas se transformer aussi simplement. Ainsi,  $Y = a + X^b$  n'est pas équivalent à un modèle linéaire. Dans ce cas, des techniques beaucoup plus évoluées, généralement itératives, doivent être utilisées pour estimer les paramètres de ce type de modèle.

## 1.4. TABLEUR

On suppose connus les principes de base et les manipulations élémentaires du tableur Excel™ et on explique quelques particularités relatives aux formules matricielles ainsi qu'à quelques fonctions utiles en calcul matriciel et en statistique.

### 1.4.1. Notion de formule matricielle

Une *formule matricielle* peut effectuer plusieurs calculs et renvoyer des résultats simples ou multiples. Les formules matricielles interviennent sur un ou plusieurs ensembles de valeurs appelés arguments matriciels. On crée des formules matricielles avec la même syntaxe que pour des formules standard simples.

Lorsqu'on tape une formule matricielle, Excel™ insère automatiquement la formule entre accolades ( { } ). L'utilisateur n'a pas à taper ces accolades ; elles ne sont là que pour permettre de reconnaître une formule matricielle d'une formule standard.

Le mode de saisie d'une formule matricielle est le suivant :

- si la formule matricielle renvoie un seul résultat, cliquer sur la cellule dans laquelle on souhaite taper la formule matricielle ; si la formule matricielle renvoie plusieurs résultats, sélectionner la plage de cellules dans laquelle on souhaite taper la formule matricielle ;
- taper la formule matricielle ;
- appuyer sur « CTRL + MAJ + ENTRÉE ».

Pour modifier une formule matricielle,

- cliquer sur une cellule de la plage matricielle ;
- cliquer sur la barre de formule ; lorsque la barre de formule est active, les accolades ( { } ) n'apparaissent pas dans la formule matricielle ;
- modifier la formule matricielle ;
- appuyer sur « CTRL + MAJ + ENTRÉE ».

### 1.4.2. Quelques fonctions matricielles

#### *Déterminant d'une matrice*

DETERMAT(matrice)

« matrice » représente une matrice numérique comportant un nombre égal de lignes et de colonnes. Cet argument peut être donné sous la forme d'une plage de cellules (par exemple A1:C3), d'une constante matricielle (par exemple {1.2.3;4.5.6;7.8.9}) ou d'un nom se référant à l'un ou l'autre de ces types de données.

#### *Inversion matricielle*

INVERSEMAT(matrice)

« matrice » représente une matrice numérique comportant un nombre égal de lignes et de colonnes. Cet argument peut être donné sous la forme d'une plage de cellules (par exemple A1:C3), d'une constante matricielle (par exemple {1.2.3;4.5.6;7.8.9}) ou d'un nom se référant à l'un ou l'autre de ces types de données.

#### *Produit matriciel*

PRODUITMAT(matrice1;matrice2)

« matrice1 », « matrice2 » représentent les matrices dont on souhaite obtenir le produit. Le nombre de colonnes de l'argument matrice1 doit être identique au nombre de lignes de l'argument matrice2, et les deux matrices ne doivent contenir que des nombres.

#### *Transposition matricielle*

TRANSPOSE(tableau)

« tableau » représente une matrice ou une plage de cellules dans une feuille de calcul que l'on veut transposer.

### 1.4.3. Quelques fonctions statistiques

#### ***Coefficient de corrélation***

**COEFFICIENT.CORRELATION**(matrice1;matrice2)

« matrice1 » représente une plage de cellules de valeurs. « matrice2 » représente une seconde plage de cellules de valeurs. Les arguments doivent être des nombres ou des noms, des matrices ou des références contenant des nombres.

#### ***Covariance***

**COVARIANCE**(matrice1;matrice2)

« matrice1 » représente la première plage de cellules de nombres entiers. « matrice2 » représente la seconde plage de cellules de nombres entiers. Les arguments doivent être soit des nombres, soit des noms, matrices ou références contenant des nombres.

#### ***Ecart type d'échantillon***

**ECARTYPE**(nombre1;nombre2;...)

« nombre1,nombre2,... » représentent de 1 à 30 arguments numériques qui correspondent à un échantillon de population. On peut également substituer une matrice unique ou une référence à une matrice à des arguments séparés par des points-virgules. L'écart type est calculé à l'aide de la méthode « sans biais », ou «  $n - 1$  ».

#### ***Ecart type de population***

**ECARTYPEP**(nombre1;nombre2;...)

« nombre1,nombre2,... » représentent de 1 à 30 arguments numériques qui correspondent à une population entière. On peut également substituer une matrice unique ou une référence à une matrice à des arguments séparés par des points-virgules. L'écart type est calculé à l'aide de la méthode «  $n$  ».

#### ***Fréquences***

**FREQUENCE**(tableau\_données;matrice\_intervalles)

« tableau\_données » représente une matrice de valeurs ou une référence à la série de valeurs dont on souhaite calculer les fréquences. « matrice\_intervalles » représente une matrice d'intervalles ou une référence aux intervalles dans lesquels on veut regrouper les valeurs de l'argument « tableau\_données ».

Le nombre d'éléments de la matrice renvoyée est égal au nombre d'éléments de l'argument « matrice\_intervalles » plus 1. Les éléments de la matrice renvoyée donnent les fréquences

d'apparition des valeurs de l'argument « tableau\_données » dans les différentes pages délimitées par les valeurs de l'argument « matrice\_intervalles ».

## **Moyenne**

MOYENNE(nombre1;nombre2;...)

« nombre1, nombre2, ... » représentent les 1 à 30 arguments numériques dont on veut obtenir la moyenne. Les arguments doivent être soit des nombres, soit des noms, des matrices ou des références contenant des nombres.

## **Régression linéaire multiple**

DROITEREG(y\_connus;x\_connus;constante;statistiques)

$$Y = \alpha_0 + \alpha_1 X_1 + \dots \alpha_k X_k$$

« y\_connus » est la série des valeurs de la variable à expliquer. Si la matrice définie par l'argument « y\_connus » occupe une seule colonne, chaque colonne de l'argument « x\_connus » est interprétée comme étant une variable distincte ; si la matrice définie par l'argument « y\_connus » occupe une seule ligne, chaque ligne de la « x\_connus » est interprétée comme étant une variable distincte.

« x\_connus » est une série de valeurs facultatives des variables explicatives. La matrice définie par l'argument « x\_connus » peut contenir une ou plusieurs séries de variables. Si l'argument « x\_connus » est omis, la matrice par défaut est la matrice {1, 2, 3, ...}, de même ordre que l'argument « y\_connus ».

« constante » est une valeur logique qui indique si la constante  $\alpha_0$  doit être égale à 0. Si l'argument « constante » est VRAI ou omis, la constante  $\alpha_0$  est calculée normalement. Si l'argument « constante » est FAUX,  $\alpha_0$  est égal à 0 et les autres coefficients sont ajustés de sorte que  $Y = \alpha_1 X_1 + \dots \alpha_k X_k$ .

« statistiques » représente une valeur logique indiquant si des statistiques de régression supplémentaires doivent être renvoyées. Si l'argument « statistiques » est VRAI, la fonction DROITEREG renvoie des statistiques de régression supplémentaires et la matrice renvoyée devient :

$\alpha_k$	$\alpha_{k-1}$	...	$\alpha_1$	$\alpha_0$
$s_k$	$s_{k-1}$	...	$s_1$	$s_0$
$R^2$	$s_y$			
$F$	$d.l.$			
$SCE_{reg}$	$SCR$			

Si l'argument « statistiques » est FAUX ou omis, la fonction DROITEREG renvoie uniquement les coefficients  $\alpha_j$  ( $j = k, k - 1, \dots, 1, 0$ ) et le tableau ne contient que la première ligne. Si l'argument « constante » est FAUX, il ne contient pas la dernière colonne.

Les statistiques de régression supplémentaires sont les suivantes :

- $s_j$  : écart type de l'estimateur du coefficient  $\alpha_j$  ;
- $R^2$  : coefficient de détermination ; il compare les valeurs de  $Y$  estimées par le modèle aux valeurs réelles et varie entre 0 et 1. Un coefficient de détermination égal à 1 indique une corrélation parfaite de l'échantillon (aucune différence entre les valeurs estimées et réelles) ; à l'inverse, un coefficient de détermination égal à 0 (zéro) indique que l'équation de régression ne peut servir à expliquer  $Y$  ;
- $s_y$  : écart type de la valeur de  $Y$  estimée par le modèle ;
- $F$  : statistique de Fisher-Snedecor ; on utilise ce paramètre pour déterminer si la relation observée entre les variables dépendantes et indépendantes est due au hasard ;
- $d.l.$  : nombre de degrés de liberté ; ils permettent de trouver les valeurs critiques de la statistique  $F$  dans une table statistique pour un niveau d'incertitude donné ;
- $SCE_{reg}$  : somme des carrés des écarts expliqués par la régression ;
- $SCR$  : somme des carrés des résidus.

### ***Somme des carrés***

SOMME.CARRES(nombre1;nombre2;...)

« nombre1, nombre2, ... » représentent de 1 à 30 arguments pour lesquels on veut obtenir la somme des carrés. On peut aussi utiliser une matrice ou une référence à une matrice plutôt que des arguments séparés par des points-virgules.

### ***Somme des produits***

SOMMEPROD(matrice1;matrice2;matrice3;...)

« matrice1, matrice2, matrice3, ... » représentent de 2 à 30 matrices dont on veut multiplier les valeurs pour ensuite additionner leurs produits. Les matrices spécifiées comme arguments doivent avoir la même dimension.

### ***Variance d'échantillon***

VAR(nombre1;nombre2;...)

« nombre1,nombre2,... » représentent de 1 à 30 arguments numériques correspondant à un échantillon de population. On peut également substituer une matrice unique ou une référence à une matrice à des arguments séparés par des points-virgules. La variance est calculée à l'aide de la méthode « sans biais », ou «  $n - 1$  ».

### ***Variance de population***

VAR.P(nombre1;nombre2;...)

« nombre1,nombre2,... » représentent de 1 à 30 arguments numériques qui correspondent à la population entière. On peut également substituer une matrice unique ou une référence à une matrice à des arguments séparés par des points-virgules. La variance est calculé à l'aide de la méthode «  $n$  ».



### *Distribution normale réduite*

$$\Phi(t) = \Pr[Y \leq t] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx$$

<i>t</i>	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998

## CHAPITRE 2

### INTRODUCTION

#### Deux éléments fondamentaux

L'évaluation des actifs financiers doit prendre en compte deux aspects fondamentaux : le hasard et le temps.

#### *Aspect aléatoire*

Il est clair que l'évolution de la valeur d'un actif financier ne peut se prévoir de manière déterministe grâce aux seules observations du passé. C'est évident pour une action, dont le cours fluctue en fonction de la loi de l'offre et de la demande, celles-ci étant elles-mêmes dictées par l'idée que les acteurs du marché se font de la valeur de l'entreprise. Il en va de même pour des produits parfois qualifiés de « sans risque », comme les obligations ; que l'on songe par exemple au risque de défaillance, au risque de change éventuel et au risque dû aux fluctuations des taux d'intérêt.

C'est pourquoi l'évaluation des actifs financiers ne peut se réaliser que dans un contexte aléatoire et les modèles que nous allons construire ne peuvent se passer de l'outil probabiliste (voir l'annexe 2 pour les rappels essentiels).

#### *Aspect temporel*

Certains modèles d'évaluation d'actifs financiers, appelés *monopériodiques*, comme la théorie du portefeuille de Markowitz, examinent la situation « photographiée » à un instant et utilisent les observations historiques pour analyser celle-ci.

On peut au contraire vouloir tenir compte de l'évolution dans le temps, avec des décisions possibles à tout instant en fonction de l'information disponible à ce moment-là. Les variables aléatoires dont il était question au paragraphe précédent deviennent alors des processus stochastiques et les théories associées sont significativement plus complexes.

C'est pourquoi, dans ce qui suit (chapitres 3, 4 et 5), on trouvera à la fois des *modèles d'évaluation* (point de vue *statique*) et des *modèles d'évolution* (point de vue *dynamique*). D'ailleurs, pour la seule évaluation des options, les modèles d'évolution du sous-jacent sont indispensables en raison du lien intrinsèque entre ce produit et la variable temps.

Parmi les modèles dynamiques, on peut distinguer les *modèles discrets* (l'évolution est observée à des époques espacées dans le temps) et les *modèles continus* (la variable temps prend ses valeurs dans un ensemble continu, comme un intervalle), les outils mathématiques utilisés pour ces derniers étant nettement plus évolués.

## Deux principes de base

Les modèles d'évaluation (ou d'évolution) reposent, comme tous les modèles, sur un certain nombre d'hypothèses.

Certaines d'entre elles sont purement techniques et ont pour objectif de garantir le sens des expressions mathématiques qui les représentent ; elles varient fortement suivant le modèle utilisé (statique ou dynamique, discret ou continu) et peuvent prendre la forme de conditions d'intégrabilité, de restriction sur les lois de probabilité et les processus stochastiques...

D'autres hypothèses sont dictées par la réalité économique et par le comportement des investisseurs<sup>14</sup> ; c'est à deux « principes » économiques généralement admis en modélisation financière que nous allons faire allusion ici.

### *Marché parfait*

On formulera souvent une hypothèse simplificatrice au point d'être irréaliste : l'hypothèse d'un marché parfait. Malgré son caractère réducteur, elle définit un contexte dans lequel il est possible de modéliser les actifs financiers et de nombreuses études s'attachent à permettre d'affaiblir les différents éléments de cette hypothèse.

Le *marché parfait*<sup>15</sup> est un marché régi par la loi de l'offre et de la demande sur lequel

- une égale information est disponible pour l'ensemble de tous les investisseurs ;
- il n'y a ni frais de transaction, ni frais d'émission associés aux actifs financiers ;
- il n'y a aucun prélèvement fiscal sur le rendement des actifs financiers (qu'il s'agisse de plus-value ou de dividende par exemple) ;
- les ventes à découvert sont autorisées sans restriction.

### *Absence d'opportunité d'arbitrage*

Une *opportunité d'arbitrage* est un portefeuille défini dans un contexte où

- aucun flux financier n'est intervenu en son sein pendant la période considérée ;
- le taux d'intérêt sans risque ne varie pas pendant cette période et est valable pour n'importe quelle échéance (courbe des taux plate et constante).

Il s'agit d'un portefeuille dont la valeur initiale (à la constitution) est négative et présente, à une époque ultérieure, une valeur positive certaine. De manière plus précise, si on note  $V_t$  la valeur du portefeuille à l'instant  $t$ , il s'agit d'un portefeuille pour lequel on a

$$V_0 < 0 \text{ et } V_T \geq 0 \quad \text{ou} \quad V_0 \leq 0 \text{ et } V_T > 0.$$

---

<sup>14</sup> Nous évoquerons ce dernier aspect au parag. 3.2.6.

<sup>15</sup> Voir par exemple MILLER M.H. et MODIGLIANI F., Dividend policy, growth and the valuation of shares, *Journal of Business*, 1961.

On formule généralement, en modélisation financière, l'hypothèse d'*absence d'opportunité d'arbitrage*. En effet, s'il était possible de construire de tels portefeuilles, on aurait tout intérêt à en construire un grand nombre. Mais les nombreuses opérations (achat/vente) sur le marché que nécessiterait ce comportement conduiraient par le jeu de l'offre et de la demande à la modification des prix des différentes composantes du portefeuille jusqu'à en annuler les bénéfices obtenus par cette position d'arbitrage.

Sous cette hypothèse, on peut donc dire que, pour un portefeuille de valeur  $V$  constitué à l'instant 0, si  $V_T = 0$ , qu'aucun flux financier n'est intervenu en son sein entre 0 et  $T$ , et que le taux d'intérêt sans risque ne varie pas pendant cette période et est valable pour n'importe quelle échéance (courbe des taux plate et constante), alors  $V_t = 0$  pour tout  $t \in [0 ; T]$ .

Cette hypothèse d'absence d'opportunité d'arbitrage peut s'exprimer de la manière équivalente suivante : dans le contexte décrit ci-dessus, un portefeuille constitué de manière à ne pas contenir de composante aléatoire présente nécessairement un rendement égal au taux d'intérêt sans risque.

## Notion de modèle d'évaluation

Un *modèle d'évaluation* pour un actif financier est une relation qui exprime, de manière tout à fait générale, le prix  $p$  (ou le rendement) de l'actif en fonction de différentes variables explicatives<sup>16</sup>  $X_1, X_2, \dots, X_n$  représentant les (des) éléments de marché influençant le prix :

$$p = f(X_1, X_2, \dots, X_n) + \varepsilon.$$

où le résidu  $\varepsilon$  correspond à l'écart existant entre la réalité (le prix effectif  $p$ ) et le modèle d'évaluation (la fonction  $f$ ).

Dans le cas où le modèle d'évaluation du prix est un *modèle linéaire* (comme pour les actions), les facteurs de risque se combinent pour donner, par l'intermédiaire du théorème de tendance normale, une distribution de la variable  $p$  qui est normale (au moins en première approximation), et par conséquent définie par les deux seuls paramètres espérance-variance.

En revanche, pour certains types d'actifs comme les options, le modèle d'évaluation n'est plus linéaire. Le raisonnement précédent ne convient donc plus et ses conclusions ne sont plus valides.

Signalons qu'à côté des facteurs de risque que nous venons de citer, on peut faire entrer parmi les éléments explicatifs du risque de marché :

- l'imperfection des modèles d'évaluation ;
- la connaissance imparfaite des règles et limites propres à l'institution ;
- l'impossibilité d'anticiper les modifications de réglementations légales.

<sup>16</sup> Dans les circonstances où c'est essentiellement le risque du titre qui fait l'objet de l'étude, ces variables explicatives portent le nom de *facteurs de risque*.

Remarquons par ailleurs qu'à côté de ce risque de marché, l'investisseur est confronté à d'autres types de risques qui correspondent à la survenance d'événements exceptionnels tels que guerres, crises pétrolières, ... Cette famille de risques ne peut bien évidemment pas être évaluée par des techniques conçues pour le risque de marché. La technique présentée dans cet ouvrage ne traitera donc pas de ces risques « événementiels ». Il n'empêche que le risk manager prudent non seulement analysera ses positions à l'aide des méthodes conçues pour les risques de marché, mais utilisera également des « *scénarios catastrophes* » pour tenter d'appréhender ce type de risques exceptionnels.

Nous allons, dans ce paragraphe, rappeler quelques principes généraux relatifs aux modèles d'évaluation et citer l'un ou l'autre modèle spécifique<sup>17</sup> qui sera analysé plus en détail dans cette deuxième partie.

### ***Modèles linéaires***

Nous considérons tout d'abord le cas simple où la fonction  $f$  du modèle d'évaluation est linéaire, ou, de manière plus précise, le cas où la variation de prix

$$\Delta p = p_t - p_0,$$

est une fonction du premier degré des variations  $\Delta X_1, \dots, \Delta X_n$  des différentes variables explicatives et de celle  $\Delta \varepsilon$  du résidu :

$$\Delta p = a_0 + a_1 \Delta X_1 + \dots + a_n \Delta X_n + \Delta \varepsilon.$$

Un exemple de modèle d'évaluation linéaire est le *modèle à indice simple* de Sharpe pour les actions (voir § 3.2.4). Ce modèle postule que la variation<sup>18</sup> de prix d'une action est une fonction du premier degré de la variation d'un indice général du marché (bien entendu, les coefficients de cette fonction du premier degré varient d'un titre à l'autre) :

$$\Delta p = \alpha + \beta \Delta I + \Delta \varepsilon.$$

En pratique, les coefficients  $\alpha$  et  $\beta$  s'évaluent par une technique de régression<sup>19</sup>.

### ***Modèles non linéaires indépendants du temps***

Un cas plus complexe est celui où la fonction  $f$  de la relation

$$p = f(X_1, X_2, \dots, X_n) + \varepsilon,$$

<sup>17</sup> BREALEY R.A. et MYERS S.C., *Principles of corporate finance*, McGraw-Hill, 1991.

BROQUET C. et VAN DEN BERG A., *Gestion de portefeuille*, De Boeck, 1990.

COPELAND T.E. et WESTON J.F., *Financial theory and corporate policy*, Addison-Wesley, 1988.

DEVOLDER P., *Finance stochastique*, Editions de l'ULB, 1993.

ROGER P., *L'évaluation des actifs financiers*, De Boeck, 1996.

<sup>18</sup> Il s'agit ici d'une variation relative du prix, c'est-à-dire d'un return. Il en va de même de l'indice.

<sup>19</sup> On trouvera en annexe 3 les éléments de base de statistique nécessaires à la compréhension de cette notion.

n'est pas linéaire.

Lorsque le temps n'est pas pris en considération, on évalue  $\Delta p$  à l'aide d'un développement de Taylor :

$$\Delta p = \sum_{k=1}^n f'_{X_k}(X_1, \dots, X_n) \Delta X_k + \frac{1}{2!} \sum_{k=1}^n \sum_{l=1}^n f''_{X_k X_l}(X_1, \dots, X_n) \Delta X_k \Delta X_l + \dots + \Delta \varepsilon.$$

Dès l'instant où les variations  $\Delta X_k$  des variables explicatives sont faibles, on peut négliger les termes du second ordre et au-delà, ce qui permet d'écrire

$$\Delta p \approx \sum_{k=1}^n f'_{X_k}(X_1, \dots, X_n) \Delta X_k + \Delta \varepsilon,$$

ce qui nous ramène à un modèle linéaire, qui sera alors traité comme au paragraphe précédent.

Par exemple, pour les obligations, lorsque l'on exprime le prix du titre en fonction du taux d'intérêt, on a affaire à un modèle non linéaire. Si on se contente de l'approximation à l'aide du seul paramètre de *duration* (voir § 4.2.2), on utilise une approximation linéaire. Si en outre on fait intervenir la *convexité* (voir § 4.2.3), le développement de Taylor utilisé prend en compte le terme du second degré.

### ***Modèles linéaires dépendant du temps***

Pour certains types d'actifs, une durée joue un rôle fondamental et le temps est un des arguments de la fonction  $f$ .

C'est par exemple le cas des actifs conditionnels, où la durée de vie du contrat est un élément essentiel. Dans ce cas, il est nécessaire de construire des modèles spécifiques qui tiennent compte de cette composante supplémentaire. On n'aura donc plus simplement un *modèle aléatoire stationnaire* (comme le modèle de Sharpe par exemple), mais un modèle qui combine les aspects aléatoire et temporel, c'est-à-dire un *processus stochastique*.

Un exemple de tel modèle est le *modèle de Black et Sholes* pour les options sur action (voir § 5.3.2), où le prix  $p$  est fonction de différentes variables (prix du sous-jacent, prix d'exercice, maturité, volatilité du sous-jacent, taux d'intérêt sans risque). Dans ce modèle, le prix du sous-jacent est lui-même modélisé par l'intermédiaire d'un processus stochastique (mouvement brownien standard).

## CHAPITRE 3

### ACTIONS

#### 3.1. FONDEMENTS

L'*action* est un actif financier correspondant à une part de propriété d'une société, sa valeur étant indicative de la santé de cette société. Elle peut donner lieu à vente et achat de gré à gré ou sur un marché organisé. C'est la loi de l'offre et de la demande sur ce marché qui en détermine le prix (cotation). L'action peut également donner droit au versement périodique de *dividendes*.

##### 3.1.1. Rendement et risque

###### *Return d'une action*

Considérons une action pendant un intervalle de temps  $]t - 1 ; t]$ , dont la durée peut par exemple être un jour, une semaine, un mois, un an. La valeur (le cours) de cette action à la fin de cette période et le dividende versé au cours de celle-ci sont des variables aléatoires<sup>20</sup> qui seront respectivement notées  $C_t$  et  $D_t$ .

Le *return* de ce titre pendant l'intervalle en question est défini par

$$R_t = \frac{C_t - C_{t-1} + D_t}{C_{t-1}}.$$

Il s'agit donc d'une grandeur sans dimension. Celle-ci peut aisément se décomposer en la somme de deux termes :

$$R_t = \frac{C_t - C_{t-1}}{C_{t-1}} + \frac{D_t}{C_{t-1}},$$

- le premier étant la *plus-value*, qui est fictive dans la mesure où le détenteur de l'action n'en profitera qu'en cas de vente à l'instant  $t$  ;
- le second étant le *taux de rendement*, qui est réel puisqu'il représente un revenu.

---

<sup>20</sup> On trouvera en annexe 2 les éléments de base de la théorie des probabilités nécessaires à la compréhension de ces notions.

Si on désire, dans la définition de ce paramètre return, tenir compte du taux d'inflation, on introduit le return nominal  $R_t^{(n)}$  (sans inflation), le return réel  $R_t^{(r)}$  (avec inflation) et le taux d'inflation  $\tau$ , qui sont liés par la relation

$$1 + R_t^{(n)} = (1 + R_t^{(r)}) \cdot (1 + \tau).$$

Le return réel s'obtient alors aisément :

$$R_t^{(r)} = \frac{1 + R_t^{(n)}}{1 + \tau} - 1.$$

### *Illustration*

Une action est cotée 1000 fin mai et 1050 fin juin ; elle a versé un dividende de 80 le 12 juin. Son return (mensuel) sur cette période est de

$$R_{juin} = \frac{1050 - 1000 + 80}{1000} = 0,13 = 13 \% .$$

Qui se compose d'une plus-value de 5 % et d'un rendement de 8 %.

Il s'agit là du return nominal. Si le taux annuel d'inflation cette année-là est de 5 %, le return réel vaut

$$R_{juin}^{(r)} = \frac{1,13}{(1,05)^{1/12}} - 1 = 0,1254 = 12,54 \% . \clubsuit$$

Dans le cas de certaines opérations particulières pendant la période de calcul du return, comme le fractionnement ou le regroupement de titres, une attribution gratuite ou une augmentation de capital, le principe de la définition du return est conservé, mais on prendra soin de n'introduire dans la formule que des grandeurs comparables. Ainsi, lorsqu'une action est fractionnée en  $X$  nouvelles, le return se déterminera par :

$$R_t = \frac{X \cdot C_t - C_{t-1} + D_t}{C_{t-1}} \quad \text{ou} \quad \frac{X \cdot C_t - C_{t-1} + X \cdot D_t}{C_{t-1}}$$

selon que les dividendes sont versés avant ou après la date du fractionnement.

Si un return est estimé à partir de plusieurs returns relatifs à la même durée mais à des époques différentes (return mensuel « moyen » estimé à partir des 12 returns mensuels de cette année-là par exemple), le bon sens mathématique voudrait que l'on tienne le raisonnement suivant :

$$1 + R_{1an} = (1 + R_1) \cdot (1 + R_2) \cdot \dots \cdot (1 + R_{12})$$



et par conséquent

$$R_{1\text{mois}} = \sqrt[12]{(1 + R_1) \cdot \dots \cdot (1 + R_{12})} - 1.$$

L'expression  $(1 + R_{1\text{mois}})$  est donc la *moyenne géométrique* des expression correspondantes pour les différents mois. On rencontre cependant – et on utilise généralement en pratique – la *moyenne arithmétique* :

$$R_{1\text{mois}} = \frac{R_1 + \dots + R_{12}}{12}.$$

Cette dernière relation n'est en fait pas correcte, comme le montre l'exemple d'un titre coté respectivement 1000, 1100 et 1000 aux instants 0, 1 et 2. Le return moyen de ce titre est bien évidemment nul. Les returns sur les deux sous-périodes valent respectivement 10 % et – 9,09 %, ce qui donne pour le return moyen : 0 % par la moyenne géométrique et 0,45 % par la moyenne arithmétique.

D'une manière générale, la moyenne arithmétique surestime toujours le return et ce d'autant plus que les fluctuations des returns partiels sont importantes. On utilise cependant plus volontiers la moyenne arithmétique en raison de sa simplicité<sup>21</sup> et parce que ce type de moyenne est celui qui est généralement utilisé pour les estimateurs en statistique<sup>22</sup>, et qu'il serait difficile de travailler avec des variances et des covariances (voir plus loin) estimées d'une autre manière.

### Remarque

On utilise aussi une autre formule de calcul lorsqu'aucun dividende n'est versé, celle du *return logarithmique* :

$$R_t^* = \ln\left(\frac{C_t}{C_{t-1}}\right).$$

Celle-ci n'est en fait que très légèrement différente de celle introduite précédemment puisqu'on peut la développer par la formule de Taylor en

$$\begin{aligned} R_t^* &= \ln\left(1 + \frac{C_t - C_{t-1}}{C_{t-1}}\right) \\ &= \ln(1 + R_t) \\ &\approx R_t \end{aligned}$$

si on ne prend pas en considération les termes du second degré et au-delà, presque toujours négligeables.

<sup>21</sup> Argument qui n'a plus guère de sens à l'heure de l'informatique.

<sup>22</sup> Voir par exemple le return d'un portefeuille ci-après.

L'intérêt de  $R_t^*$  par rapport à  $R_t$  est que :

- seul  $R_t^*$  peut prendre des valeurs aussi petites que l'on veut : si  $C_{t-1} > 0$ , on a

$$\lim_{C_t \rightarrow 0^+} \ln\left(\frac{C_t}{C_{t-1}}\right) = -\infty,$$

ce qui est compatible avec certaines hypothèses distributionnelles pour le return, alors que

$$\frac{C_t - C_{t-1}}{C_{t-1}} = \frac{C_t}{C_{t-1}} - 1 \geq -1,$$

- $R_t^*$  permet le calcul simple de la variation sur plusieurs périodes consécutives :

$$\ln\left(\frac{C_t}{C_{t-2}}\right) = \ln\left(\frac{C_t}{C_{t-1}} \cdot \frac{C_{t-1}}{C_{t-2}}\right) = \ln\left(\frac{C_t}{C_{t-1}}\right) + \ln\left(\frac{C_{t-1}}{C_{t-2}}\right),$$

ce qui n'est pas possible avec  $R_t$ .

Nous utiliserons cependant  $R_t$  dans la suite des raisonnements. ♣

### Illustration

Calculons, pour quelques valeurs de  $C_t$ , les quantités  $R_t$  et  $R_t^*$  :

$C_t$	$R_t$	$R_t^*$
12750		
12800	0.0039	0.0039
13150	0.0273	0.0271
12150	-0.0760	-0.0794
11100	-0.0864	-0.0907

Tableau 3.1.-  $R_t$  et  $R_t^*$  en fonction de  $C_t$

Les différences constatées sont faibles ; de plus, on a bien

$$\ln\left(\frac{11100}{12750}\right) = 0,0039 + 0,0271 - 0,0794 - 0,0907 = -0,1391. \clubsuit$$

### ***Return d'un portefeuille***

Considérons un portefeuille composé d'un ensemble de  $N$  actions et notons respectivement  $n_j$ ,  $C_{jt}$  et  $R_{jt}$  le nombre d'actions ( $j$ ), le cours de celle-ci à la fin de la  $t$ -ème période et le dividende versé pour une telle action au cours de cette période.

La valeur totale  $V_t$  du portefeuille à l'instant  $t$  et la totalité  $D_t$  des dividendes versés pendant la  $t$ -ème période sont donc donnés par

$$V_t = \sum_{j=1}^N n_j C_{jt}$$
$$D_t = \sum_{j=1}^N n_j D_{jt}$$

Le *return du portefeuille* sera donc donné par

$$\begin{aligned} R_{P,t} &= \frac{V_t - V_{t-1} + D_t}{V_{t-1}} \\ &= \frac{\sum_{j=1}^N n_j C_{jt} - \sum_{j=1}^N n_j C_{j,t-1} + \sum_{j=1}^N n_j D_{jt}}{\sum_{k=1}^N n_k C_{k,t-1}} \\ &= \frac{\sum_{j=1}^N n_j (C_{jt} - C_{j,t-1} + D_{jt})}{\sum_{k=1}^N n_k C_{k,t-1}} \\ &= \sum_{j=1}^N \frac{n_j C_{j,t-1}}{\sum_{k=1}^N n_k C_{k,t-1}} R_{jt} \end{aligned}$$

La quantité

$$X_j = \frac{n_j C_{j,t-1}}{\sum_{k=1}^N n_k C_{k,t-1}},$$

représente la proportion du titre ( $j$ ) investie dans le portefeuille à l'instant  $t - 1$ , exprimée en terme de capitalisation boursière et on a  $\sum X_j = 1$ . Avec cette notation, le return du portefeuille prend la forme suivante :

$$R_{P,t} = \sum_{j=1}^N X_j R_{jt} .$$

### Remarque

Les relations qui précèdent supposent bien entendu que le nombre de chacun des titres en portefeuille ne varie pas pendant la période considérée. Même si cette condition est vérifiée, les proportions  $X_j$ , elles, dépendent de  $t$  par l'intermédiaire des cours. Si on voulait donc considérer un portefeuille qui ait des proportions identiques à deux instants différents, il faudrait modifier les  $n_j$  en conséquence. Ceci n'est guère imaginable en pratique, en raison entre autres des frais de transaction, et nous n'en tiendrons plus compte à l'avenir et les raisonnements seront menés comme si les proportions demeuraient inchangées. ♣

Comme pour une action isolée, lorsque l'on considère un return est estimé à partir de plusieurs returns relatifs à la même durée mais à des époques différentes, on utilise – au lieu de la moyenne géométrique – la moyenne géométrique, ce qui donne

$$\begin{aligned} R_{P,1\text{ mois}} &= \frac{1}{12} \sum_{t=1}^{12} R_{P,t} \\ &= \frac{1}{12} \sum_{t=1}^{12} \sum_{j=1}^N X_j R_{jt} \\ &= \sum_{j=1}^N X_j \left( \frac{1}{12} \sum_{t=1}^{12} R_{jt} \right) \end{aligned}$$

et donc, d'après ce qui a été dit plus haut<sup>23</sup>,

$$R_{P,1\text{ mois}} = \sum_{j=1}^N X_j R_{j,1\text{ mois}} .$$

### Return du marché

D'un point de vue théorique, le marché peut être considéré comme le portefeuille constitué de la totalité des titres en circulation. Le return du marché se définit donc par

$$R_{M,t} = \sum_{j=1}^N X_j R_{jt} ,$$

où  $X_j$  représente le rapport de la capitalisation boursière globale du titre ( $j$ ) et de celle de l'ensemble des titres.

Ces chiffres étant souvent difficiles à traiter, on remplace généralement dans la pratique cette notion par le return d'un *indice boursier* représentatif du marché concerné :

$$R_{I,t} = \frac{I_t - I_{t-1}}{I_{t-1}} .$$

---

<sup>23</sup> Remarquons que cette relation n'aurait pu exister sans l'utilisation de la moyenne arithmétique.

Un indice statistique est un paramètre permettant de décrire l'évolution d'une grandeur  $X$  entre l'époque de base  $t$  et l'époque de calcul  $s$  par

$$I_t(s) = \frac{X(s)}{X(t)}.$$

Dans le cas où  $X$  est composite comme pour la valeur d'un marché boursier, plusieurs modes d'évaluation sont envisageables. Disons simplement que

- certains portent sur les cours, d'autres sur les returns ;
- certains utilisent des moyennes arithmétiques de cours, d'autres des capitalisations boursières ;
- certains tiennent compte des dividendes versés, d'autres pas ;
- certains portent sur l'ensemble des titres cotés, d'autres sont sectoriels ;
- ...

Les indices boursiers les plus célèbres sont le Dow Jones (USA), le S&P 500 (USA), le Nikkei (Japon), l'Eurostoxx 50 (Europe).

### ***Return espéré et estimateur ergodique***

Comme nous l'avons signalé plus haut, le return d'une action est une variable aléatoire, dont la distribution est généralement imparfaitement connue. L'élément essentiel de cette loi de probabilité est bien sûr son espérance<sup>24</sup> : le *return espéré*  $E_j = E(R_j)$ .

Il s'agit là d'une moyenne « ex ante », a priori inaccessible. C'est pourquoi on estime celle-ci sur base des observations historiques disponibles, calculées pour les  $T$  dernières périodes. Un tel estimateur « ex post » portant sur des données historiques est qualifié d'*ergodique*. Celui du return espéré du titre  $(j)$  est donc

$$\bar{R}_j = \frac{1}{T} \sum_{t=1}^T R_{jt}.$$

De la même manière, pour un portefeuille, le return espéré est égal à

$$E_p = E(R_p) = \sum_{j=1}^N X_j E_j = X^t E,$$

en introduisant les vecteurs  $X$  et  $E$  des proportions et des returns espérés des  $N$  titres

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix} \quad E = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_N \end{pmatrix}.$$

<sup>24</sup> A partir d'ici, on n'utilise plus l'indice  $t$  pour la variable aléatoire return relatif à la  $t$ -ème période, mais pour repérer une observation historique (valeur réalisée de cette variable aléatoire).

L'estimateur ergodique associé est alors donné par

$$\bar{R}_P = \frac{1}{T} \sum_{t=1}^T R_{P_t} = \sum_{j=1}^N X_j \bar{R}_j .$$

Dans les développement théorique qui suivent nous utiliserons le formalisme probabiliste (espérance), tout en sachant qu'au niveau des calculs pratiques il faut utiliser le formalisme statistique (estimateur ergodique).

### ***Risque d'une action***

La performance d'une action ne peut se mesurer par son seul return espéré. Il convient également de tenir compte de l'amplitude des fluctuations de ce return autour de sa valeur moyenne, cette amplitude étant une mesure du *risque* associé au titre en question. L'importance des variations d'une variable autour de sa moyenne se mesure par des indices de dispersion ; ceux qui sont retenus ici sont la variance  $\sigma_j^2$  et l'écart type  $\sigma_j$  du return :

$$\sigma_j^2 = \text{var}(R_j) = E[(R_j - E_j)^2] = E(R_j^2) - E_j^2 .$$

Celui-ci est évalué en pratique par son estimateur ergodique

$$s_j^2 = \frac{1}{T} \sum_{t=1}^T (R_{j_t} - \bar{R}_j)^2 = \frac{1}{T} \sum_{t=1}^T R_{j_t}^2 - \bar{R}_j^2 .$$

### ***Remarque***

On connaît à présent deux valeurs typiques du return d'une action : son return (espéré) et son risque. En ce qui concerne la distribution de cette variable aléatoire, s'il est possible d'admettre une *loi normale*, aucun autre paramètre n'est nécessaire, cette loi de probabilité étant caractérisée par sa moyenne et son écart type.

La raison d'être de l'omniprésence de cette distribution est le *théorème de tendance normale* (CLT), qui demande que la variable concernée soit la somme d'un très grand nombre de « petits » effets indépendants.

C'est probablement la raison pour laquelle (nombre de transactions) on constate empiriquement que des returns portant sur de longues périodes (le mois ou l'année) sont souvent distribués normalement, alors qu'il n'en va pas nécessairement de même des returns journaliers par exemple. On constate en effet, dans ce dernier cas, des queues de distribution généralement plus épaisses<sup>25</sup> que celles de la loi normale. Nous aurons l'occasion de revenir sur ce phénomène dans la troisième partie, la Value at Risk s'intéressant précisément à ces queues de distribution.

Néanmoins, dans cette partie, nous considérerons que la distribution du return est caractérisée par le couple « return espéré – risque », ce qui suffit pour la théorie du portefeuille de

<sup>25</sup> On parle alors de distribution *leptocurtique*.

Markowitz<sup>26</sup> ; dans d'autres cas (modèles d'évolution), on supposera en outre qu'elle est normale.♣

D'autres indices de dispersion auraient pu être utilisés comme mesure du risque, comme l'écart moyen  $E(|R_j - E_j|)$  ou la semi-variance qui se définit comme la variance, mais en ne prenant en considération que les valeurs du return inférieures au return espéré. C'est néanmoins la variance (et son équivalent l'écart type) qui est presque toujours utilisée en raison de ses propriétés probabilistes et statistiques, comme on le verra lors de la définition du risque d'un portefeuille.

### ***Covariance et corrélation***

Le risque d'un portefeuille dépendra bien sûr du risque des titres constitutifs, mais également des liens existant entre les différents titres, et ce par le jeu de la diversification. L'évaluation d'une dépendance linéaire entre le return du titre  $(i)$  et celui du titre  $(j)$  se mesure par la covariance :

$$\sigma_{ij} = \text{cov}(R_i, R_j) = E[(R_i - E_i)(R_j - E_j)] = E(R_i R_j) - E_i E_j,$$

qui s'évalue par l'estimateur ergodique

$$s_{ij} = \frac{1}{T} \sum_{t=1}^T (R_{it} - \bar{R}_i)(R_{jt} - \bar{R}_j) = \frac{1}{T} \sum_{t=1}^T (R_{it} R_{jt}) - \bar{R}_i \bar{R}_j.$$

L'interprétation du signe de la covariance est bien connue, mais son ordre de grandeur est difficile à traduire. Pour éviter ce problème, on utilise le coefficient de corrélation

$$\rho_{ij} = \text{corr}(R_i, R_j) = \frac{\sigma_{ij}}{\sigma_i \cdot \sigma_j},$$

dont l'estimateur ergodique est bien sûr donné par

$$r_{ij} = \frac{s_{ij}}{s_i \cdot s_j}.$$

Rappelons que ce dernier paramètre est un nombre pur compris entre  $-1$  et  $1$ , dont le signe indique le sens de la dépendance entre les deux variables, et dont les valeurs proches de  $\pm 1$  correspondent à des relations linéaires presque parfaites entre les variables.

### ***Risque d'un portefeuille***

Si on se rappelle que

---

<sup>26</sup> MARKOWITZ H., Portfolio selection, *Journal of Finance*, Vol. 7, n° 1, 1952, pp. 419-433.

$$R_{P,t} = \sum_{j=1}^N X_j R_{jt}$$

et compte tenu de la formule de la variance d'une combinaison linéaire de variables aléatoires, la variance du return du portefeuille prend la forme

$$\sigma_P^2 = \text{var}(R_P) = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} = X' V X ,$$

où  $\sigma_{ii} = \sigma_i^2$  et où on a posé

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix} \quad V = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{pmatrix} .$$

Si on désire faire apparaître les coefficients de corrélation, la formule ci-dessus devient

$$\sigma_P^2 = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_i \sigma_j \rho_{ij} .$$

### Illustration

Le risque d'un portefeuille composé de deux titres en proportions respectives 30 % et 70 % et tels que

$$\sigma_1^2 = 0,03 \quad \sigma_2^2 = 0,02 \quad \sigma_{12} = 0,01 ,$$

se calcule indifféremment par

$$\sigma_P^2 = 0,3^2 \cdot 0,03 + 0,7^2 \cdot 0,02 + 2 \cdot 0,3 \cdot 0,7 \cdot 0,01 = 0,0167$$

ou par

$$\sigma_P^2 = (0,3 \quad 0,7) \begin{pmatrix} 0,03 & 0,01 \\ 0,01 & 0,02 \end{pmatrix} \begin{pmatrix} 0,3 \\ 0,7 \end{pmatrix} = 0,0167 .$$

Il est intéressant de comparer le risque du portefeuille et le risque des titres individuels. ♣

L'approche « return espéré – risque » au niveau du portefeuille nécessite donc de connaître les returns espérés et les variances individuels ainsi que toutes les covariances deux à deux. Rappelons que la distribution multinormale est caractérisée par ces éléments, mais que la théorie du portefeuille de Markowitz n'exige pas cette loi de probabilité.



### *Risque d'un titre au sein d'un portefeuille*

Le risque d'un portefeuille peut aussi s'écrire

$$\sigma_P^2 = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} = \sum_{i=1}^N X_i \left( \sum_{j=1}^N X_j \sigma_{ij} \right).$$

Le *risque total* du titre ( $i$ ) au sein du portefeuille dépend donc de  $\sigma_i^2$  mais également des covariances avec les autres titres du portefeuille. Il peut se développer de la manière suivante :

$$\begin{aligned} \sum_{j=1}^N X_j \sigma_{ij} &= \sum_{j=1}^N X_j \text{cov}(R_i, R_j) \\ &= \text{cov}\left(R_i, \sum_{j=1}^N X_j R_j\right) \\ &= \text{cov}(R_i, R_P) \\ &= \sigma_{iP} \end{aligned}$$

L'importance relative du risque total du titre ( $i$ ) dans le risque du portefeuille se mesure donc par

$$\frac{\sigma_{iP}}{\sigma_P^2} = \frac{\sum_{j=1}^N X_j \sigma_{ij}}{\sigma_P^2}.$$

Ces *risques relatifs* sont bien entendu tels que

$$\sum_{i=1}^N X_i \frac{\sigma_{iP}}{\sigma_P^2} = 1.$$

### *Illustration*

Avec les données de l'illustration précédente, les risques totaux des deux titres au sein du portefeuille sont donnés par

$$\begin{cases} \sigma_{1P} = 0,3 \cdot 0,03 + 0,7 \cdot 0,01 = 0,016 \\ \sigma_{2P} = 0,3 \cdot 0,01 + 0,7 \cdot 0,02 = 0,017 \end{cases}$$

et les risques relatifs correspondants valent donc respectivement 0,958 et 1,018. Remarquons que l'on a bien

$$0,3 \cdot 0,958 + 0,7 \cdot 1,018 = 1. \clubsuit$$

La notion de risque relatif appliquée au marché dans sa totalité au d'un portefeuille particulier conduit à la notion de *risque systématique* :

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2}.$$

Il représente donc l'importance relative du risque total du titre ( $i$ ) dans le risque du marché, c'est-à-dire la volatilité de  $R_i$  par rapport à  $R_M$  puisque le quotient en question est la pente de la droite de régression où le return du titre ( $i$ ) est expliqué par le return du marché :

$$R_i = \alpha_i + \beta_i R_M.$$

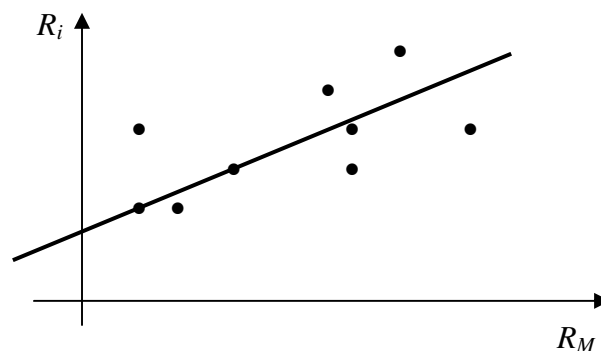


Figure 3.1 – Risque systématique

On peut retenir, en conclusion, que le risque d'un titre ne peut en aucune manière s'envisager indépendamment du portefeuille dans la composition duquel il entre.

### 3.1.2. Efficience des marchés

Nous présentons rapidement ci-après la notion d'efficience des marchés<sup>27</sup>, qui constitue une hypothèse nécessaire (ou qui doit au moins être approximativement vérifiée) pour la validité des différents modèles d'évaluation en finance et qui est fortement liée à la notion de marché parfait.

#### *Principes généraux*

C'est Eugène Fama<sup>28</sup> qui a introduit de manière explicite cette notion d'efficience. La définition qu'il en a donnée est la suivante : « Un marché financier est dit *efficient* si et

<sup>27</sup> On consultera à ce sujet GILLET P. *L'efficience des marchés financiers*, Economica, 1999.

<sup>28</sup> FAMA E.F., The behavior of stock market prices, *Journal of Business*, Vol. 38, 1965, pp. 34-105.

FAMA E.F., Random walks in stock market prices, *Financial Analysis Journal*, 1965.

FAMA E.F., Efficient capital markets : a review of theory and empirical work, *Journal of Finance*, Vol. 25, 1970.

seulement si l'ensemble des informations disponibles concernant chaque actif financier coté sur le marché est immédiatement intégré dans le prix de cet actif ».

Il revient au même de dire qu'il n'y a aucune surévaluation ou sous-évaluation de la valeur des titres, ou encore qu'aucun actif ne peut procurer aucun return supérieur à celui qui correspond à ses caractéristiques propres.

Une telle hypothèse garantit donc l'égalité de traitement des différents investisseurs : aucune catégorie de ceux-ci ne dispose d'avantage informationnel. Les informations disponibles sur un tel marché permettent alors de réaliser des allocations de ressources de manière optimale.

La justification économique de cette notion est que les différents investisseurs – en concurrence et disposant de la même information – vont, par leurs interventions et compte tenu de la loi de l'offre et de la demande, amener le prix d'un titre à coïncider avec sa valeur intrinsèque.

Il s'agit bien entendu là d'une hypothèse qui divise les tenants de l'*analyse fondamentale* et les tenants de l'*analyse technique*. Les premiers l'admettent et en font même la base de leurs raisonnements ; ils supposent que les rendements des titres sont des variables imprévisibles et proposent des techniques de gestion de portefeuille qui consistent à minimiser les risques liés à ces variables<sup>29</sup>. Les seconds proposent des méthodes<sup>30</sup> consistant à prévoir l'évolution des cours à partir des mouvements observés historiquement.

D'un point de vue plus mathématique l'efficacité des marchés consiste à supposer que la suite des cours constitue une *promenade aléatoire*, c'est-à-dire que la suite  $C_t - C_{t-1}$  ( $t = 1, 2, \dots$ ) est faite de v.a. indépendantes et identiquement distribuées. Dans ces circonstances, la prévision d'une telle variation ne peut se réaliser à partir des observations disponibles.

Les conditions économiques qui définissent un marché efficient sont

- les agents économiques intervenant sur ce marché ont un comportement rationnel : ils utilisent l'information disponible de manière cohérente et visent à maximiser l'utilité espérée de leur richesse ;
- l'information est disponible simultanément pour l'ensemble des investisseurs et la réaction de ces derniers à ladite information est instantanée ;
- l'information est disponible gratuitement ;
- il n'y a sur ce marché ni frais de transaction, ni taxes ;
- le marché en question est parfaitement liquide.

Il est bien clair que ces conditions ne peuvent être rigoureusement satisfaites par un marché réel. Se pose alors la question de savoir si les écarts sont à ce point importants qu'ils conduisent à invalider l'hypothèse d'efficacité. C'est à cette question que sont consacrés les paragraphes qui suivent et l'analyse est réalisée à trois niveaux, en fonction de l'accessibilité de l'information. Le moins que l'on puisse dire c'est que les conclusions des recherches réalisées pour tester l'efficacité sont mitigées et ne permettent pas de se faire une idée claire et définitive.

---

<sup>29</sup> C'est l'attitude qui est adoptée dans cet ouvrage.

<sup>30</sup> On consultera par exemple BECHU T., BERTRAND E., *L'analyse technique*, Economica, 1998.

### ***Forme faible de l'efficience***

La *forme faible* de l'hypothèse d'efficience postule qu'il n'est pas possible d'obtenir un avantage particulier de l'ensemble des observations historiques ; les cours intègrent donc purement et simplement les valeurs antérieures des cours.

Les tests réalisés pour vérifier cette hypothèse sont relatifs à la possibilité de prévision des cours à partir des historiques. Voici quelques analyses menées :

- test d'autocorrélation : existe-t-il une corrélation (positive ou négative) entre les valeurs successives du return des titres, permettant de réaliser des prévisions ?
- test des changements de signe : la distribution des longueurs de séquences de returns positifs et de returns négatifs est-elle normale ?
- tests statistiques de promenade aléatoire ;
- tests de simulation des méthodes de l'analyse technique : les techniques de spéculation donnent-elles de meilleurs résultats que la gestion passive ?

D'une manière générale, une partie importante de ces tests conduit à accepter l'hypothèse d'efficience faible, même si les tests les plus exigeants du point de vue statistique l'invalident parfois.

### ***Forme semi-forte de l'efficience***

La *forme semi-forte* de l'hypothèse d'efficience postule qu'il n'est pas possible d'obtenir un avantage particulier d'informations rendues publiques concernant les titres ; les cours s'adaptent donc instantanément et correctement à l'annonce publique d'événement comme augmentation de capital, fractionnement d'un titre, modification de politique de dividende, publication de bilan, OPA, ...

Les tests réalisés pour vérifier cette hypothèse portent donc sur l'effet d'événements annoncés. Ils consistent successivement à

- déterminer le return théorique d'un titre  $R_{it} = \alpha_i + \beta_i R_{Mt}$  sur base d'observations historiques relatives à une période ne comportant pas de tels événements ;
- comparer, lorsque de tels événements surviennent, l'écart entre le return théorique et le return réel ;
- mesurer le temps de réaction pour que les valeurs s'ajustent à nouveau.

Pour la majorité des familles d'événements analysés, l'ajustement se réalise correctement, même si celui-ci n'est pas toujours immédiat après la publication de l'événement.

### ***Forme forte de l'efficience***

La *forme forte* de l'hypothèse d'efficience postule qu'il n'est pas possible d'obtenir un avantage particulier d'informations non publiques concernant les titres ; les cours s'adaptent donc instantanément et correctement à la survenance non rendue publique d'un événement, c'est-à-dire à une information privilégiée.

Les tests réalisés pour vérifier cette hypothèse portent donc sur l'existence d'information privilégiée. Ils suivent des méthodologies similaires à celle utilisée pour la forme semi-forte, mais dans des circonstances particulières :

- dans des cas de délit d'initié avérés ;
- dans des cas où des transactions intensives sur le marché sans qu'une information publique n'existe ;
- dans des cas de transactions intensives de la part d'initiés ;
- dans des cas de portefeuilles gérés par des professionnels susceptibles de disposer d'informations spécifiques avant le grand public, comme pour les organismes de placement collectif.

Les conclusions de ces tests sont plus délicates à interpréter, leur champ d'action étant par définition nettement plus restreint. Cependant, ces tests ne permettent pas de rejeter de manière formelle l'hypothèse d'efficience forte.

### ***Cas d'inefficience systématique observée***

Bien que, d'après les analyses précédentes, l'hypothèse d'efficience peut globalement être acceptée, des cas d'inefficience systématique ont été mis à jour. C'est ainsi qu'on a parfois observé

- des rentabilités supérieures à la moyenne en fin de semaine, en fin de mois ou en fin d'année ;
- des rentabilités pour les entreprises à faible capitalisation boursière supérieures à celles des sociétés à forte capitalisation.

A côté de ces écarts, des *poches d'inefficience* permettant de réaliser des arbitrages peuvent se présenter. Elles peuvent avoir pour origine

- des *bulles spéculatives*, où le cours d'un titre s'écarte significativement et durablement de sa valeur intrinsèque avant de finir par retrouver sa valeur d'origine, et ce sans que des mouvements de variables économiques de marché puissent expliquer ces écarts ;
- le comportement irrationnel de certains investisseurs.

Ces différents éléments, bien que s'écartant de l'hypothèse d'efficience, ne conduisent pas à remettre cette dernière en cause. Par ailleurs, l'investisseur désirant en bénéficier consommera souvent son avantage en frais de transactions.

### ***Conclusion***

Donnons la parole à Ph. Gillet pour conclure cette analyse.

« L'efficience des marchés financiers apparaît à la fois comme une abstraction intellectuelle, un mythe et un objectif.

Une abstraction intellectuelle. Mise en évidence par des chercheurs, la théorie de l'efficience des marchés financiers remet en cause un certain nombre de pratiques couramment utilisées par les professionnels des marchés financiers, comme l'analyse technique. (...) Elle suggère une gestion passive alors que l'analyse technique conduit à une gestion active. (...) Par ailleurs, elle constitue l'un des fondements de la théorie financière moderne. (...).

Un mythe. L'ensemble des hypothèses nécessaires à l'acceptation de la théorie de l'efficience sont acceptées par les tenants de la théorie. Outre les hypothèses classiques de circulation de l'information ou d'absence de coûts de transaction, qui ont été résolues, d'autres hypothèses sous-jacentes restent peu explorées, en particulier celles liées au comportement des investisseurs et à la liquidité. (...).

Un objectif. Les autorités de marché sont conscientes que les caractéristiques de l'efficience assainissent le marché, rendent celui-ci plus crédible et donc attirent les investisseurs et les entreprises. Rendre un marché plus efficient, c'est diminuer l'ensemble des risques systémiques liés à ce marché ; c'est également diminuer les risques de bulles spéculatives (...). L'objectif des autorités est donc d'accroître l'efficience des marchés financiers (...). »

### 3.1.3. Modèles d'évaluation d'une action

Le principe d'équivalence, base de la mathématique financière, permet d'exprimer que la valeur intrinsèque  $V_0$  d'une action à l'instant 0 est égale à la somme des valeurs actualisées des flux financiers futurs que le titre va engendrer. Pour simplifier, si on suppose que les dividendes (flux financiers futurs) sont versés aux époques 1, 2, ... et ont un montant respectif de  $D_1, D_2, \dots$  et si on note  $k$  le taux d'actualisation, on obtient la relation

$$V_0 = \sum_{t=1}^{\infty} D_t (1+k)^{-t}.$$

#### *Remarque 1*

L'utilisation directe de cette relation peut être délicate. En effet,

- on ne connaît généralement pas la valeur de tous les dividendes futurs ;
- cette formule suppose le taux d'actualisation constant (jusqu'à l'infini !) ;
- elle ne permet pas de prendre en considération des opérations particulières comme le fractionnement ou le regroupement de titres, une attribution gratuite ou une augmentation de capital.

Cette formule rend néanmoins de nombreux services et nous présentons ci-après une relation simplifiée qui s'en déduit.♣

#### *Remarque 2*

Cette formule reliant  $V_0$  et  $k$  peut s'utiliser de deux manières :

- si  $V_0$  est connu (valeur intrinsèque sur un marché efficient), on en déduit la valeur de  $k$ , qui représente alors le taux de rendement espéré du titre en question ;
- si c'est  $k$  qui est donné, la formule fournit une évaluation de la valeur du titre, qui peut donc être comparée au cours réel  $C_0$ , ce qui permet de constater la sur- ou sous-évaluation du titre.♣

### Formule de Gordon-Shapiro

Cette relation<sup>31</sup> repose sur les hypothèses suivantes :

- la croissance de la firme est autofinancée ;
- le taux de rentabilité  $r$  des investissements le taux de distribution  $d$  des bénéfices sont constants de période en période.

Sous ces hypothèses, si on note  $B_t$  le bénéfice par action réalisé au cours de la  $t$ -ième période et  $E_t$  la valeur comptable par action à l'instant  $t$  (capital divisé par le nombre d'actions), on a

$$\begin{cases} \frac{D_t}{B_t} = d \\ B_t = r \cdot E_{t-1} \end{cases}$$

et, par conséquent,

$$\begin{aligned} B_{t+1} &= B_t + r \cdot (B_t - D_t) \\ &= B_t [1 + r(1 - d)] \end{aligned}$$

Les bénéfices croissent donc à un taux constant

$$g = r(1 - d),$$

qui est le taux de rentabilité des investissements, diminué de la proportion distribuée. Les dividendes croissent également à ce taux constant et on peut écrire  $D_{t+1} = g \cdot D_t$ , d'où

$$D_t = D_1(1 + g)^{t-1}.$$

La valeur actuelle peut alors se développer de la manière suivante :

$$\begin{aligned} V_0 &= \sum_{t=1}^{\infty} D_1(1 + g)^{t-1}(1 + k)^{-t} \\ &= \frac{D_1}{1 + k} \sum_{t=0}^{\infty} \left( \frac{1 + g}{1 + k} \right)^t \\ &= \frac{\frac{D_1}{1 + k}}{1 - \frac{1 + g}{1 + k}} \end{aligned}$$

à condition que le taux d'actualisation  $k$  soit supérieur au taux de croissance  $g$ , ce conduit à la formule de Gordon-Shapiro :

---

<sup>31</sup> Voir GORDON M., SHAPIRO E., Capital equipment analysis : the required rate profit, *Management Science*, Vol 3, octobre, 1956.

$$V_0 = \frac{D_1}{k - g} = \frac{dB_1}{k - g} = \frac{drE_0}{k - g}$$

### *Illustration*

Le capital d'une entreprise est constitué de 50000 actions, pour un montant total de 10 000 000. Le taux de rentabilité des investissements est de 15 %, le taux de distribution des bénéfices de 40 % et le taux d'actualisation de 12 %.

Le bénéfice par action vaut

$$B = 0,15 \cdot \frac{10\,000\,000}{50\,000} = 30$$

et le dividende par action est donc égal à  $D = 0,4 \cdot 30 = 12$ . Par ailleurs, le taux de croissance est donné par

$$g = 0,15 \cdot (1 - 0,4) = 0,09.$$

La formule de Gordon-Shapiro conduit donc à

$$V_0 = \frac{12}{0,12 - 0,09} = \frac{12}{0,03} = 400.$$

La valeur de marché de cette entreprise est donc de  $50\,000 \times 400 = 20\,000\,000$ , alors que sa valeur comptable est seulement de 10 000 000.♣

La formule de Gordon-Shapiro permet d'écrire

$$k = g + \frac{D_1}{V_0},$$

ce qui montre que le rendement  $k$  peut se décomposer en le taux de croissance des dividendes et le taux de rémunération par titre.

Le « price-earning » ratio

Un des indicateurs d'évaluation les plus répandus est le *PER*. Il est égal au rapport du cours de l'action sur le bénéfice net attendu par action :

$$PER_0 = \frac{C_0}{B_1}.$$



Son interprétation est claire : à l'achat d'un titre, on paie  $PER_0 \times 1$  EUR pour un bénéfice de 1 EUR. Son inverse (bénéfice sur prix) est souvent considéré comme une mesure du rendement du titre et les titres dont le  $PER$  est plus faible que la moyenne du marché sont considérés comme sous-évalués, et donc intéressants.

Cet indicateur peut s'interpréter à l'aide de la formule de Gordon-Shapiro, si les hypothèses relatives à cette dernière sont satisfaites. En effet, en remplaçant le cours par la valeur  $V_0$  donnée par cette formule

$$C_0 = \frac{D_1}{k - g} = \frac{dB_1}{k - r(1 - d)},$$

on obtient directement

$$PER_0 = \frac{d}{k - r(1 - d)}.$$

Ceci permet de déduire une expression pour le taux de rendement  $k$  :

$$\begin{aligned} k &= r(1 - d) + \frac{d}{PER_0} \\ &= r(1 - d) + \frac{1}{PER_0} - \frac{1 - d}{PER_0} \end{aligned}$$

et comme  $PER_0 = C_0/rE_0$ , on trouve

$$k = \frac{1}{PER_0} + \frac{r(1 - d)(C_0 - E_0)}{C_0}.$$

### *Illustration*

Si l'on reprend les mêmes chiffres qu'au paragraphe précédent :

$$r = 15 \%$$

$$d = 40 \%$$

$$E_0 = \frac{10\,000\,000}{50\,000} = 200$$

et que le cours effectivement observé est de 360, on a

$$PER_0 = \frac{360}{30} = 12,$$

ce qui permet de déterminer le taux de rendement<sup>32</sup> :

$$\begin{aligned}k &= \frac{1}{12} + \frac{0,15 \cdot (1 - 0,4) \cdot (360 - 200)}{360} \\&= 0,0833 + 0,04 \quad \clubsuit \\&= 12,33 \%\end{aligned}$$

## 3.2. DIVERSIFICATION ET GESTION DE PORTEFEUILLE

### 3.2.1. Principes de diversification

La constitution d'un portefeuille d'actions optimal consiste à répondre à la double question suivante : étant une liste de  $N$  actions disponibles sur le marché,

- quelles actions choisir parmi celles-ci ?
- en quelles quantités (nombres ou proportions) ?

On pourrait envisager de rechercher le portefeuille dont le return attendu est le plus grand possible. Cette approche conduirait à disposer d'un portefeuille constitué d'une seul titre, celui dont le return espéré est le plus grand. Malheureusement, ce comportement ignore totalement l'aspect risque et peut conduire à des catastrophes en cas de chute du cours du titre retenu.

Les corrélations existant entre les returns des différents titres disponibles peuvent par ailleurs contribuer à compenser les fluctuations des diverses composantes du portefeuille. Cette propriété permettra, à l'opposé de l'attitude précédente, de diminuer le risque du portefeuille sans trop détériorer son return espéré.

C'est ce phénomène que nous analysons ici et que nous exploiterons ensuite pour constituer des portefeuilles optimaux.

#### *Portefeuille à deux actions*

D'après ce qui a été dit précédemment, le return espéré et la variance d'un portefeuille à deux actions représentées en proportions<sup>33</sup>  $X_1$  et  $X_2$  sont donnés par

$$\begin{cases} E_p = X_1 E_1 + X_2 E_2 \\ \sigma_p^2 = X_1^2 \sigma_1^2 + X_2^2 \sigma_2^2 + 2X_1 X_2 \sigma_1 \sigma_2 \rho \end{cases}$$

<sup>32</sup> Bien entendu, si le cours avait été égal à la valeur intrinsèque  $V_0 = 400$ , on trouverait  $k = 12 \%$ .

<sup>33</sup> On suppose implicitement dans ce paragraphe que les proportions sont comprises entre 0 et 1, c'est-à-dire qu'il n'y a pas de ventes à découvert.

Afin de bien mettre en évidence l'effet de la diversification (impact de la corrélation sur le risque), considérons tout d'abord le cas où les deux titres ont le même return espéré ( $E_1 = E_2 = E$ ) et le même risque ( $\sigma_1 = \sigma_2 = \sigma$ ). Puisque  $X_1 + X_2 = 1$ , les équations ci-dessus deviennent

$$\begin{cases} E_p = E \\ \sigma_p^2 = (X_1^2 + X_2^2 + 2X_1X_2\rho)\sigma^2 \end{cases}$$

Le return espéré du portefeuille est égal à celui des titres mais le risque, lui, est inférieur puisque la valeur maximale qu'il peut prendre correspond à  $\rho = 1$  pour laquelle  $\sigma_p = \sigma$  et que lorsque  $\rho < 1$ , on a  $\sigma_p < \sigma$ . Remarquons que dans le cas d'une corrélation parfaite négative ( $\rho = -1$ ), le risque peut s'écrire

$$\sigma_p^2 = (X_1 - X_2)^2 \sigma^2,$$

qui vient même à s'annuler lorsque l'on choisit  $X_1 = X_2 = 1/2$  : dans ce cas, le return espéré est conservé, mais le risque est complètement annulé.

Envisageons maintenant le cas plus général où les returns espérés et les risques sont quelconques. Un titre est caractérisé par un couple  $(E_i, \sigma_i)$  pour  $i = 1$  ou  $2$  et peut donc se représenter par un point dans l'espace  $(E, \sigma)$  ; il en va bien sûr de même pour le portefeuille, correspondant au point  $(E_p, \sigma_p)$ . Suivant les valeurs que l'on donne à  $X_1$  (et donc à  $X_2$ ), le point représentatif du portefeuille va décrire, dans ce plan, une courbe. Étudions rapidement l'allure de cette courbe en fonction des valeurs du coefficient de corrélation  $\rho$ .

Lorsque  $\rho = 1$ , la variance du portefeuille<sup>34</sup> devient

$$\sigma_p^2 = (X_1\sigma_1 + X_2\sigma_2)^2.$$

En éliminant  $X_1$  et  $X_2$  entre les trois équations

$$\begin{cases} E_p = X_1E_1 + X_2E_2 \\ \sigma_p = X_1\sigma_1 + X_2\sigma_2 \\ X_1 + X_2 = 1 \end{cases}$$

on obtient la relation

$$\sigma_p = \frac{E_p - E_2}{E_1 - E_2} \sigma_1 + \frac{E_1 - E_p}{E_1 - E_2} \sigma_2,$$

qui exprime  $\sigma_p$  en fonction de  $E_p$  comme une fonction du premier degré et l'ensemble des portefeuilles est donc le segment de droite reliant les points représentatifs des deux titres (voir figure 3.2).

<sup>34</sup> En toute rigueur, on devrait dire la variance du return du portefeuille.

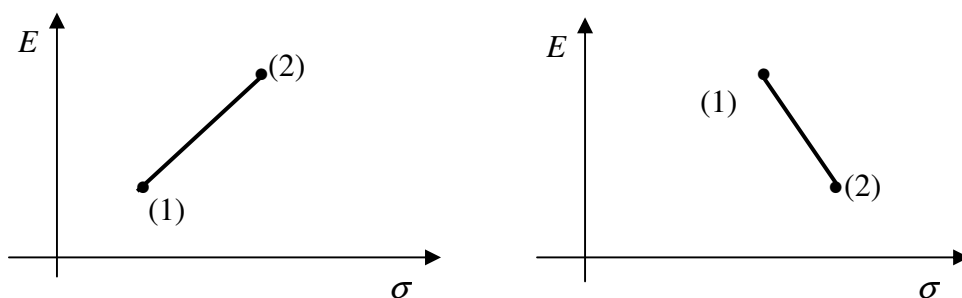


Figure 3.2 – Portefeuille à deux actions (cas  $\rho = 1$ )

Face à la situation représentée à gauche, l'investisseur choisira un portefeuille situé sur le segment en fonction de son attitude vis-à-vis du risque, le portefeuille (1) donnant un return espéré faible mais présentant peu de risque, (2) étant exactement à l'opposé. Face à une situation représentée sur le graphique de droite, le doute n'est par contre plus permis : (2) est meilleur que (1), tant du point de vue du return espéré que du risque couru.

Lorsque  $\rho = -1$ , la variance du portefeuille devient

$$\sigma_p^2 = (X_1\sigma_1 - X_2\sigma_2)^2,$$

c'est-à-dire  $\sigma_p = |X_1\sigma_1 - X_2\sigma_2|$ . Le même raisonnement que ci-avant conduit à la conclusion suivante : les portefeuilles qu'il est possible de construire constituent deux segments de droite issus des points (1) et (2) et concourant en un point de l'axe vertical ( $\sigma = 0$ ), et de pentes égales au signe près (voir figure 3.3).

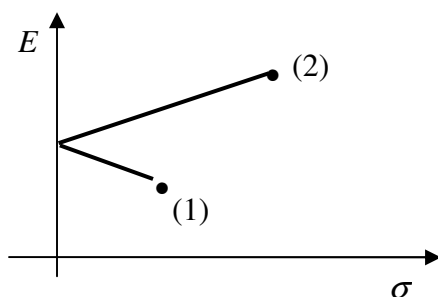


Figure 3.3 – Portefeuille à deux actions (cas  $\rho = -1$ )

Bien entendu, parmi ces portefeuilles, seuls ceux situés sur le segment supérieur sont intéressants : ceux du segment inférieur sont moins bons du double point de vue return espéré et risque.

Dans le cas général,  $-1 < \rho < 1$ , et on peut montrer que l'ensemble des portefeuilles qu'il est possible de construire est un arc de courbe reliant les points (1) et (2), situé entre les graphes des cas extrêmes  $\rho = \pm 1$ , tel que représenté sur la figure 3.4.

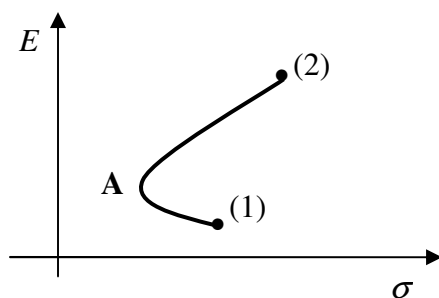


Figure 3.4 – Portefeuille à deux actions (cas général)

Si, comme on l'a fait dans le cas  $\rho = 1$ , on exprime  $\sigma_p^2$  en fonction de  $E_p$ , on obtient une fonction du second degré. Par conséquent, la courbe obtenue dans le plan  $(E, \sigma)$  est une branche d'hyperbole.

On appelle *portefeuille efficient* un portefeuille qui, parmi ceux qu'il est possible de construire avec les deux titres, ne peut être amélioré du double point de vue return espéré et risque.

Graphiquement, il s'agit donc des portefeuilles situés au-dessus du point A de contact<sup>35</sup> de la tangente verticale à la courbe des portefeuilles. En effet, entre A et (2), il n'est pas possible d'améliorer  $E_p$  sans augmenter le risque, ni de diminuer  $\sigma_p$  sans détériorer le return espéré ; par ailleurs, tout portefeuille situé sur l'arc reliant A et (1) est moins bon que ceux qui sont situés à sa gauche.

### Portefeuille à plus de deux actions

Un portefeuille composé de trois actions<sup>36</sup> peut être considéré comme un mélange d'un des titres et d'un portefeuille s-composé des deux autres. Par exemple, le portefeuille dont la composition est  $X_1 = 0,5$ ,  $X_2 = 0,2$  et  $X_3 = 0,3$  peut également être considéré comme composé du titre (1) et d'un portefeuille, lui-même composé des titres (2) et (3), à raison de 40 % et 60 % respectivement. Donc, pour des covariances  $\sigma_{12}$ ,  $\sigma_{13}$ ,  $\sigma_{23}$  fixées, l'ensemble des portefeuilles qu'on peut construire par ce procédé correspond à un ensemble continu de courbes telles que représentées sur la figure 3.5.

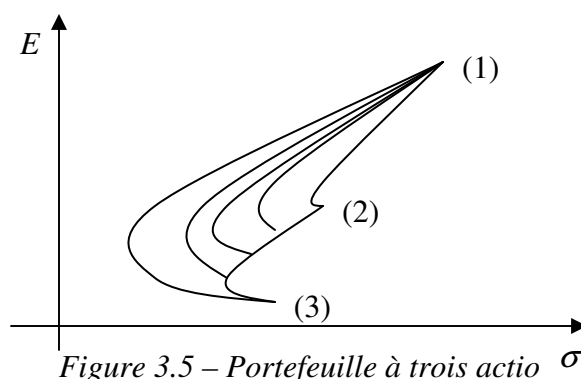


Figure 3.5 – Portefeuille à trois actions

<sup>35</sup> Ce point de contact correspondant au portefeuille de risque minimal.

<sup>36</sup> Le passage de deux à trois actions est général : les résultats obtenus sont valables pour  $N$  titres. Dans le CD-Rom joint, on trouvera des illustrations plus réalistes des différents modèles dans des fichiers Excel™ rassemblés dans le dossier nommé « PfAct ».

L'ensemble des portefeuilles que l'on peut construire avec trois titres (ou davantage) est donc une surface dans le plan  $(E, \sigma)$ .

La notion de portefeuille efficient se définit de la même manière que dans le cas de deux titres. L'ensemble des portefeuilles efficients est donc la partie de la frontière de cette surface, limitée par le titre (1) et le point de contact de la tangente verticale à cette surface, correspondant au portefeuille de risque minimal. Cet arc de courbe porte le nom de *frontière efficiente*.

La fin de ce paragraphe 3.2. est consacrée aux différentes techniques de détermination de cette frontière efficiente, en fonction de diverses contraintes et hypothèses.

Le choix, pour un investisseur, d'un portefeuille sur la frontière efficiente se fera en fonction de son attitude par rapport au risque. S'il adopte l'attitude la plus prudente, il choisira le portefeuille situé à l'extrémité gauche de cette frontière efficiente (portefeuille le moins risqué, très bien diversifié) ; un goût prononcé pour le risque le fera se diriger vers de portefeuilles situés sur la partie droite de la frontière efficiente (acceptation d'un risque élevé pour atteindre un return espéré plus important ; généralement obtenu pour un portefeuille composé de très peu de titres très rentables mais très volatiles)<sup>37</sup>.

### 3.2.2. Diversification et taille du portefeuille

On vient de voir que la diversification a pour effet de réduire le risque d'un portefeuille grâce à la présence de différents titres non parfaitement corrélés. Examinons maintenant les limites de cette diversification : jusqu'à quel point, pour une structure de corrélation donnée, la diversification peut-elle réduire le risque ?

#### *Formulation mathématique*

Pour simplifier l'analyse, considérons un portefeuille constitué de  $N$  titres en égales proportions

$$X_j = \frac{1}{N} \quad j = 1, \dots, N.$$

Le risque du portefeuille peut alors se développer en

$$\begin{aligned} \sigma_P^2 &= \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} \\ &= \frac{1}{N^2} \left\{ \sum_{i=1}^N \sigma_i^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sigma_{ij} \right\} \end{aligned}$$

<sup>37</sup> Cette question sera examinée au § 3.2.6.

La double somme contient  $N(N - 1)$  termes ; il est donc naturel de définir la variance moyenne et la covariance moyenne par

$$\begin{aligned}\overline{\text{var}} &= \frac{1}{N} \sum_{i=1}^N \sigma_i^2 \\ \overline{\text{cov}} &= \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sigma_{ij}\end{aligned}$$

Ces deux quantités, dès l'instant où  $N$  est suffisamment grand, ne varient plus guère avec  $N$ . Elle permettent alors d'écrire la variance du portefeuille :

$$\sigma_P^2 = \frac{1}{N} \overline{\text{var}} + \frac{N-1}{N} \overline{\text{cov}}.$$

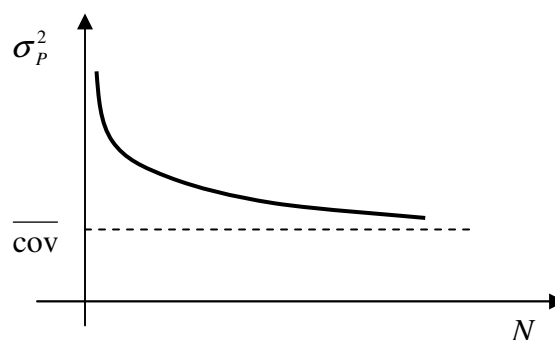
### *Comportement asymptotique*

Lorsque  $N$  devient très grand, le premier terme décroît vers 0 alors que le second, assez stable, converge vers  $\overline{\text{cov}}$ . Le risque du portefeuille, même très bien diversifié, ne descendra jamais sous cette dernière valeur, qui correspond à

$$\overline{\text{cov}} = \lim_{N \rightarrow \infty} \left[ \frac{1}{N} \overline{\text{var}} + \frac{N-1}{N} \overline{\text{cov}} \right] = \lim_{N \rightarrow \infty} \sigma_P^2 = \sigma_M^2,$$

c'est-à-dire au risque de marché.

On peut représenter le comportement de la variance du portefeuille en fonction du nombre de titres par le graphe de la figure 3.6.



*Figure 3.6 – Diversification et taille du portefeuille*

Les effets de la diversification sont très rapides au début (le premier terme perd 80 % de sa valeur lorsque l'on passe de 1 à 5 titres) pour se stabiliser rapidement autour de la valeur  $\overline{\text{cov}}$ .

### 3.2.3. Modèle de Markowitz et algorithme de la ligne critique

#### *Première formulation*

La frontière efficiente est la partie « nord-ouest » de la courbe composée des portefeuilles définis par le principe : pour chaque valeur fixée  $r$  de  $E_P$ , on détermine les proportions  $X_j$  ( $j = 1, \dots, N$ ) pour lesquelles  $\sigma_P^2$  est minimal. On définira donc la frontière efficiente en donnant à  $r$  toutes les valeurs possibles.

Mathématiquement, le problème se présente donc comme la recherche du minimum par rapport aux  $X_1, \dots, X_N$  de la fonction

$$\sigma_P^2 = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij},$$

sous la double contrainte

$$\begin{cases} \sum_{j=1}^N X_j E_j = r \\ \sum_{j=1}^N X_j = 1 \end{cases}$$

Le lagrangien<sup>38</sup> du problème peut donc s'écrire

$$L(X_1, \dots, X_N; m_1, m_2) = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} + m_1 \cdot \left( \sum_{j=1}^N X_j E_j - r \right) + m_2 \cdot \left( \sum_{j=1}^N X_j - 1 \right).$$

L'annulation des dérivées partielles par rapport aux variables  $X_1, \dots, X_N$  et aux multiplicateurs de Lagrange  $m_1$  et  $m_2$  conduit au système de  $N + 2$  équations à  $N + 2$  inconnues

$$\begin{cases} L'_{X_j}(X_1, \dots, X_N; m_1, m_2) = 2 \sum_{i=1}^N X_i \sigma_{ij} + m_1 E_j + m_2 = 0 & (j = 1, \dots, N) \\ L'_{m_1}(X_1, \dots, X_N; m_1, m_2) = \sum_{i=1}^N X_i E_i - r = 0 \\ L'_{m_2}(X_1, \dots, X_N; m_1, m_2) = \sum_{i=1}^N X_i - 1 = 0 \end{cases}$$

qui peut se réécrire matriciellement sous la forme

<sup>38</sup> On consultera, pour la théorie des extrema, l'annexe 1.



$$\begin{pmatrix} 2\sigma_1^2 & 2\sigma_{12} & \cdots & 2\sigma_{1N} & E_1 & 1 \\ 2\sigma_{21} & 2\sigma_2^2 & \cdots & 2\sigma_{2N} & E_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 2\sigma_{N1} & 2\sigma_{N2} & \cdots & 2\sigma_N^2 & E_N & 1 \\ E_1 & E_2 & \cdots & E_N & . & . \\ 1 & 1 & \cdots & 1 & . & . \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \\ m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} . \\ . \\ \vdots \\ . \\ r \\ 1 \end{pmatrix}.$$

En notant respectivement  $M$ ,  $X^*$  et  $G$  la matrice des coefficients<sup>39</sup>, le vecteur des inconnues<sup>40</sup> et le vecteur des seconds membres, le système prend la forme

$$MX^* = G.$$

La résolution de ce système passe par la matrice inverse de  $M$  :

$$X^* = M^{-1}G.$$

### Remarque 1

Ce vecteur ne fournit en réalité qu'un point stationnaire du lagrangien ; on peut montrer (mais nous ne le ferons pas ici) qu'il constitue la solution du problème de minimisation qui nous occupe.♣

### Remarque 2

Il faut donc appliquer cette relation pour les différentes valeurs possibles de  $r$  pour trouver la frontière, dont on ne gardera que la partie efficiente (« nord-ouest »). L'intérêt de ce résultat est que, si  $r$  intervient bien dans le vecteur  $G$ , il n'apparaît pas dans la matrice  $M$ , que l'on ne doit donc inverser qu'une seule fois.♣

### Illustration<sup>41</sup>

Déterminons la frontière efficiente qu'il est possible de construire avec trois titres dont les caractéristiques sont

$$\begin{array}{lll} E_1 = 0,05 & E_2 = 0,08 & E_3 = 0,10 \\ \sigma_1 = 0,10 & \sigma_2 = 0,12 & \sigma_3 = 0,15 \\ \rho_{12} = 0,3 & \rho_{13} = 0,1 & \rho_{23} = 0,4 \end{array}$$

<sup>39</sup> Celle-ci contient, dans sa zone d'ordre  $N$  du coin supérieur gauche la matrice  $2V$ , où  $V$  est la matrice des variances-covariances.

<sup>40</sup> Le vecteur des inconnues ne contient pas uniquement les proportions ; il fait également intervenir les multiplicateurs de Lagrange (qui nous seront plus utiles pour la suite) ; c'est pourquoi nous utilisons la notation  $X^*$  au lieu de  $X$  (réservée au vecteur des proportions). Cette remarque est valable pour les différents modèles développés par la suite.

<sup>41</sup> Dans le CD-Rom joint, on trouvera des illustrations plus réalistes des différents modèles dans un fichier Excel<sup>TM</sup> nommé « PfAct ». Les modèles sont repérés au niveau des onglets par référence aux paragraphes correspondants.

La matrice des variances-covariances est donnée par

$$V = \begin{pmatrix} 0,0100 & 0,0036 & 0,0015 \\ 0,0036 & 0,0144 & 0,0072 \\ 0,0015 & 0,0072 & 0,0225 \end{pmatrix}$$

et la matrice  $M$  est par conséquent égale à

$$M = \begin{pmatrix} 0,0200 & 0,0072 & 0,0030 & 0,05 & 1 \\ 0,0072 & 0,0288 & 0,0144 & 0,08 & 1 \\ 0,0030 & 0,0144 & 0,0450 & 0,10 & 1 \\ 0,05 & 0,08 & 0,10 & . & . \\ 1 & 1 & 1 & . & . \end{pmatrix},$$

qui s'inverse en

$$M^{-1} = \begin{pmatrix} 6,02 & -15,04 & 9,02 & -23,37 & 2,13 \\ -15,04 & 37,59 & -22,56 & 8,42 & -0,32 \\ 9,02 & -22,56 & 13,53 & 14,95 & -0,81 \\ -23,37 & 8,42 & 14,95 & -21,71 & 1,45 \\ 2,13 & -0,32 & -0,81 & 1,45 & -0,11 \end{pmatrix}.$$

En appliquant cette matrice au vecteur

$$G = \begin{pmatrix} . \\ . \\ . \\ r \\ 1 \end{pmatrix},$$

pour différentes valeurs de  $r$ , on trouve un ensemble de vecteurs  $X^*$ , dont les trois premières composantes fournissent la composition des portefeuilles (voir tableau 3.1) :

$r$	$X_1$	$X_2$	$X_3$
0,00	2,1293	-0,3233	-0,8060
0,01	1,8956	-0,2391	-0,6565
0,02	1,6620	-0,1549	-0,5071
0,03	1,4283	-0,0707	-0,3576
0,04	1,1946	0,0135	-0,2081
0,05	0,9609	0,0977	-0,0586
0,06	0,7272	0,1820	0,0908
0,07	0,4935	0,2662	0,2403
0,08	0,2598	0,3504	0,3898
0,09	0,0262	0,4346	0,5392
0,10	-0,2075	0,5188	0,6887
0,11	-0,4412	0,6030	0,8382
0,12	-0,6749	0,6872	0,9877
0,13	-0,9086	0,7714	1,1371
0,14	-1,1423	0,8556	1,2866
0,15	-1,3759	0,9398	1,4361

*Tableau 3.1 – Composition des portefeuilles*

Ces proportions permettent de calculer<sup>42</sup>  $\sigma_P$  pour les différents portefeuilles (tableau 3.2) :

$E_P$	$\sigma_P$
0,00	0,2348
0,01	0,2043
0,02	0,1746
0,03	0,1465
0,04	0,1207
0,05	0,0994
0,06	0,0857
0,07	0,0835
0,08	0,0937
0,09	0,1130
0,10	0,1376
0,11	0,1651
0,12	0,1943
0,13	0,2245
0,14	0,2554
0,15	0,2868

*Tableau 3.2 – Return espéré et risque des portefeuilles*

<sup>42</sup> Le return espéré est bien entendu connu.

On peut donc, à partir de là, construire la courbe représentative de ces portefeuilles (figure 3.7).

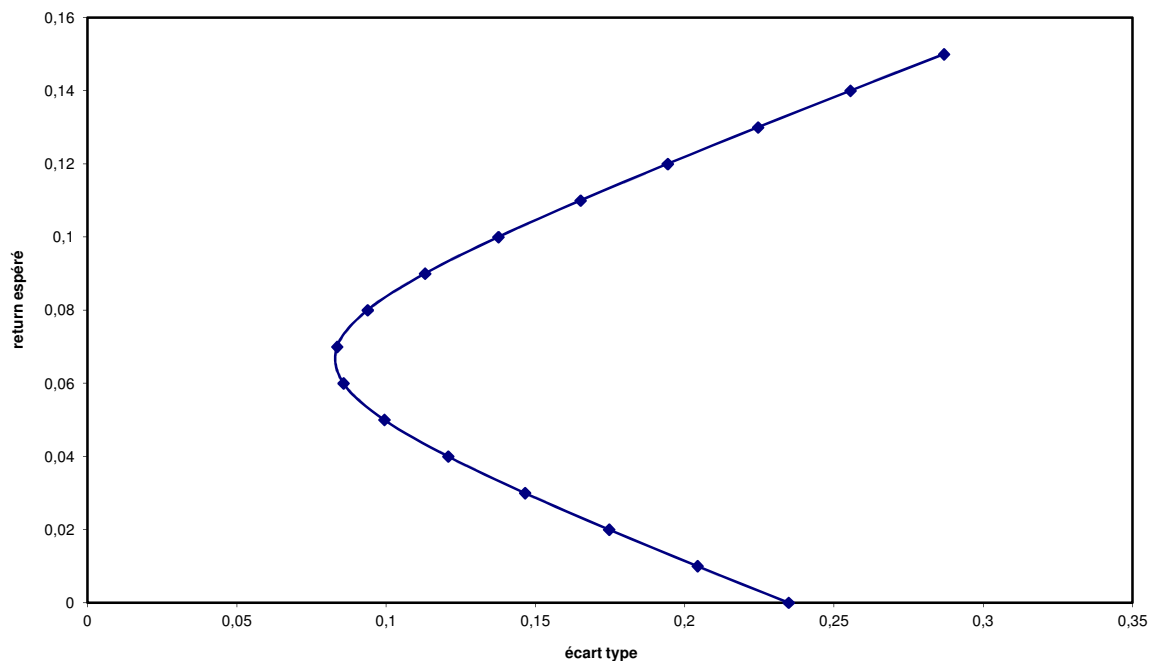


Figure 3.7 – Frontière efficiente

La partie efficiente de cette frontière en est donc la partie « nord-ouest », dont la limite inférieure correspond au portefeuille de risque minimal ; pour ce dernier, on a  $E_P = 0,0667$  et  $\sigma_P = 0,0828$ .♣

La méthode qui vient d'être présentée n'impose pas aux proportions d'être positives. On constate d'ailleurs en examinant l'illustration précédente que des valeurs négatives (et des valeurs supérieures à 1) sont parfois obtenues, les portefeuilles « classiques » ( $0 \leq X_j \leq 1$  pour tout  $j$ ) correspondant seulement aux valeurs du return espéré comprises entre 0,06 et 0,09.

Une valeur négative pour une proportion correspond à une *vente à découvert*. Ce genre d'opération – extrêmement hasardeuse – n'est pas toujours autorisée, particulièrement dans la gestion des fonds de placement. Symétriquement, une proportion supérieure à 1 consiste en l'acquisition d'un titre pour un montant supérieur au montant total investi.

Par ailleurs, dans de nombreux portefeuilles, des contraintes réglementaires ou internes imposent que certains types de titres ne soient pas représentés à une hauteur supérieure à un pourcentage fixé. Il faut donc, dans ce cas, pouvoir résoudre le problème de construire des portefeuilles où on impose en outre aux proportions du type

$$B_j^- \leq X_j \leq B_j^+ \quad j = 1, \dots, N.$$

Nous envisagerons cette problématique plus tard.

## Reformulation du problème

Nous continuons ici à considérer le problème où on n'impose pas des contraintes d'inégalité aux proportions. Nous modifions simplement légèrement l'approche, qui fournira la même solution mais qui pourra plus aisément se généraliser aux divers modèles envisagés par la suite.

Si, au lieu de représenter graphiquement les portefeuilles en portant  $\sigma_P$  en abscisse et  $E_P$  en ordonnée (comme sur la figure 3.7), on porte  $E_P$  en abscisse et  $\sigma_P^2$  en ordonnée, le graphe de la frontière efficiente prend l'allure représentée sur la figure 3.8.

Une droite, dans ce plan, a pour équation  $\sigma^2 = a + \lambda E$  où  $a$  représente l'ordonnée à l'origine et  $\lambda$  la pente de cette droite. Considérons en particulier une droite tangente à la frontière efficiente. Si la pente de cette droite est nulle ( $\lambda = 0$ ), le point de contact de la tangente est le portefeuille le moins risqué de la frontière efficiente. A l'opposé, plus  $\lambda$  augmente, plus le point de contact s'éloigne sur la frontière efficiente vers des portefeuilles risqués. Le paramètre  $\lambda$  peut varier de 0 à  $+\infty$  et est donc représentatif du risque du portefeuille correspondant au point de contact de la tangente ayant cette valeur  $\lambda$  pour pente.

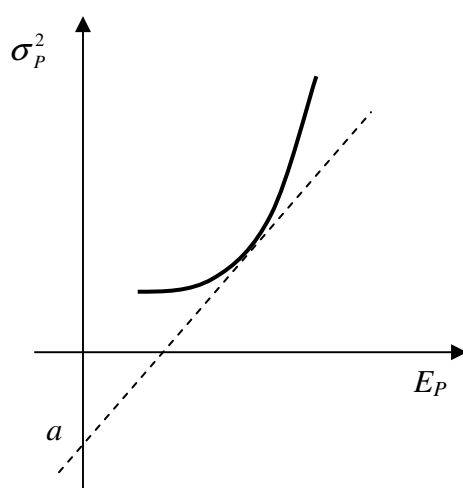


Figure 3.8 – Reformulation du problème

Pour une valeur de  $\lambda$  fixée, la tangente de pente  $\lambda$  à la frontière efficiente est, parmi toutes les droites ayant cette pente et ayant au moins un point en commun avec la frontière efficiente, celle qui est située le plus à droite, c'est-à-dire pour laquelle l'ordonnée à l'origine  $a = \sigma^2 - \lambda E$  est la plus petite.

Le problème se reformule donc de la manière suivante : pour les différentes valeurs de  $\lambda$  entre 0 et  $+\infty$ , minimiser par rapport aux proportions  $X_1, \dots, X_N$  l'expression

$$\sigma_P^2 - \lambda E_P = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} - \lambda \sum_{j=1}^N X_j E_j,$$

sous la contrainte

$$\sum_{j=1}^N X_j = 1.$$

Une fois trouvée la solution, qui dépendra de  $\lambda$ , il suffira de faire varier ce dernier paramètre de 0 à  $+\infty$  pour obtenir la frontière efficiente.

Le lagrangien du problème peut s'écrire

$$L(X_1, \dots, X_N; m) = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} - \lambda \sum_{j=1}^N X_j E_j + m \cdot \left( \sum_{j=1}^N X_j - 1 \right).$$

Un raisonnement en tous points similaire à celui tenu dans la première formulation permet de déduire, de l'annulation des dérivées partielles, l'expression matricielle

$$MX^* = \lambda E^* + F,$$

où on a posé<sup>43</sup>

$$M = \begin{pmatrix} 2\sigma_1^2 & 2\sigma_{12} & \cdots & 2\sigma_{1N} & 1 \\ 2\sigma_{21} & 2\sigma_2^2 & \cdots & 2\sigma_{2N} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2\sigma_{N1} & 2\sigma_{N2} & \cdots & 2\sigma_N^2 & 1 \\ 1 & 1 & \cdots & 1 & . \end{pmatrix} \quad X^* = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \\ m \end{pmatrix} \quad E^* = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_N \\ . \end{pmatrix} \quad F = \begin{pmatrix} . \\ . \\ \vdots \\ . \\ 1 \end{pmatrix}.$$

La solution de ce système d'équations est donc fournie par

$$X^* = \lambda(M^{-1}E^*) + (M^{-1}F).$$

Comme pour la première formulation, la matrice  $M$  est indépendante du paramètre  $\lambda$  qui doit varier ; elle ne doit donc être inversée qu'une seule fois.

### Illustration

Reprenons les mêmes données que celles utilisées pour la première formulation, à savoir

$$\begin{array}{lll} E_1 = 0,05 & E_2 = 0,08 & E_3 = 0,10 \\ \sigma_1 = 0,10 & \sigma_2 = 0,12 & \sigma_3 = 0,15 \\ \rho_{12} = 0,3 & \rho_{13} = 0,1 & \rho_{23} = 0,4 \end{array}$$

On a bien entendu la même matrice  $V$  des variances-covariances que ci-dessus et la matrice  $M$  peut se mettre sous la forme

<sup>43</sup> A l'instar de ce qui a été fait pour  $X^*$ , nous utilisons ici la notation  $E^*$  car  $E$  est réservé au vecteur de dimension  $N$  des returns espérés.

$$M = \begin{pmatrix} 0,0200 & 0,0072 & 0,0030 & 1 \\ 0,0072 & 0,0288 & 0,0144 & 1 \\ 0,0030 & 0,0144 & 0,0450 & 1 \\ 1 & 1 & 1 & . \end{pmatrix},$$

qui s'inverse en

$$M^{-1} = \begin{pmatrix} 31,16 & -24,10 & -7,06 & 0,57 \\ -24,10 & 40,86 & -16,76 & 0,24 \\ -7,06 & -16,76 & 23,82 & 0,19 \\ 0,57 & 0,24 & 0,19 & -0,01 \end{pmatrix}.$$

Les vecteurs  $E^*$  et  $F$  étant donnés par

$$E^* = \begin{pmatrix} 0,05 \\ 0,08 \\ 0,10 \\ . \end{pmatrix} \quad F = \begin{pmatrix} . \\ . \\ . \\ 1 \end{pmatrix},$$

les solutions du problème pour différentes valeurs de  $\lambda$  sont données dans le tableau 3.3.

$\lambda$	$X_1$	$X_2$	$X_3$	$E_P$	$\sigma_P$
2,0	-1,5810	1,0137	1,5672	0,1588	0,3146
1,9	-1,4734	0,9750	1,4984	0,1542	0,3000
1,8	-1,3657	0,9362	1,4296	0,1496	0,2854
1,7	-1,2581	0,8974	1,3607	0,1450	0,2709
1,6	-1,1505	0,8586	1,2919	0,1404	0,2565
1,5	-1,0429	0,8198	1,2231	0,1357	0,2422
1,4	-0,9353	0,7810	1,1542	0,1311	0,2280
1,3	-0,8276	0,7423	1,0854	0,1265	0,2139
1,2	-0,7200	0,7035	1,0165	0,1219	0,2000
1,1	-0,6124	0,6647	0,9477	0,1173	0,1863
1,0	-0,5048	0,6259	0,8789	0,1127	0,1729
0,9	-0,3972	0,5871	0,8100	0,1081	0,1597
0,8	-0,2895	0,5484	0,7412	0,1035	0,1470
0,7	-0,1819	0,5096	0,6723	0,0989	0,1347
0,6	-0,0743	0,4708	0,6035	0,0943	0,1231
0,5	0,0333	0,4320	0,5347	0,0897	0,1123
0,4	0,1409	0,3932	0,4658	0,0851	0,1027
0,3	0,2486	0,3544	0,3970	0,0805	0,0945
0,2	0,3562	0,3157	0,3282	0,0759	0,0882
0,1	0,4638	0,2769	0,2593	0,0713	0,0842
0,0	0,5714	0,2381	0,1905	0,0667	0,0828

Tableau 3.3 – Solution pour le problème reformulé

Le graphe de la frontière efficiente se présente alors sous la forme présentée à la figure 3.9.

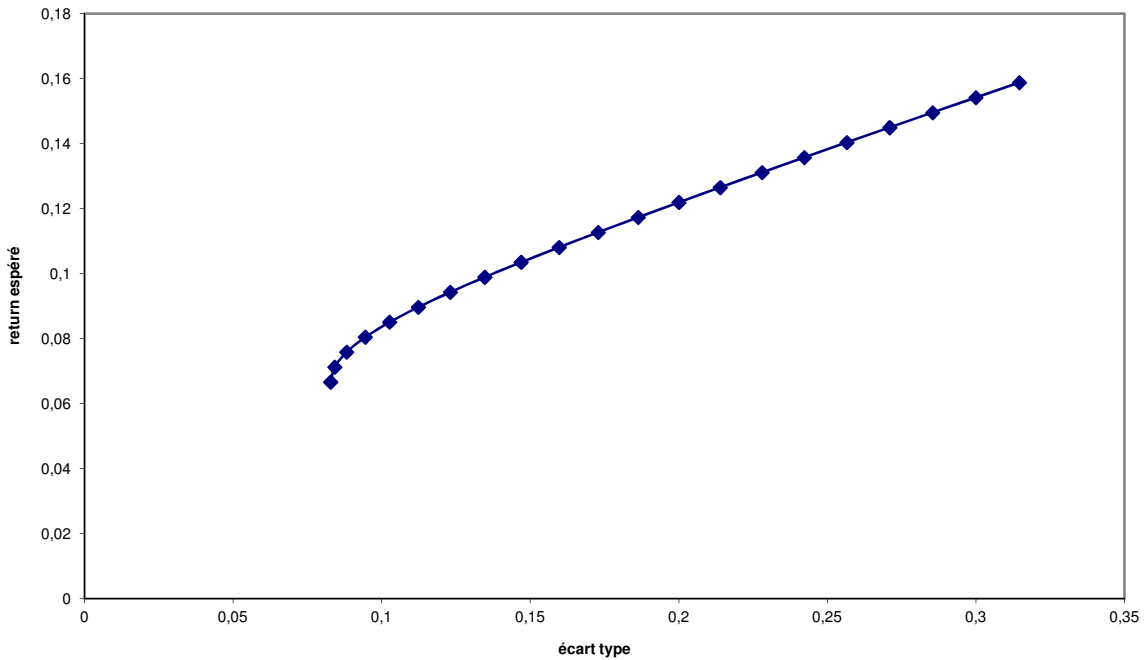


Figure 3.9 – Frontière efficiente pour le problème reformulé♣

L'avantage de cette nouvelle formulation est double.

D'une part, elle ne fournit plus que les portefeuilles vraiment efficients, au lieu de la frontière de l'ensemble des portefeuilles qu'il est possible de construire et donc il fallait sélectionner la partie supérieure.

D'autre part, elle se prête aisément à la généralisation au cas du problème avec contraintes d'inégalité, ainsi qu'aux modèles à indice simple et avec titre non risqué.

### Modèle de Markowitz avec contraintes

Le problème à résoudre se formule ici de la manière suivante : pour les différentes valeurs de  $\lambda$  entre 0 et  $+\infty$ , minimiser par rapport aux proportions  $X_1, \dots, X_N$  l'expression

$$\sigma_p^2 - \lambda E_p = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} - \lambda \sum_{j=1}^N X_j E_j,$$

sous les contraintes

$$\begin{cases} \sum_{j=1}^N X_j = 1 \\ B_j^- \leq X_j \leq B_j^+ \end{cases} \quad j = 1, \dots, N$$



Introduisons tout d'abord la notion de statut d'un titre. Le titre ( $j$ ) est qualifié de « down » (resp. « up ») si sa proportion est égale à la borne inférieure (resp. supérieure) qui lui est imposée :  $X_j = B_j^-$  (resp.  $X_j = B_j^+$ ). Pour un portefeuille efficient (c'est-à-dire pour lequel le lagrangien est minimum), la dérivée partielle du lagrangien par rapport à  $X_j$  n'est pas nulle à l'optimum : elle est strictement positive (resp. strictement négative), comme on le voit sur la figure 3.10.

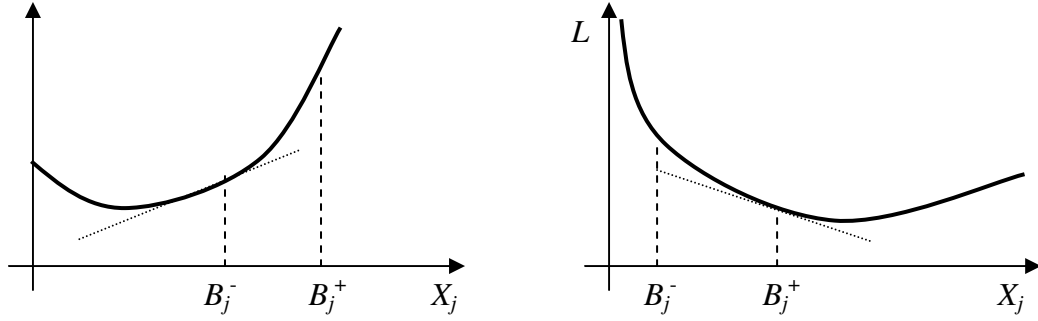


Figure 3.10 – Titre « down » et titre « up »

Dans le système d'équations issu de l'annulation des dérivées partielles du lagrangien, les équations relatives aux titres « down » (resp. « up ») devront donc ici être remplacées par  $X_j = B_j^-$  (resp.  $X_j = B_j^+$ ).

Les autres titres seront qualifiés de « in » ; ils sont tels que  $B_j^- < X_j < B_j^+$  et, à l'optimum, la dérivée partielle du lagrangien par rapport à  $X_j$  est nulle. Les équations relatives à ces titres ne doivent pas être modifiées.

L'adaptation du système d'équations issu de l'annulation des dérivées partielles du lagrangien

$$MX^* = \lambda E^* + F,$$

va donc consister à ne pas modifier les composantes correspondant aux titres « in » et, si le titre ( $j$ ) est « down » ou « up », à modifier la  $j$ -ème ligne de  $M$ , et la  $j$ -ème composante de  $E^*$  et de  $F$ , de la manière suivante :

$$M = \begin{pmatrix} 2\sigma_1^2 & \cdots & 2\sigma_{1j} & \cdots & 2\sigma_{1N} & 1 \\ \vdots & \ddots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \cdots & \vdots & \vdots \\ 2\sigma_{N1} & \cdots & 2\sigma_{jN} & \cdots & 2\sigma_N^2 & 1 \\ 1 & \cdots & 1 & \cdots & 1 & 0 \end{pmatrix} \quad E^* = \begin{pmatrix} E_1 \\ \vdots \\ 0 \\ \vdots \\ E_N \\ 0 \end{pmatrix} \quad F = \begin{pmatrix} 0 \\ \vdots \\ B_j^\pm \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

En effet, avec cette modification, la  $j$ -ème équation devient  $X_j = B_j^\pm$ . D'ailleurs, en considérant la  $j$ -ème ligne de l'égalité  $M^{-1}M = I$ , on voit aisément que  $M^{-1}$  a la même  $j$ -ème ligne que  $M$  et la  $j$ -ème composante de la solution

$$X^* = \lambda(M^{-1}E^*) + (M^{-1}F),$$

s'écrit aussi  $X_j = B_j^\pm$ .

Si  $(j)$  a le statut « in », cette même  $j$ -ème composante peut bien sûr s'écrire  $X_j = \lambda u_j + v_j$ , quantité qui est strictement comprise entre  $B_j^-$  et  $B_j^+$ .

La méthode procède par étapes et nous noterons  $M_0, E_0^*$  et  $F_0$  les éléments matriciels tels qu'ils étaient définis dans le cas « sans contrainte ». L'indice évoluera d'étape en étape.

La méthode commence avec les grandes valeurs de  $\lambda$  ( $+\infty$  idéalement). Comme il s'agit de minimiser  $\sigma_P^2 - \lambda E_P$ , il faut que  $E_P$  soit le plus grand possible, ce qui est bien cohérent avec une grande valeur du paramètre de risque  $\lambda$ . Le premier portefeuille sera donc constitué des titres offrant les plus grands returns espérés, en proportions égales aux bornes supérieures  $B_j^+$ , jusqu'à ce que – avec les titres en proportions égales à  $B_j^-$  – la somme des proportions soit égale à 1<sup>44</sup>. Ce portefeuille est appelé premier *portefeuille-coin*.

Au moins un titre est donc « up », un est « in » et les autres sont « down ». On modifie la matrice  $M$  et les vecteurs  $E^*$  et  $F$  comme indiqué ci-dessus. On obtient alors  $M_1, E_1^*$  et  $F_1$  et on calcule

$$X^* = \lambda(M_1^{-1}E_1^*) + (M_1^{-1}F_1)$$

On fait alors décroître le paramètre  $\lambda$  jusqu'à ce qu'un des titres change de statut<sup>45</sup>. Ce premier changement se produira pour une valeur de  $\lambda$  égale à  $\lambda_c^{(1)}$ , appelé premier  $\lambda$  critique. Pour déterminer cette valeur critique ainsi que le titre changeant de statut, on examine les différents titres pour chacun desquels on définit un  $\lambda_j$  potentiellement critique.

Pour un titre  $(j)$  « down » ou « up », il changera de statut si l'équation qui lui correspond devient  $L'_{X_j} = 0$ , c'est-à-dire

$$2 \sum_{k=1}^N X_k \sigma_{jk} - \lambda_j E_j + m = 0,$$

qui n'est autre que la  $j$ -ème composante de l'équation  $M_0 X^* = \lambda E_0^* + F_0$  dans laquelle on donne aux  $X_k$  et à  $m$  les valeurs obtenues par  $X^* = \lambda(M_1^{-1}E_1^*) + (M_1^{-1}F_1)$ .

<sup>44</sup> Si les contraintes d'inégalités sont simplement  $0 \leq X_j \leq 1 \quad \forall j$  (absence de ventes à découvert), le premier portefeuille sera constitué du seul titre ayant le plus grand return espéré.

<sup>45</sup> Pour les contraintes  $0 \leq X_j \leq 1 \quad \forall j$ , le premier portefeuille-coin est constitué d'un seul titre « up », les autres étant tous « down ». Le premier changement de statut consistera en le passage à « in » du titre qui était « up » et d'un des titres qui étaient « down ». Dans ce cas, d'une part, les éléments matriciels  $M_1, E_1^*$  et  $F_1$  sont obtenus en effectuant la modification prescrite pour les titres « down », mais pas pour celui dont on sait qu'il passe au statut « in » et, d'autre part, il n'y a pas d'équation de détermination des  $\lambda$  critiques potentiels pour ce titre.

Pour un titre ( $j$ ) « in », on sait que  $X_j = \lambda_j u_j + v_j$  et celui-ci changera de statut s'il devient « down » ( $u_j > 0$ , car  $\lambda$  diminue) ou « up » ( $u_j < 0$ ), auquel cas on a

$$B_j^\pm = \lambda_j u_j + v_j$$

qui n'est autre que la  $j$ -ème composante de la relation  $X^* = \lambda(M_1^{-1}E_1^*) + (M_1^{-1}F_1)$  où le membre de gauche est remplacé par la borne inférieure ou supérieure suivant le cas.

On obtient donc  $N$  équations pour  $N$  valeurs de  $\lambda_j$  critiques potentiels. Le plus grand de ceux-ci est le premier  $\lambda$  critique :  $\lambda_c^{(1)}$ . Les proportions des différents titres n'ont pas varié entre  $\lambda = +\infty$  et  $\lambda = \lambda_c^{(1)}$  ; Le portefeuille correspondant est donc toujours le premier portefeuille-coin.

Le titre correspondant au  $\lambda$  critique change donc de statut, ce qui permet de construire  $M_2, E_2^*$  et  $F_2$  et de déterminer le deuxième  $\lambda$  critique  $\lambda_c^{(2)}$  ainsi que l'ensemble des portefeuilles correspondant aux valeurs de  $\lambda$  comprises entre  $\lambda_c^{(1)}$  et  $\lambda_c^{(2)}$  ; le portefeuille correspondant à  $\lambda_c^{(2)}$  est bien entendu le deuxième portefeuille-coin.

On recommence le processus jusqu'à ce que tous les  $\lambda$  critiques potentiels soient négatifs, auquel cas le dernier  $\lambda$  critique est égal à 0, auquel correspond le dernier portefeuille-coin (le moins risqué), qui est l'extrémité gauche de la frontière efficiente.

Les portefeuilles-coins font bien sûr partie de la frontière efficiente. Entre deux portefeuilles-coins consécutifs, les statuts des titres n'évoluent pas : seules les proportions changent. Bien sûr, ces dernières se calculent, entre  $\lambda_c^{(k-1)}$  et  $\lambda_c^{(k)}$ , à l'aide de la relation

$$X^* = \lambda(M_k^{-1}E_k^*) + (M_k^{-1}F_k).$$

Les différents tronçons de courbes ainsi construits se raccordent de manière continue et dérivable<sup>46</sup> et constituent la frontière efficiente.

### Illustration

Reprenons les mêmes données que celles traitées précédemment :

$$\begin{array}{lll} E_1 = 0,05 & E_2 = 0,08 & E_3 = 0,10 \\ \sigma_1 = 0,10 & \sigma_2 = 0,12 & \sigma_3 = 0,15 \\ \rho_{12} = 0,3 & \rho_{13} = 0,1 & \rho_{23} = 0,4 \end{array}$$

et imposons les contraintes d'absence de ventes à découvert :  $0 \leq X_j \leq 1$  ( $j = 1, 2, 3$ ).

On a les éléments matriciels de base :

<sup>46</sup> C'est-à-dire avec la même tangente.

$$M_0 = \begin{pmatrix} 0,0200 & 0,0072 & 0,0030 & 1 \\ 0,0072 & 0,0288 & 0,0144 & 1 \\ 0,0030 & 0,0144 & 0,0450 & 1 \\ 1 & 1 & 1 & . \end{pmatrix} \quad E_0^* = \begin{pmatrix} 0,05 \\ 0,08 \\ 0,10 \\ . \end{pmatrix} \quad F_0 = \begin{pmatrix} . \\ . \\ . \\ 1 \end{pmatrix},$$

Le premier portefeuille-coin est constitué du seul titre (3), son return espéré étant le plus grand. Les titres (1) et (2) étant « down », on construit

$$M_1 = \begin{pmatrix} 1 & . & . & . \\ . & 1 & . & . \\ 0,0030 & 0,0144 & 0,0450 & 1 \\ 1 & 1 & 1 & . \end{pmatrix} \quad E_1^* = \begin{pmatrix} . \\ . \\ 0,10 \\ . \end{pmatrix} \quad F_1 = \begin{pmatrix} . \\ . \\ . \\ 1 \end{pmatrix}.$$

On a

$$M_1^{-1} = \begin{pmatrix} 1 & . & . & . \\ . & 1 & . & . \\ -1 & -1 & . & 1 \\ 0,0420 & 0,0306 & 1 & -0,0450 \end{pmatrix}$$

et par conséquent

$$X^* = \lambda(M_1^{-1}E_1^*) + (M_1^{-1}F_1) = \lambda \begin{pmatrix} . \\ . \\ . \\ 0,1 \end{pmatrix} + \begin{pmatrix} . \\ . \\ 1 \\ -0,045 \end{pmatrix}.$$

Les deux premières composantes de  $M_0X^* = \lambda E_0^* + F_0$ , avec le vecteur  $X^*$  obtenu ci-dessus donnent

$$\begin{cases} 0,003 + (0,1\lambda_1 - 0,045) = 0,05\lambda_1 \\ 0,0144 + (0,1\lambda_2 - 0,045) = 0,08\lambda_2 \end{cases}$$

ce qui fournit les deux  $\lambda$  critiques potentiels :  $\lambda_1 = 0,84$  et  $\lambda_2 = 1,53$ . Le premier  $\lambda$  critique est donc  $\lambda_c^{(1)} = 1,53$  et le titre (2) devient « in » avec (3), tandis que (1) reste « down ».

On construit donc

$$M_2 = \begin{pmatrix} 1 & . & . & . \\ 0,0072 & 0,0288 & 0,0144 & 1 \\ 0,0030 & 0,0144 & 0,0450 & 1 \\ 1 & 1 & 1 & . \end{pmatrix} \quad E_2^* = \begin{pmatrix} . \\ 0,08 \\ 0,10 \\ . \end{pmatrix} \quad F_2 = \begin{pmatrix} . \\ . \\ . \\ 1 \end{pmatrix}$$

ce qui donne successivement

$$M_1^{-1} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ -0,7733 & 22,22 & -22,22 & 0,68 \\ -0,2267 & -22,22 & 22,22 & 0,32 \\ 0,0183 & 0,68 & 0,32 & -0,0242 \end{pmatrix}$$

$$X^* = \lambda(M_2^{-1}E_2^*) + (M_2^{-1}F_2) = \lambda \begin{pmatrix} \cdot \\ -0,4444 \\ 0,4444 \\ 0,0864 \end{pmatrix} + \begin{pmatrix} \cdot \\ 0,68 \\ 0,32 \\ -0,0242 \end{pmatrix}.$$

La première composante de  $M_0 X^* = \lambda E_0^* + F_0$ , avec le vecteur  $X^*$  obtenu ci-dessus donne

$$0,0072 \cdot (-0,4444 \lambda_1 + 0,68) + 0,0030 \cdot (0,4444 \lambda_1 + 0,32) + (0,0864 \lambda_1 - 0,0242) = 0,05 \lambda_1,$$

ce qui fournit un  $\lambda$  critique potentiel :  $\lambda_1 = 0,5312$ .

Les deuxième et troisième composantes de la relation  $X^* = \lambda(M_2^{-1}E_2^*) + (M_2^{-1}F_2)$  où le membre de gauche est remplacé par la borne adéquate donnent

$$\begin{cases} -0,4444\lambda_2 + 0,68 = 1 \\ 0,4444\lambda_3 + 0,32 = 0 \end{cases}$$

et par conséquent  $\lambda_2 = \lambda_3 = -0,7201$ . Le deuxième  $\lambda$  critique est donc  $\lambda_c^{(2)} = 0,5312$  et les trois titres acquièrent le statut « in ».

Les éléments matriciels  $M_3$ ,  $E_3^*$  et  $F_3$  sont donc les mêmes que ceux de base et on est ramené au problème sans contrainte. On a

$$M_3^{-1} = \begin{pmatrix} 31,16 & -24,10 & -7,06 & 0,57 \\ -24,10 & 40,86 & -16,76 & 0,24 \\ -7,06 & -16,76 & 23,82 & 0,19 \\ 0,57 & 0,24 & 0,19 & -0,01 \end{pmatrix}$$

et donc

$$X^* = \lambda(M_3^{-1}E_3^*) + (M_3^{-1}F_3) = \lambda \begin{pmatrix} -1,0762 \\ 0,3878 \\ 0,6884 \\ 0,0667 \end{pmatrix} + \begin{pmatrix} 0,5714 \\ 0,2381 \\ 0,1905 \\ -0,0137 \end{pmatrix},$$

dont les trois premières composantes fournissent, avec les bornes adéquates,

$$\begin{cases} -1,0762\lambda_1 + 0,5714 = 1 \\ 0,3878\lambda_2 + 0,2381 = 0 \\ 0,6884\lambda_3 + 0,1905 = 0 \end{cases}$$

On obtient donc  $\lambda_1 = -0,3983$ ,  $\lambda_2 = -0,6140$  et  $\lambda_3 = -0,2767$ . Le dernier  $\lambda$  critique est donc  $\lambda_c^{(3)} = 0$  et les trois titres restent « in » jusqu'à la fin du processus<sup>47</sup>.

Les différents portefeuilles de la frontière efficiente, ainsi que le return espéré et le risque, sont donnés dans le tableau 3.4.

$\lambda$	$X_1$	$X_2$	$X_3$	$E_P$	$\sigma_P$
1,53	0	0	1	0,1000	0,1500
1,5	0	0,0133	0,9867	0,0997	0,1486
1,4	0	0,0578	0,9422	0,0988	0,1442
1,3	0	0,1022	0,8978	0,0980	0,1400
1,2	0	0,1467	0,8533	0,0971	0,1360
1,1	0	0,1911	0,8089	0,0962	0,1322
1,0	0	0,2356	0,7644	0,0953	0,1286
0,9	0	0,2800	0,7200	0,0944	0,1253
0,8	0	0,3244	0,6756	0,0935	0,1222
0,7	0	0,3689	0,6311	0,0926	0,1195
0,6	0	0,4133	0,5867	0,0917	0,1170
0,5312	0	0,4439	0,5561	0,0911	0,1155
0,5	0,0333	0,4320	0,5347	0,0897	0,1123
0,4	0,1409	0,3932	0,4658	0,0851	0,1027
0,3	0,2486	0,3544	0,3970	0,0805	0,0945
0,2	0,3562	0,3157	0,3282	0,0759	0,0882
0,1	0,4638	0,2769	0,2593	0,0713	0,0842
0,0	0,5714	0,2381	0,1905	0,0667	0,0828

Tableau 3.4 – Solution pour le modèle de Markowitz avec contraintes

Bien entendu, entre  $\lambda = 0,5312$  et  $\lambda = 0$ , les proportions obtenues ici sont les mêmes que celles obtenues dans le modèle « sans contrainte » puisque tous les titres sont « in ».

Le graphe de la frontière efficiente se présente alors sous la forme donnée par la figure 3.11.

<sup>47</sup> Il est assez logique d'avoir, pour le portefeuille efficient le moins risqué, une diversification importante.

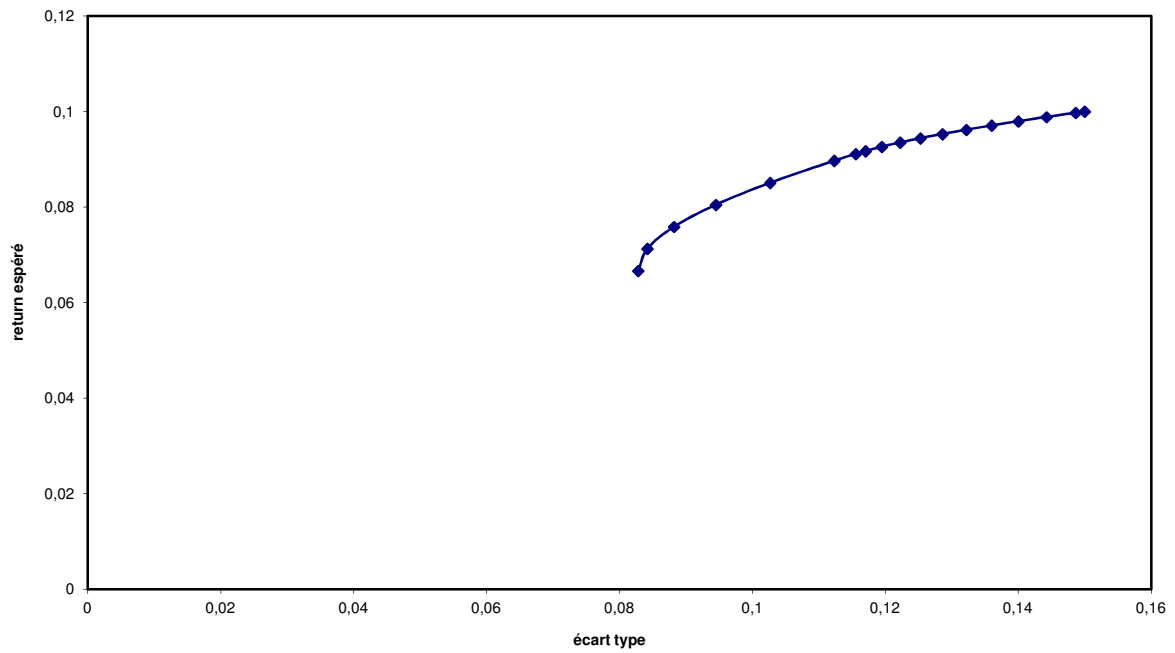


Figure 3.11 – Frontière efficiente pour le modèle de Markowitz avec contraintes

La figure 3.12 superpose les deux frontières efficientes (sans et avec contraintes). On y voit clairement les zones correspondant aux ventes à découvert et celles où tous les titres sont « in ».

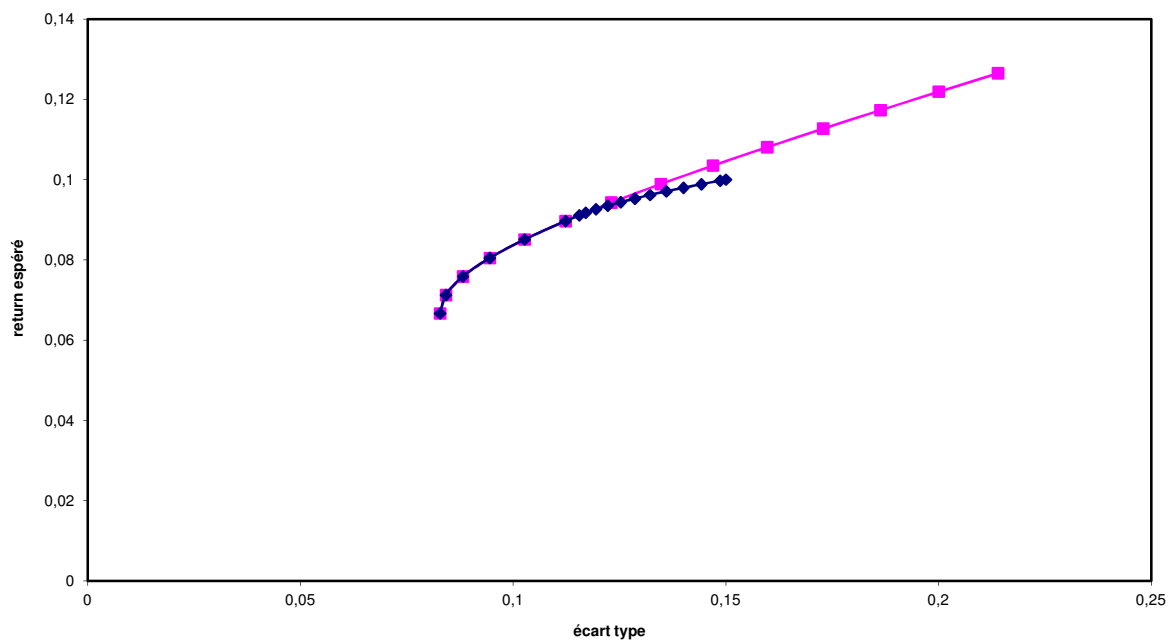


Figure 3.12 – Comparaison des frontières efficientes sans et avec contraintes♣

### Algorithme de la ligne critique

H. Markowitz a proposé une méthode algorithmique de résolution du problème avec les contraintes  $X_j \geq 0$  ( $j = 1, \dots, N$ ). Il porte le nom d'*algorithme de la ligne critique*.

Celui-ci démarre avec le premier portefeuille-coin, qui est bien sûr constitué du seul titre ayant le return espéré le plus élevé. Il parcourt alors les portefeuilles-coins successifs en testant à chaque étape l'évolution de la fonction à minimiser lorsque l'on

- fait entrer un nouveau titre dans le portefeuille ; ou
- fait sortir un titre du portefeuille ; ou
- remplace un titre du portefeuille par un qui n'y était pas déjà présent.

Le développement de cet algorithme dépasse le cadre de cet ouvrage et se trouve dans la littérature spécialisée<sup>48</sup>. On se contentera simplement d'illustrer le cheminement pour un problème à trois titres tel que celui qui est illustré dans cette section. Les contraintes

$$\begin{cases} \sum_{j=1}^3 X_j = 1 \\ 0 \leq X_j \leq 1 \end{cases} \quad j = 1, 2, 3$$

définissent dans l'espace à trois dimensions le triangle dont les sommets sont  $(1, 0, 0)$ ,  $(0, 1, 0)$  et  $(0, 0, 1)$ , comme représenté sur la figure 3.13. La ligne critique est représentée en trait gras et les points  $A$ ,  $B$  et  $C$  correspondent aux portefeuilles-coins obtenus pour  $\lambda = \lambda_c^{(1)}$ ,  $\lambda_c^{(2)}$  et  $\lambda_c^{(3)}$  respectivement.

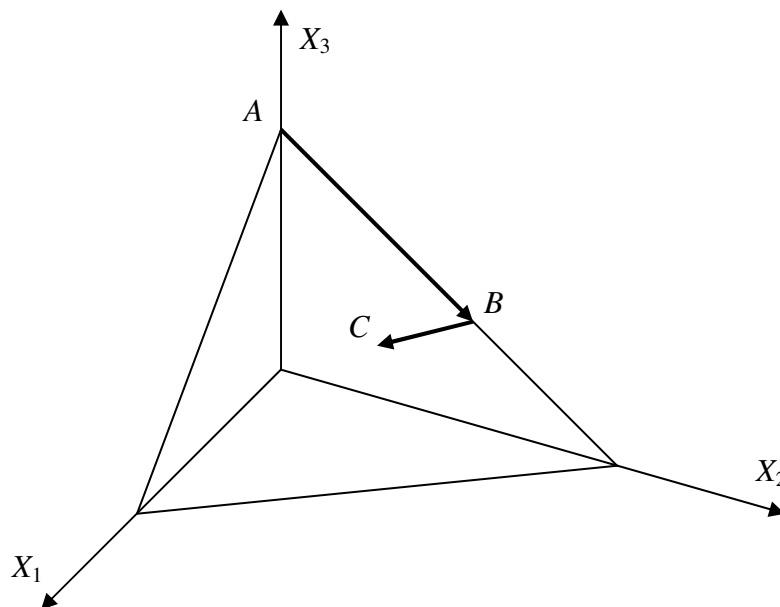


Figure 3.13 – Ligne critique

Dans cet algorithme, seuls les portefeuilles-coins sont déterminés. Ceux qui sont situés entre deux portefeuilles-coins consécutifs sont estimés comme combinaisons linéaires de ces derniers.

<sup>48</sup> Par exemple, MARKOWITZ H., *Mean-variance analysis in portfolio choice and capital markets*, Basil Blackwell, 1987.



### 3.2.4. Modèle à indice simple de Sharpe

#### *Principes*

La détermination de la frontière efficiente dans le modèle de Markowitz n'est pas très aisée. De plus, le nombre de données nécessaires est important puisque la matrice des variances-covariances est requise.

C'est la raison pour laquelle W. Sharpe<sup>49</sup> a proposé une simplification du modèle de Markowitz qui repose sur les deux hypothèses suivantes.

- a) Les returns des différents titres s'expriment comme des fonctions du premier de gré du return d'un indice représentatif du marché :

$$R_{jt} = a_j + b_j R_{It} + \varepsilon_{jt} \quad j = 1, \dots, N.$$

On suppose également que les résidus vérifient les hypothèses classiques de la régression linéaire<sup>50</sup>, entre autres que les résidus sont d'espérance nulle et qu'ils sont non corrélés avec la variable explicative  $R_{It}$ .

- b) Les résidus des régressions relatives aux différents titres sont non corrélés :  $\text{cov}(\varepsilon_{it}, \varepsilon_{jt}) = 0$  pour tout  $i$  et  $j$  différents.

En reprenant la convention d'omettre l'indice  $t$ , le return d'un portefeuille s'écrira donc, dans ce cas,

$$\begin{aligned} R_p &= \sum_{j=1}^N X_j R_j \\ &= \sum_{j=1}^N X_j (a_j + b_j R_I + \varepsilon_j) \\ &= \sum_{j=1}^N X_j a_j + \left( \sum_{j=1}^N X_j b_j \right) R_I + \sum_{j=1}^N X_j \varepsilon_j \\ &= \sum_{j=1}^N X_j a_j + Y R_I + \sum_{j=1}^N X_j \varepsilon_j \end{aligned}$$

où on a posé  $Y = \sum X_j b_j$ .

Le return espéré et la variance du portefeuille peuvent s'écrire, compte tenu des hypothèses du modèle,

<sup>49</sup> SHARPE W., A simplified model for portfolio analysis, *Management Science*, Vol. 9, n° 1, 1963, pp.277-293.

<sup>50</sup> On consultera, à ce sujet, l'annexe 3.

$$E_p = \sum_{j=1}^N X_j a_j + Y E_I$$

$$\sigma_p^2 = \sum_{j=1}^N X_j^2 \sigma_{\varepsilon_j}^2 + Y^2 \sigma_I^2$$

### Remarque 1

La variance du portefeuille peut s'écrire matriciellement à l'aide d'une forme quadratique :

$$\sigma_p^2 = \begin{pmatrix} X_1 & \dots & X_N & Y \end{pmatrix} \begin{pmatrix} \sigma_{\varepsilon_1}^2 & & & \\ & \ddots & & \\ & & \sigma_{\varepsilon_N}^2 & \\ & & & \sigma_I^2 \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_N \\ Y \end{pmatrix}.$$

La structure de cette matrice justifie le nom de *modèle diagonal* qui est aussi parfois donné au modèle à indice simple.

Cependant, contrairement à ce que cette expression pourrait laisser croire, la simplification n'est pas excessive : on ne suppose pas que les returns des différents titres sont non corrélés puisque

$$\begin{aligned} \sigma_{ij} &= \text{cov}(a_i + b_i R_I + \varepsilon_i, a_j + b_j R_I + \varepsilon_j) \\ &= b_i b_j \sigma_I^2 \end{aligned} \clubsuit$$

### Remarque 2

En pratique, les coefficients  $a_j$  et  $b_j$  des différentes régressions seront estimés par la méthode des moindres carrés :  $(\hat{a}_j; \hat{b}_j)$ . Les résidus seront estimés par la relation

$$\hat{\varepsilon}_{jt} = R_{jt} - (\hat{a}_j + \hat{b}_j R_{It}).$$

C'est à partir de ces estimations que l'on déterminera les variances résiduelles  $\sigma_{\varepsilon_j}^2$ , par leur estimateur ergodique. ♣

### Modèle à indice simple

Nous devons donc résoudre le problème suivant : pour les différentes valeurs de  $\lambda$  entre 0 et  $+\infty$ , minimiser par rapport aux proportions  $X_1, \dots, X_N$  et à la variable  $Y$ , l'expression

$$\sigma_p^2 - \lambda E_p = \sum_{j=1}^N X_j^2 \sigma_{\varepsilon_j}^2 + Y^2 \sigma_I^2 - \lambda \cdot \left( \sum_{j=1}^N X_j a_j + Y E_I \right),$$

sous les contraintes

$$\begin{cases} \sum_{j=1}^N X_j b_j = Y \\ \sum_{j=1}^N X_j = 1 \end{cases}$$

Le lagrangien du problème s'écrit

$$\begin{aligned} L(X_1, \dots, X_N, Y; m_1, m_2) \\ = \sum_{j=1}^N X_j^2 \sigma_{\varepsilon_j}^2 + Y^2 \sigma_I^2 - \lambda \cdot \left( \sum_{j=1}^N X_j a_j + Y E_I \right) + m_1 \cdot \left( \sum_{j=1}^N X_j b_j - Y \right) + m_2 \cdot \left( \sum_{j=1}^N X_j - 1 \right) \end{aligned}$$

L'annulation des dérivées partielles de ce lagrangien s'exprime matriciellement par

$$MX^* = \lambda E^* + F,$$

où on a posé

$$M = \begin{pmatrix} 2\sigma_{\varepsilon_1}^2 & & . & . & b_1 & 1 \\ & \ddots & & \vdots & \vdots & \vdots \\ . & & 2\sigma_{\varepsilon_N}^2 & . & b_N & 1 \\ . & \cdots & . & 2\sigma_I^2 & -1 & . \\ b_1 & \cdots & b_N & -1 & . & . \\ 1 & \cdots & 1 & . & . & . \end{pmatrix} \quad X^* = \begin{pmatrix} X_1 \\ \vdots \\ X_N \\ Y \\ m_1 \\ m_2 \end{pmatrix} \quad E^* = \begin{pmatrix} a_1 \\ \vdots \\ a_N \\ E_I \\ . \\ . \end{pmatrix} \quad F = \begin{pmatrix} . \\ \vdots \\ . \\ . \\ . \\ 1 \end{pmatrix}.$$

La solution de ce système s'écrit encore

$$X^* = \lambda(M^{-1}E^*) + (M^{-1}F).$$

### Illustration

Reprenons les mêmes données que celles utilisées pour la première formulation, à savoir<sup>51</sup>

$$\begin{array}{lll} E_1 = 0,05 & E_2 = 0,08 & E_3 = 0,10 \\ \sigma_1 = 0,10 & \sigma_2 = 0,12 & \sigma_3 = 0,15 \\ \rho_{12} = 0,3 & \rho_{13} = 0,1 & \rho_{23} = 0,4 \end{array}$$

et supposons que les relations de régression ainsi que les variances résiduelles estimées sont données par

<sup>51</sup> Ces valeurs ne sont bien évidemment pas nécessaires pour déterminer les proportions par le modèle de Sharpe (d'ailleurs, une des motivations était d'éviter de devoir calculer la matrice des variances-covariances). Nous les utiliserons ici uniquement pour calculer la frontière efficiente.

$$\begin{aligned}
R_1 &= 0,014 + 0,60R_I & (\sigma_{\varepsilon_1}^2 &= 0,0060) \\
R_2 &= -0,020 + 1,08R_I & (\sigma_{\varepsilon_2}^2 &= 0,0040) \\
R_3 &= 0,200 + 1,32R_I & (\sigma_{\varepsilon_3}^2 &= 0,0012)
\end{aligned}$$

et que le return espéré et la variance de l'indice valent respectivement  $E_I = 0,04$  et  $\sigma_I^2 = 0,0045$ .

Ces données permettent d'écrire

$$M = \begin{pmatrix} 0,0120 & . & . & . & 0,60 & 1 \\ . & 0,0080 & . & . & 1,08 & 1 \\ . & . & 0,0024 & . & 1,32 & 1 \\ . & . & . & 0,0090 & -1 & . \\ 0,60 & 1,08 & 1,32 & -1 & . & . \\ 1 & 1 & 1 & . & . & . \end{pmatrix} \quad E^* = \begin{pmatrix} 0,014 \\ -0,020 \\ 0,200 \\ 0,040 \\ . \\ . \end{pmatrix} \quad F = \begin{pmatrix} . \\ . \\ . \\ . \\ . \\ 1 \end{pmatrix}$$

et donc de calculer

$$M^{-1}E^* = \begin{pmatrix} -7,46 \\ -18,32 \\ 25,79 \\ 9,77 \\ 0,05 \\ 0,07 \end{pmatrix} \quad M^{-1}F = \begin{pmatrix} 0,513 \\ 0,295 \\ 0,192 \\ 0,880 \\ 0,008 \\ -0,011 \end{pmatrix}.$$

Les portefeuilles pour différentes valeurs de  $\lambda$  sont donnés dans le tableau 3.5.

$\lambda$	$X_1$	$X_2$	$X_3$	$E_P$	$\sigma_P$
0,100	-0,2332	-1,5375	2,7706	0,1424	0,3829
0,095	-0,1958	-1,4458	2,6417	0,1387	0,3647
0,090	-0,1585	-1,3542	2,5127	0,1350	0,3465
0,085	-0,1212	-1,2626	2,3838	0,1313	0,3284
0,080	-0,0839	-1,1710	2,2548	0,1276	0,3104
0,075	-0,0465	-1,0793	2,1259	0,1239	0,2924
0,070	-0,0092	-0,9877	1,9969	0,1202	0,2746
0,065	0,0281	-0,8961	1,8680	0,1165	0,2568
0,060	0,0654	-0,8045	1,7390	0,1128	0,2392
0,055	0,1028	-0,7129	1,6101	0,1091	0,2218
0,050	0,1401	-0,6212	1,4812	0,1054	0,2046
0,045	0,1774	-0,5296	1,3522	0,1017	0,1877
0,040	0,2147	-0,4380	1,2233	0,0980	0,1711
0,035	0,2521	-0,3464	1,0943	0,0943	0,1551
0,030	0,2894	-0,2547	0,9654	0,0906	0,1397
0,025	0,3267	-0,1631	0,8364	0,0869	0,1252
0,020	0,3640	-0,0715	0,7075	0,0832	0,1119
0,015	0,4014	0,0201	0,5785	0,0795	0,1003
0,010	0,4387	0,1118	0,4496	0,0758	0,0912
0,005	0,4760	0,2034	0,3206	0,0721	0,0853
0,000	0,5133	0,2950	0,1917	0,0684	0,0832

Tableau 3.5 – Solution pour le modèle à indice simple de Sharpe

La frontière efficiente est représentée dans la figure 3.14.

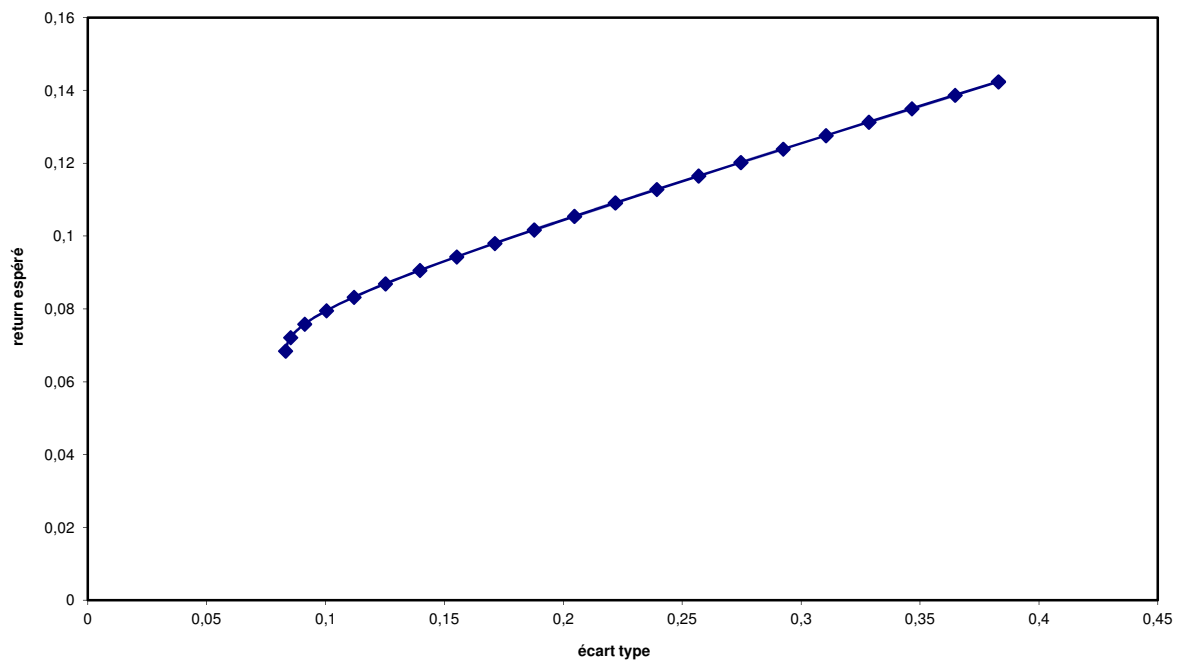


Figure 3.14 – Frontière efficiente pour le modèle à indice simple de Sharpe

Signalons que si la frontière efficiente a la même allure que pour le modèle de Markowitz, il ne faut pas comparer ici les proportions puisque les équations de régression sur lesquelles on s'est appuyé sont arbitraires et ne résultent pas d'une analyse effective de la relation entre les returns des titres et ceux de l'indice. ♣

### *Remarque 1*

Le gain en nombre de données nécessaires par rapport au modèle de Markowitz est considérable : dans ce dernier, les returns espérés, les variances et les covariances deux à deux correspondent à

$$N + N + \frac{N(N-1)}{2} = \frac{N(N+3)}{2},$$

alors que dans le modèle à indice simple, on a besoin des coefficients de régression, des variances résiduelles, ainsi que du return espéré et de la variance de l'indice, à savoir

$$2N + N + 2 = 3N + 2.$$

Par exemple, sur un marché pour lequel on a le choix entre 100 titres, le nombre de données nécessaires est de 5150 dans le premier cas et de seulement 302 dans le second. ♣

### *Remarque 2*

Si, en plus des contraintes envisagées ci-dessus, on impose les contraintes d'inégalités

$$B_j^- \leq X_j \leq B_j^+ \quad j = 1, \dots, N,$$

le modèle à indice simple peut encore s'utiliser en appliquant les mêmes principes que pour le modèle de Markowitz (modification des éléments matriciels suivant le statut « down », « in » et « up » des différents titres, calcul des  $\lambda$  critiques et des portefeuilles-coins).

### **Modèle multi-indices**

Une critique que l'on peut formuler à l'encontre du modèle à indice simple est la représentation du comportement de n'importe quel titre en fonction d'un seul indice. Une manière probablement plus cohérente de procéder est de partager l'ensemble des titres du marché en secteurs et d'exprimer le return de chaque titre d'un même secteur comme fonction du premier degré du return d'un indice sectoriel.

L'écriture générale de ce modèle est quelque peu encombrante ; nous allons l'exprimer dans le cas de deux secteurs, le premier correspondant aux titres  $j = 1, \dots, N_1$  et le second à  $j = N_1 + 1, \dots, N_1 + N_2 = N$ , les indices sectoriels étant notés  $I_1$  et  $I_2$  respectivement.

Les équations de régression prennent la forme

$$\begin{cases} R_{jt} = a_j + b_j R_{I_1 t} + \varepsilon_{jt} & j = 1, \dots, N_1 \\ R_{jt} = a_j + b_j R_{I_2 t} + \varepsilon_{jt} & j = N_1 + 1, \dots, N_1 + N_2 = N \end{cases}$$

Le return du portefeuille, ainsi que son espérance et sa variance, sont donnés par

$$\begin{aligned} R_p &= \sum_{j=1}^N X_j R_j \\ &= \sum_{j=1}^N X_j a_j + Y_1 R_{I_1} + Y_2 R_{I_2} + \sum_{j=1}^N X_j \varepsilon_j \\ E_p &= \sum_{j=1}^N X_j a_j + Y_1 E_{I_1} + Y_2 E_{I_2} \\ \sigma_p^2 &= \sum_{j=1}^N X_j^2 \sigma_{\varepsilon_j}^2 + Y_1^2 \sigma_{I_1}^2 + Y_2^2 \sigma_{I_2}^2 + 2Y_1 Y_2 \sigma_{I_1 I_2} \\ &= (X_1 \quad \dots \quad X_N \quad Y_1 \quad Y_2) \begin{pmatrix} \sigma_{\varepsilon_1}^2 & & & & \\ & \ddots & & & \\ & & \sigma_{\varepsilon_N}^2 & & \\ & & & \sigma_{I_1}^2 & \sigma_{I_1 I_2} \\ & & & \sigma_{I_2 I_1} & \sigma_{I_2}^2 \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_N \\ Y_1 \\ Y_2 \end{pmatrix} \end{aligned}$$

où on a introduit les paramètres

$$\begin{cases} Y_1 = \sum_{j=1}^{N_1} X_j b_j \\ Y_2 = \sum_{j=N_1+1}^N X_j b_j \end{cases}$$

Le raisonnement habituel conduit à nouveau à la relation

$$MX^* = \lambda E^* + F,$$

Qui se résout en

$$X^* = \lambda(M^{-1}E^*) + (M^{-1}F),$$

avec les notations

$$M = \begin{pmatrix} 2\sigma_{\varepsilon_1}^2 & & & & & & & & & b_1 & & 1 \\ & \ddots & & & & & & & & \vdots & & \vdots \\ & & 2\sigma_{\varepsilon_{N_1}}^2 & & & & & & & \vdots & & \vdots \\ & & & 2\sigma_{\varepsilon_{N_1+1}}^2 & & & & & & \vdots & & \vdots \\ & & & & \ddots & & & & & \vdots & & \vdots \\ & & & & & 2\sigma_{\varepsilon_N}^2 & & & & \vdots & & \vdots \\ . & & & & & & 2\sigma_{I_1}^2 & 2\sigma_{I_1 I_2} & -1 & . & . \\ . & \dots & . & . & \dots & . & 2\sigma_{I_2 I_1} & 2\sigma_{I_2}^2 & . & -1 & . \\ . & \dots & . & . & \dots & . & -1 & . & . & . & . \\ b_1 & \dots & b_{N_1} & . & \dots & . & . & . & . & . & . \\ . & \dots & . & b_{N_1+1} & \dots & b_N & . & -1 & . & . & . \\ 1 & \dots & 1 & 1 & \dots & 1 & . & . & . & . & . \end{pmatrix}$$

$$X^* = \begin{pmatrix} X_1 \\ \vdots \\ X_N \\ Y_1 \\ Y_2 \\ m_1 \\ m_2 \\ m_3 \end{pmatrix} \quad E^* = \begin{pmatrix} a_1 \\ \vdots \\ a_N \\ E_{I_1} \\ E_{I_2} \\ . \\ . \\ . \end{pmatrix} \quad F = \begin{pmatrix} . \\ \vdots \\ . \\ . \\ . \\ . \\ . \\ 1 \end{pmatrix}.$$

Remarquons que, par rapport au modèle à indice simple, le modèle à deux indices ne requiert que trois données supplémentaires : un return espéré, une variance et une covariance pour le deuxième indice.

### 3.2.5. Modèle avec titre non risqué

#### *Modélisation et résolution*

Examinons maintenant le cas où le portefeuille est constitué d'un certain nombre  $N$  d'actions (de returns  $R_1, \dots, R_N$ ) en proportions  $X_1, \dots, X_N$ , ainsi que d'un *titre sans risque* dont le rendement est  $R_F$  et qui est en proportion  $X_{N+1}$ , avec

$$X_1 + \dots + X_N + X_{N+1} = 1.$$

Ce titre non risqué s'interprète comme une hypothèse qui se formule comme suit : l'investisseur a la possibilité de placer ou prêter ( $X_{N+1} > 0$ ) ou encore d'emprunter ( $X_{N+1} < 0$ ) des fonds au même taux  $R_F$ .



A côté des returns des actions, qui sont les v.a. que nous avons traitées dans les paragraphes précédents (avec leurs espérances  $E_j$  et leur matrice des variances-covariances  $V$ ), le return du titre sans risque est une v.a. dégénérée :

$$\begin{cases} E_{N+1} = R_F \\ \sigma_{N+1}^2 = 0 \\ \sigma_{j,N+1} = 0 \quad (j = 1, \dots, N) \end{cases}$$

### Remarque

Nous allons étudier l'influence de la présence d'un titre sans risque dans le portefeuille sur base du modèle de Markowitz sans contrainte d'inégalité. On pourrait sans difficulté adapter la présentation au modèle de Sharpe, ou prendre en considération des contraintes d'inégalité.

Le résultat relatif à la forme de la frontière efficiente (voir plus loin) est valable dans tous les cas et une seule présentation suffit. ♣

Le return du portefeuille s'écrit

$$R_p = X_1 R_1 + \dots + X_N R_N + X_{N+1} R_F,$$

ce qui permet de calculer le return espéré et la variance :

$$\begin{cases} E_p = \sum_{j=1}^N X_j E_j + X_{N+1} R_F \\ \sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} \end{cases}$$

Nous devons donc résoudre le problème, pour les différentes valeurs de  $\lambda$  entre 0 et  $+\infty$ , de la minimisation par rapport aux proportions  $X_1, \dots, X_N, X_{N+1}$  de l'expression

$$\sigma_p^2 - \lambda E_p,$$

sous la contrainte

$$\sum_{j=1}^N X_j + X_{N+1} = 1.$$

Le lagrangien du problème peut s'écrire

$$L(X_1, \dots, X_N, X_{N+1}; m) = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} - \lambda \cdot \left( \sum_{j=1}^N X_j E_j + X_{N+1} R_F \right) + m \cdot \left( \sum_{j=1}^N X_j + X_{N+1} - 1 \right).$$

L'annulation de ses dérivées partielles conduit au système d'équations

$$MX^* = \lambda E^* + F,$$

où on a posé

$$M = \begin{pmatrix} 2\sigma_1^2 & 2\sigma_{12} & \cdots & 2\sigma_{1N} & . & 1 \\ 2\sigma_{21} & 2\sigma_2^2 & \cdots & 2\sigma_{2N} & . & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 2\sigma_{N1} & 2\sigma_{N2} & \cdots & 2\sigma_N^2 & . & 1 \\ . & . & \cdots & . & . & 1 \\ 1 & 1 & \cdots & 1 & 1 & . \end{pmatrix} \quad X^* = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \\ X_{N+1} \\ m \end{pmatrix} \quad E^* = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_N \\ R_F \\ . \end{pmatrix} \quad F = \begin{pmatrix} . \\ . \\ . \\ . \\ . \\ 1 \end{pmatrix}.$$

La solution de ce système s'écrit bien entendu

$$X^* = \lambda(M^{-1}E^*) + (M^{-1}F).$$

### Illustration

Reprenons les mêmes données que précédemment, à savoir

$$\begin{array}{lll} E_1 = 0,05 & E_2 = 0,08 & E_3 = 0,10 \\ \sigma_1 = 0,10 & \sigma_2 = 0,12 & \sigma_3 = 0,15 \\ \rho_{12} = 0,3 & \rho_{13} = 0,1 & \rho_{23} = 0,4 \end{array}$$

et supposons que le taux d'intérêt sans risque est  $R_F = 0,03$ .

On a donc

$$M = \begin{pmatrix} 0,0200 & 0,0072 & 0,0030 & . & 1 \\ 0,0072 & 0,0288 & 0,0144 & . & 1 \\ 0,0030 & 0,0144 & 0,0450 & . & 1 \\ . & . & . & . & 1 \\ 1 & 1 & 1 & 1 & . \end{pmatrix} \quad E^* = \begin{pmatrix} 0,05 \\ 0,08 \\ 0,10 \\ 0,03 \\ . \end{pmatrix} \quad F = \begin{pmatrix} . \\ . \\ . \\ . \\ 1 \end{pmatrix}$$

et par conséquent

$$M^{-1}E^* = \begin{pmatrix} 0,452 \\ 1,024 \\ 1,198 \\ -2,674 \\ 0,030 \end{pmatrix} \quad M^{-1}F = \begin{pmatrix} . \\ . \\ . \\ 1 \\ . \end{pmatrix},$$

ce qui conduit aux portefeuilles du tableau 3.6.

$\lambda$	$X_1$	$X_2$	$X_3$	$X(R_F)$	$E_P$	$\sigma_P$
2,0	0,9031	2,0488	2,3953	-4,3472	0,3182	0,5368
1,9	0,8580	1,9464	2,2755	-4,0799	0,3038	0,5100
1,8	0,8128	1,8439	2,1558	-3,8125	0,2894	0,4831
1,7	0,7677	1,7415	2,0360	-3,5451	0,2749	0,4563
1,6	0,7225	1,6390	1,9162	-3,2778	0,2605	0,4295
1,5	0,6774	1,5366	1,7965	-3,0104	0,2461	0,4026
1,4	0,6322	1,4342	1,6767	-2,7431	0,2317	0,3758
1,3	0,5870	1,3317	1,5569	-2,4757	0,2173	0,3489
1,2	0,5419	1,2293	1,4372	-2,2083	0,2029	0,3221
1,1	0,4967	1,1268	1,3174	-1,9410	0,1885	0,2952
1,0	0,4516	1,0244	1,1976	-1,6736	0,1741	0,2684
0,9	0,4064	0,9220	1,0779	-1,4063	0,1597	0,2416
0,8	0,3613	0,8195	0,9581	-1,1389	0,1453	0,2147
0,7	0,3161	0,7171	0,8384	-0,8715	0,1309	0,1879
0,6	0,2709	0,6146	0,7186	-0,6042	0,1165	0,1610
0,5	0,2258	0,5122	0,5988	-0,3368	0,1020	0,1342
0,4	0,1806	0,4098	0,4791	-0,0694	0,0876	0,1074
0,3	0,1355	0,3073	0,3593	0,1979	0,0732	0,0805
0,2	0,0903	0,2049	0,2395	0,4653	0,0588	0,0537
0,1	0,0452	0,1024	0,1198	0,7326	0,0444	0,0268
0,0	0,0000	0,0000	0,0000	1,0000	0,0300	0,0000

Tableau 3.6 – Solution pour le modèle avec titre non risqué

La frontière efficiente est représentée dans la figure 3.15.

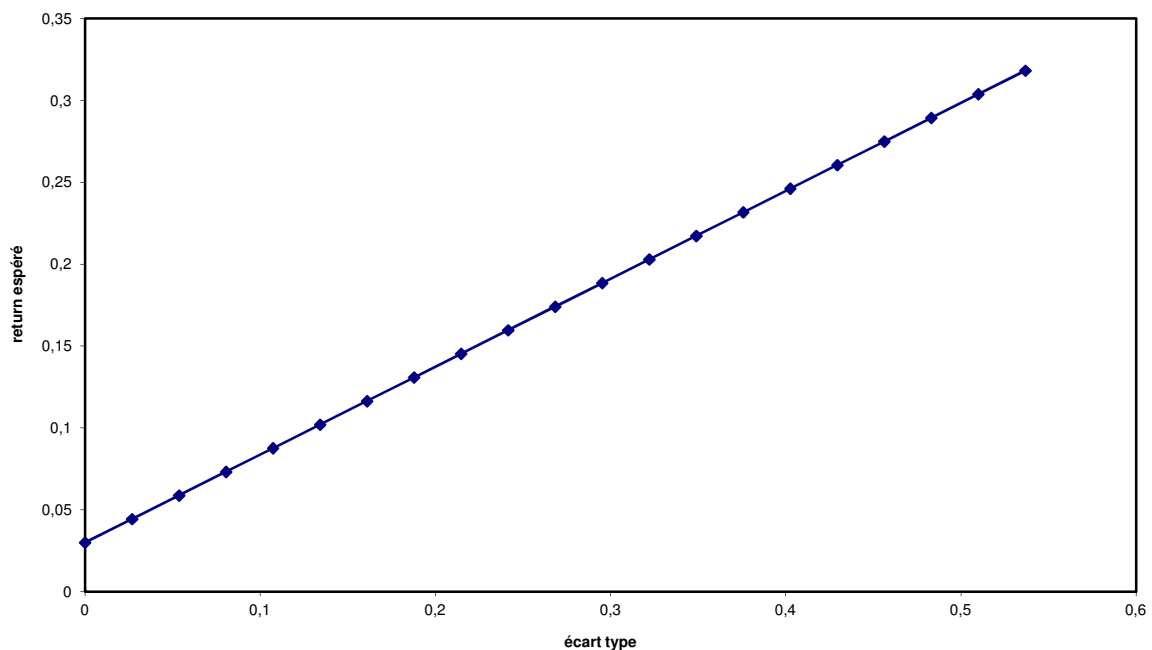


Figure 3.15 – Frontière efficiente pour le modèle avec titre non risqué

Si l'on superpose la frontière efficiente obtenue ci-dessus et celle découlant du modèle de Markowitz (sans titre sans risque), on obtient la figure 3.16.

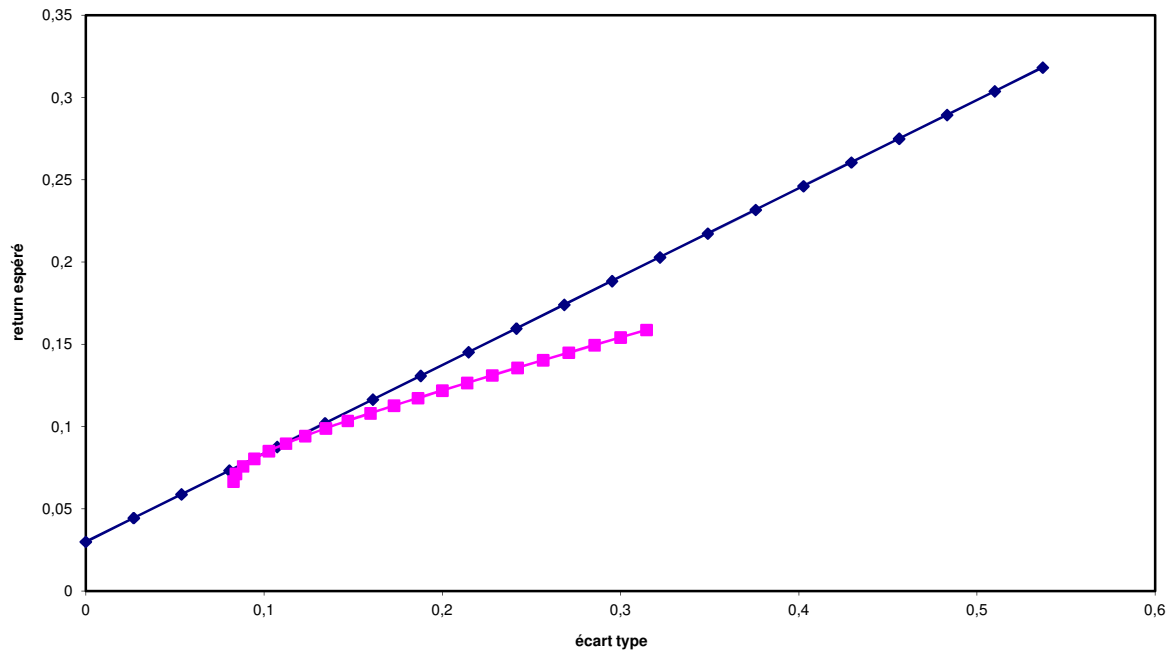


Figure 3.16 – Comparaison des frontières efficientes avec et sans titre non risqué♣

### Frontière efficiente

Le phénomène graphique qui apparaît dans l'illustration précédente est général.

En effet, un portefeuille constitué des  $N$  actions et du titre non risqué peut être considéré comme formé du titre sans risque en proportion  $X = X_{N+1}$  et d'un portefeuille d'actions dont la proportion est  $1 - X$  et le return  $R$  (de paramètres  $E$  et  $\sigma$ ). Le return du titre non risqué a une variance nulle et est non corrélé avec le portefeuille d'actions. Les paramètres du portefeuille sont donnés par

$$\begin{cases} E_p = XR_F + (1 - X)E \\ \sigma_p^2 = (1 - X)^2 \sigma^2 \end{cases}$$

ce qui donne, après élimination de  $X$ ,

$$E_p = R_F \pm \sigma_p \left( \frac{E - R_F}{\sigma} \right),$$

suivant que  $X \leq 1$  ou  $X \geq 1$ . Les équations de ces droites montrent que les portefeuilles en question sont situés sur deux demi-droites de même pente, au signe près (voir figure 3.17).

$$E_p$$

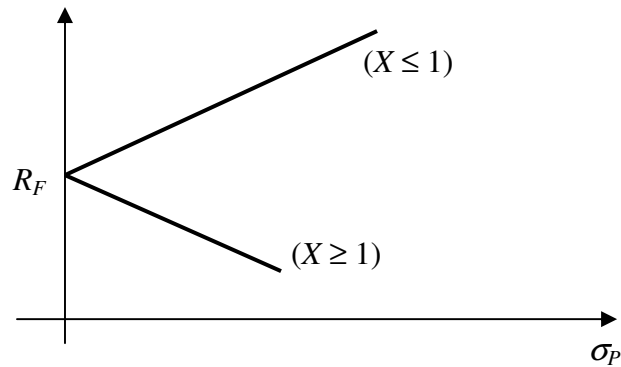


Figure 3.17 – Portefeuilles avec titre non risqué

La demi-droite inférieure ( $X \geq 1$ ) correspond à une situation où le portefeuille d'actions est vendu à découvert afin de placer davantage dans le titre sans risque. On s'intéressera désormais à la partie supérieure.

Si on connaît la frontière efficiente du portefeuille constitué uniquement des actions, la demi-droite optimale – qui maximise  $E_P$  pour  $\sigma_P$  donné – est celle qui est située le plus haut, c'est-à-dire la tangente à la frontière efficiente des actions (voir figure 3.18).

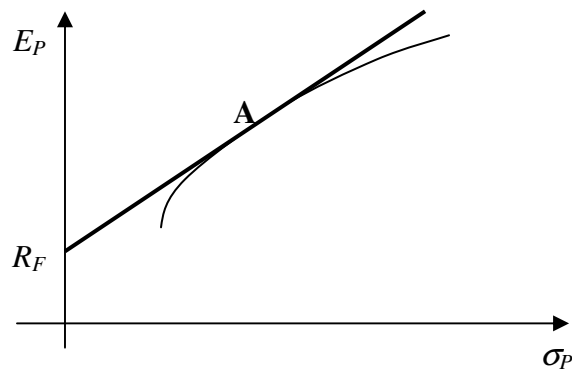


Figure 3.18 – Frontière efficiente en présence d'un titre non risqué

Les portefeuilles situés entre l'axe vertical et le point de contact  $A$  sont caractérisés par  $0 \leq X \leq 1$  et ceux au-delà de  $A$  sont tels que  $X \leq 0$  (emprunt au taux  $R_F$  pour investir davantage dans le portefeuille-contact  $A$ ).

### 3.2.6. Méthode d'Elton, Gruber et Padberg pour la gestion de portefeuille

La méthode<sup>52</sup> d'Elton, Gruber et Padberg (en abrégé méthode E.G.P.) a été développée<sup>53</sup> afin de fournir une résolution rapide et cohérente du problème d'optimisation de portefeuille. Au lieu de déterminer la frontière efficiente comme dans les modèles de Markowitz et Sharpe, cette nouvelle technique détermine simplement le portefeuille qui correspond au point de contact de la tangente à cette frontière efficiente issue du point  $(0, R_F)$ .

#### *Hypothèses*

La méthode en question suppose que

- l'approche moyenne-variance est pertinente ; ce qui permettra d'utiliser un certain nombre de résultats de la théorie de Markowitz ;
- il existe un actif sans risque dont le rendement est noté  $R_F$ .

A côté de ces hypothèses générales, Elton, Gruber et Padberg ont développé des algorithmes de résolution dans deux cas particuliers :

- corrélations constantes : dans ce premier modèle, on suppose que les coefficients de corrélation des returns des différents titres sont tous égaux :  $\rho_{ij} = \rho \quad \forall i, j$  ;
- on peut utiliser le modèle à indice simple de Sharpe.

La première de ces deux simplifications étant relativement « brutale » et assez peu réaliste, nous nous contenterons de développer ici le second cas. Rappelons qu'il repose sur les deux conditions suivantes :

- a) Les returns des différents titres s'expriment comme des fonctions du premier de gré du return d'un indice représentatif du marché :

$$R_{jt} = a_j + b_j R_{It} + \varepsilon_{jt} \quad j = 1, \dots, N.$$

On suppose également que les résidus vérifient les hypothèses classiques de la régression linéaire, entre autres que les résidus sont d'espérance nulle et qu'ils sont non corrélés avec la variable explicative  $R_{It}$ .

---

<sup>52</sup> Ou plus précisément les méthodes ; en effet, à côté d'une idée générale, divers modèles ont été développés suivant les hypothèses formulées.

<sup>53</sup> ELTON E., GRUBER M et PADBERG M., Simple criteria for optimal portfolio selection, *Journal of Finance*, Vol. XI, n° 5, 1976, pp. 1341-1357.

ELTON E., GRUBER M et PADBERG M., Optimal portfolios from single ranking devices, *Journal of portfolio management*, Vol. 4, n° 3, 1978, pp. 15-19.

ELTON E., GRUBER M et PADBERG M., Simple criteria for optimal portfolio selection : tracing out the efficient frontier, *Journal of Finance*, Vol. XIII, n° 1, 1978, pp. 296-302.

ELTON E., GRUBER M et PADBERG M., Simple criteria for optimal portfolio selection with upper bounds, *Operation Research*, 1978.

On consultera également avec intérêt ELTON E. et GRUBER M., *Modern portfolio theory and investment analysis*, Wiley, New-York, 1991.

- b) Les résidus des régressions relatives aux différents titres sont non corrélés :  $\text{cov}(\varepsilon_{it}, \varepsilon_{jt}) = 0$  pour tout  $i$  et  $j$  différents.

### **Résolution du cas où les ventes à découvert sont autorisées**

Nous analyserons tout d'abord en détail le cas où aucune contrainte d'inégalité n'est imposée aux proportions. Le raisonnement est ici plus simple<sup>54</sup> que dans le cas où les ventes à découvert sont interdites. Néanmoins, comme on le verra (mais sans démonstration), la mise en œuvre de l'algorithme est à peine plus complexe dans ce second cas.

Si on considère un portefeuille  $P$  constitué uniquement d'actions dans les proportions  $X_1, X_2, \dots, X_N$ , alors l'ensemble des portefeuilles composés en partie de  $P$  et pour le reste du taux sans risque  $R_F$  constitue la droite reliant les points  $(0, R_F)$  et  $(\sigma_P, E_P)$ , comme illustré sur la figure 3.19.

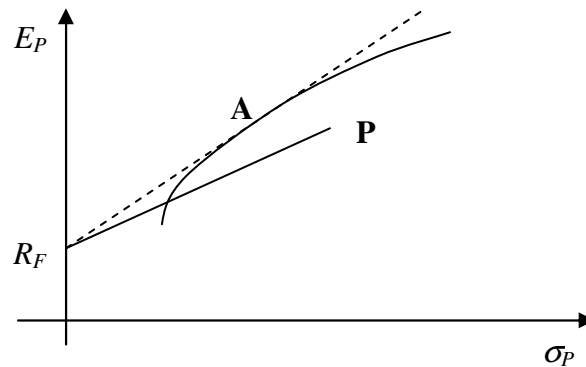


Figure 3.19 – Méthode E.G.P.

La pente de la droite en question est donnée par

$$\Psi_P = \frac{E_P - R_F}{\sigma_P},$$

qui peut s'interpréter comme une prime de risque, comme on le verra au § 3.3.1.

D'après les raisonnements tenus au paragraphe précédent, le portefeuille  $P$  idéal correspond au point de contact  $A$  de la tangente à la frontière efficiente, issue du point  $(0, R_F)$ , pour lequel la pente est maximale. On est donc à la recherche des proportions qui maximisent la pente  $\Psi_P$  ou, ce qui revient au même, qui maximisent  $\Psi_P^2$ .

Comme

<sup>54</sup> Il commence d'ailleurs comme la démonstration de l'équation du C.A.P.M. (voir § 3.3.1).

$$\begin{cases} E_P - R_F = \sum_{j=1}^N X_j E_j - \left( \sum_{j=1}^N X_j \right) R_F = \sum_{j=1}^N X_j (E_j - R_F) \\ \sigma_P^2 = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} \end{cases}$$

la dérivée de

$$\Psi_P^2 = \frac{(E_P - R_F)^2}{\sigma_P^2} = \frac{\left( \sum_{j=1}^N X_j (E_j - R_F) \right)^2}{\sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij}},$$

par rapport à  $X_k$  est donnée par

$$\begin{aligned} (\Psi_P^2)'_{X_k} &= \frac{2 \left( \sum_{j=1}^N X_j (E_j - R_F) \right) (E_k - R_F) \cdot \sigma_P^2 - \left( \sum_{j=1}^N X_j (E_j - R_F) \right)^2 \cdot 2 \sum_{j=1}^N X_j \sigma_{kj}}{\sigma_P^4} \\ &= \frac{2(E_P - R_F)(E_k - R_F) \sigma_P^2 - 2(E_P - R_F)^2 \sum_{j=1}^N X_j \sigma_{kj}}{\sigma_P^4} \\ &= \frac{2(E_P - R_F)}{\sigma_P^2} \cdot \left( (E_k - R_F) - \gamma \cdot \sum_{j=1}^N X_j \sigma_{kj} \right) \end{aligned}$$

où on a posé provisoirement  $\gamma = (E_P - R_F) / \sigma_P^2$ .

Cette dérivée est nulle si

$$E_k - R_F = \gamma \cdot \sum_{j=1}^N X_j \sigma_{kj}.$$

En introduisant  $Z_j = \gamma X_j$  ( $j = 1, \dots, N$ ), le système à résoudre, par rapport à  $Z_1, \dots, Z_N$  est donc

$$E_k - R_F = \sum_{j=1}^N Z_j \sigma_{kj} \quad k = 1, \dots, N.$$

Avant de passer à sa résolution, remarquons que trouver les  $Z_k$  permet d'obtenir les  $X_k$  puisque

$$X_k = \frac{Z_k}{\gamma} = \frac{Z_k}{\gamma \cdot \sum_{j=1}^N X_j} = \frac{Z_k}{\sum_{j=1}^N Z_j}.$$



Les hypothèses du modèle de Sharpe permettent d'écrire

$$\begin{aligned}\sigma_{kj} &= \text{cov}(a_k + b_k R_I + \varepsilon_k, a_j + b_j R_I + \varepsilon_j) \\ &= b_k b_j \sigma_I^2 + \begin{cases} \sigma_{\varepsilon_k}^2 & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases}\end{aligned}$$

La  $k$ -ème équation du système peut dès lors s'écrire

$$E_k - R_F = b_k \left( \sum_{j=1}^N Z_j b_j \right) \sigma_I^2 + Z_k \sigma_{\varepsilon_k}^2$$

ou encore, en résolvant par rapport à  $Z_k$ ,

$$\begin{aligned}Z_k &= \frac{1}{\sigma_{\varepsilon_k}^2} \left\{ (E_k - R_F) - b_k \left( \sum_{j=1}^N Z_j b_j \right) \sigma_I^2 \right\} \\ &= \frac{b_k}{\sigma_{\varepsilon_k}^2} \left\{ \theta_k - \left( \sum_{j=1}^N Z_j b_j \right) \sigma_I^2 \right\}\end{aligned}$$

où on a posé

$$\theta_k = \frac{E_k - R_F}{b_k}.$$

Il reste maintenant à déterminer la somme entre parenthèses ; à partir du dernier résultat, on trouve

$$\begin{aligned}\sum_{k=1}^N Z_k b_k &= \sum_{k=1}^N \frac{b_k^2}{\sigma_{\varepsilon_k}^2} \left\{ \theta_k - \left( \sum_{j=1}^N Z_j b_j \right) \sigma_I^2 \right\} \\ &= \sum_{k=1}^N \frac{b_k^2}{\sigma_{\varepsilon_k}^2} \theta_k - \left( \sum_{k=1}^N \frac{b_k^2}{\sigma_{\varepsilon_k}^2} \right) \left( \sum_{j=1}^N Z_j b_j \right) \sigma_I^2\end{aligned}$$

dont la résolution donne

$$\sum_{j=1}^N Z_j b_j = \frac{\sum_{k=1}^N \frac{b_k^2}{\sigma_{\varepsilon_k}^2} \theta_k}{1 + \left( \sum_{k=1}^N \frac{b_k^2}{\sigma_{\varepsilon_k}^2} \right) \sigma_I^2}.$$

En introduisant la nouvelle notation

$$\phi = \left( \sum_{j=1}^N Z_j b_j \right) \cdot \sigma_I^2 = \frac{\sum_{k=1}^N \frac{b_k^2}{\sigma_{\varepsilon_k}^2} \theta_k}{1 + \left( \sum_{k=1}^N \frac{b_k^2}{\sigma_{\varepsilon_k}^2} \right) \sigma_I^2} \cdot \sigma_I^2$$

et en substituant la somme qui vient d'être calculée dans l'expression de  $Z_k$ , on trouve

$$Z_k = \frac{b_k}{\sigma_{\varepsilon_k}^2} (\theta_k - \phi) \quad k = 1, \dots, N.$$

### Illustration

Reprenons les mêmes données que celles utilisées pour le modèle à indice simple, à savoir (on cite seulement les données nécessaires ici)

$$E_1 = 0,05 \quad E_2 = 0,08 \quad E_3 = 0,10$$

avec les relations de régression ainsi que les variances résiduelles estimées

$$\begin{aligned} R_1 &= 0,014 + 0,60R_I & (\sigma_{\varepsilon_1}^2 &= 0,0060) \\ R_2 &= -0,020 + 1,08R_I & (\sigma_{\varepsilon_2}^2 &= 0,0040) \\ R_3 &= 0,200 + 1,32R_I & (\sigma_{\varepsilon_3}^2 &= 0,0012) \end{aligned}$$

et que la variance de l'indice vaut  $\sigma_I^2 = 0,0045$  ; supposons enfin, comme pour le modèle avec titre sans risque que ce dernier est égal à  $R_F = 0,03$ .

Ces données permettent de calculer

$$\theta_1 = 0,0333 \quad \theta_2 = 0,0741 \quad \theta_3 = 0,0758$$

et par conséquent  $\phi = 0,0659$ . On en déduit la valeur des  $Z_k$  :

$$Z_1 = -3,2615 \quad Z_2 = 2,1939 \quad Z_3 = 10,7899$$

et par conséquent les proportions du portefeuille optimal :

$$X_1 = -0,3355 \quad X_2 = 0,2257 \quad X_3 = 1,1098. \clubsuit$$

### Résolution du cas où les ventes à découvert sont interdites

Examinons maintenant le cas où on ajoute des contraintes. Celles-ci sont moins générales que celles envisagées dans le modèle de Markowitz : elles s'écrivent simplement  $0 \leq X_j \leq 1$  ( $j = 1, \dots, N$ ).

La méthode, que nous livrons ici sans justification, est très semblable à celle utilisée pour le cas où les ventes à découvert sont autorisées. Comme ci-dessus, on calcule les

$$\theta_k = \frac{E_k - R_F}{b_k} \quad k = 1, \dots, N$$

et on trie alors les titres par ordre de  $\theta_k$  décroissants et on conserve cet ordre jusqu'à la fin de l'algorithme. Au lieu d'avoir un seul paramètre  $\phi$ , on en calcule un par titre :

$$\phi_k = \frac{\sum_{j=1}^k \frac{b_j^2}{\sigma_{\varepsilon_j}^2} \theta_j}{1 + \left( \sum_{j=1}^k \frac{b_j^2}{\sigma_{\varepsilon_j}^2} \right) \sigma_I^2} \cdot \sigma_I^2 \quad k = 1, \dots, N.$$

Il est possible de montrer que la suite des nombres  $\phi_k$  est d'abord croissante, passe par un maximum, pour finir par une phase décroissante. On note  $K$  la valeur de l'indice  $k$  correspondant au  $\phi_k$  maximal. Le nombre  $\phi_K$  porte le nom de *seuil d'acceptabilité* (en anglais « *cut-off rate* ») et on peut montrer que le calcul des  $Z_k$  par la même relation que précédemment<sup>55</sup> conduit à des valeurs positives pour  $k = 1, \dots, K$  et à des valeurs négatives pour  $k = K + 1, \dots, N$ . Seuls les  $K$  premiers titres seront intégrés au portefeuille. Les calculs à réaliser sont donc

$$Z_k = \frac{b_k}{\sigma_{\varepsilon_k}^2} (\theta_k - \phi_K) \quad k = 1, \dots, N,$$

ce qui donne, pour les proportions des  $K$  titres intégrés,

$$X_k = \frac{Z_k}{\sum_{j=1}^K Z_j} \quad k = 1, \dots, K.$$

### Illustration

Reprenons les mêmes données que ci-dessus. On a bien sûr toujours

$$\theta_1 = 0,0333 \quad \theta_2 = 0,0741 \quad \theta_3 = 0,0758$$

ce qui permet de classer les titres dans l'ordre (3), (2), (1). Nous renumérotions provisoirement les titres dans ce nouvel ordre. On trouve alors

$$\phi_1 = 0,0657 \quad \phi_2 = 0,0669 \quad \phi_3 = 0,0659$$

<sup>55</sup> En remplaçant  $\phi$  par  $\phi_K$ .

ce qui montre que  $K = 2$  et que le seuil d'acceptabilité est  $\phi_2 = 0,0669$ . On en déduit la valeur des  $Z_k$  :

$$Z_1 = 9,6949 \quad Z_2 = 1,9251$$

et par conséquent les proportions du portefeuille optimal :

$$X_1 = 0,8343 \quad X_2 = 0,1657.$$

Si l'on revient à l'ordre de départ, les titres à inclure dans le portefeuille sont donc les titres (2) et (3) avec pour proportions relatives :

$$X_2 = 0,1657 \quad X_3 = 0,8343. \clubsuit$$

### 3.2.7. Théorie de l'utilité et sélection du portefeuille optimal

Une fois la frontière efficiente déterminée, se pose la question pour un investisseur donné, de sélectionner parmi tous les portefeuilles efficients celui qui lui convient le mieux. Le portefeuille choisi sera différent d'un investisseur à l'autre, le choix effectué dépendant de sa psychologie, de son attitude par rapport au risque ; la frontière efficiente contient en effet aussi bien des portefeuilles prudents (return espéré et risque faibles, situés à l'extrémité gauche de la courbe) que des portefeuilles plus hasardeux (return espéré et risque élevés, situés vers l'extrémité droite).

#### *Fonction d'utilité*

La notion de *fonction d'utilité* peut s'introduire de manière tout à fait générale<sup>56</sup> pour représenter, pour un individu, l'utilité, l'intérêt qu'il trouve dans un projet, un investissement, une stratégie, ... , ces éléments présentant un certain niveau de risque. Les valeurs numériques de cette fonction d'utilité sont peu importantes, cette dernière étant essentiellement utilisée pour comparer des projets, des investissements, des stratégies, ... Nous présenterons ici la théorie de l'utilité dans le cas où on l'applique à un rendement (aléatoire, rappelons-le), celui d'un portefeuille d'actions par exemple.

En raison de la présence du risque, il est bien clair qu'on ne peut se limiter de prendre  $E(R)$  comme utilité du return  $U(R)$ . Ceci a été clairement mis en évidence par D. Bernoulli en 1732 grâce au *paradoxe de Saint-Pétersbourg*. La question est : « quelle somme seriez-vous prêt à miser pour participer au jeu suivant ? Je réalise une suite de lancers d'une pièce de monnaie et je vous donne 2 € si pile apparaît pour la première fois au premier lancer, 4 € si pile apparaît pour la première fois au deuxième lancer, 8 € si pile apparaît pour la première fois au troisième lancer, et ainsi de suite. Je vous donnerai donc  $2^n$  € si pile apparaît pour la première fois au  $n$ -ème lancer ». La plupart accepteront de miser une petite somme (au moins 2 € !),

<sup>56</sup> On trouvera une excellente présentation des notions générales sur l'attitude face au risque (non nécessairement financier) et à la notion d'utilité dans EECKHOUDT L. et GOLLIER C., *Risk*, Harvester Wheatsheaf, 1995.

mais répugneront à investir un montant important en raison du risque élevé de ce jeu : en misant 20 €, le joueur sera perdant avec probabilité

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16} = 0,9375$$

et ne gagnera donc qu'avec 6,25 chances sur 100. Cependant, le gain moyen de ce jeu est égal à

$$\sum_{n=1}^{\infty} 2^n \left(\frac{1}{2}\right)^n = 1 + 1 + 1 + \dots = \infty.$$

C'est l'aversion vis-à-vis du risque qui justifie la décision du joueur. La fonction d'utilité a pour objet de représenter ce comportement.

Dans la théorie de l'utilité, on compare des projets, des investissements, des stratégies, ... (dans notre cas des rendements) par une relation de *préférence* ( $R_1$  est préféré à  $R_2$  :  $R_1 \succ R_2$ ) et une relation d'*indifférence* (indifférence entre  $R_1$  et  $R_2$  :  $R_1 \sim R_2$ ). Le comportement de l'investisseur peut se modéliser si ces deux relations obéissent aux axiomes suivants :

- (comparabilité) : l'investisseur peut toujours comparer deux rendements

$$\forall R_1, R_2, \text{ on a toujours } R_1 \succ R_2, \quad R_1 \prec R_2 \quad \text{ou} \quad R_1 \sim R_2.$$

- (réflexivité) :  $\forall R, \quad R \sim R$ .
- (transitivité) :  $\forall R_1, R_2, R_3, \quad \text{si } R_1 \succ R_2 \text{ et } R_2 \succ R_3 \text{ alors } R_1 \succ R_3$ .
- (continuité) :  $\forall R_1, R_2, R_3, \quad \text{si } R_1 \succ R_2 \succ R_3, \quad \text{il existe un unique } X \in [0 ; 1] \text{ tel que}$

$$[X \cdot R_1 + (1 - X) \cdot R_3] \sim R_2.$$

- (indépendance) :  $\forall R_1, R_2, R_3 \quad \text{et} \quad \forall X \in [0 ; 1], \quad \text{si } R_1 \succ R_2, \quad \text{alors}$

$$[X \cdot R_1 + (1 - X) \cdot R_3] \succ [X \cdot R_2 + (1 - X) \cdot R_3].$$

Von Neumann et Mognstern<sup>57</sup>, ont démontré le *théorème de l'utilité espérée*, qui exprime que si les préférences d'un investisseur obéissent aux axiomes ci-dessus, il existe une fonction  $U$  telle que

$$\forall R_1, R_2, \quad R_1 \succ R_2 \Leftrightarrow E[U(R_1)] > E[U(R_2)].$$

Cette fonction d'utilité est manifestement une fonction croissante. Nous avons fait remarquer que ses valeurs numériques ne sont pas essentielles puisqu'elle n'est utilisée que pour réaliser des comparaisons de rendements. Le théorème de l'utilité espérée permet de préciser cette notion : si les préférences d'un investisseur sont modélisées par la fonction d'utilité  $U$ , on

<sup>57</sup> VON NEUMANN J., MORGENSTERN O., *Theory of games and economic behavior*, Princeton University Press, 1947.

aura le même système de préférences à partir de la fonction  $aU + b$  avec  $a > 0$ . En effet, si  $R_1 \succ R_2$  se traduit par  $E[U(R_1)] > E[U(R_2)]$ , on a

$$\begin{aligned} E[U^*(R_1)] &= E[aU(R_1) + b] \\ &= aE[U(R_1)] + b \\ &> aE[U(R_2)] + b \\ &= E[aU(R_2) + b] \\ &= E[U^*(R_2)] \end{aligned}$$

La fonction d'utilité est un élément associé intrinsèquement à chaque investisseur (et d'ailleurs susceptible d'évoluer au cours du temps et en fonction des circonstances). Il n'est guère facile, ni même utile, de connaître cette fonction. Si l'on voulait en obtenir une estimation approximative, il faudrait définir une liste de valeurs possibles  $R_1 < R_2 < \dots < R_n$  du rendement, puis, pour  $i = 2, \dots, n - 1$ , demander à l'investisseur quelle est pour lui la probabilité  $p_i$  pour laquelle il est indifférent d'obtenir un rendement  $R_i$  de manière certaine ou de jouer à une loterie qui donne les rendements  $R_1$  et  $R_n$  avec les probabilités respectives  $(1 - p_i)$  et  $p_i$ . Si on choisit arbitrairement  $U(R_1) = 0$  et  $U(R_n) = 100$ , alors  $U(R_i) = 100 p_i$  ( $i = 2, \dots, n - 1$ ).

### *Attitude face au risque*

Pour la plupart des investisseurs, une augmentation de rendement de 0,5 % est considérée comme plus intéressante si le rendement actuel est de 2 % que s'il vaut 5 %. Un tel comportement est une *aversion pour le risque*. L'attitude opposée est un *goût pour le risque*, la situation intermédiaire étant la *neutralité vis-à-vis du risque*. Comment ces comportements se traduisent-ils au niveau de la fonction d'utilité ?

Raisonnons par exemple pour le cas de l'aversion. D'une manière plus générale, si on désire exprimer que l'utilité du rendement  $U(R)$  doit croître avec  $R$  en donnant un poids moindre à des mêmes variations de rendement lorsque le niveau de rendement est important, on aura

$$R_1 < R_2 \quad \Rightarrow \quad U(R_1 + \Delta R) - U(R_1) > U(R_2 + \Delta R) - U(R_2),$$

qui traduit le caractère décroissant de l'*utilité marginale*. Dans ce cas, la dérivée de la fonction d'utilité est une fonction décroissante et la dérivée seconde est donc négative : la fonction d'utilité est concave.

Les résultats que l'on déduit de ces considérations sont résumés dans le tableau 3.7 et une représentation de la fonction d'utilité dans ces différents cas est donnée par la figure 3.20.

	<i>utilité marginale</i>	$U''$	$U$
<i>aversion pour le risque</i>	décroissante	$< 0$	concave
<i>neutralité vis-à-vis du risque</i>	constante	$= 0$	linéaire
<i>goût du risque</i>	croissante	$> 0$	convexe

*Tableau 3.7 – Attitude face au risque*

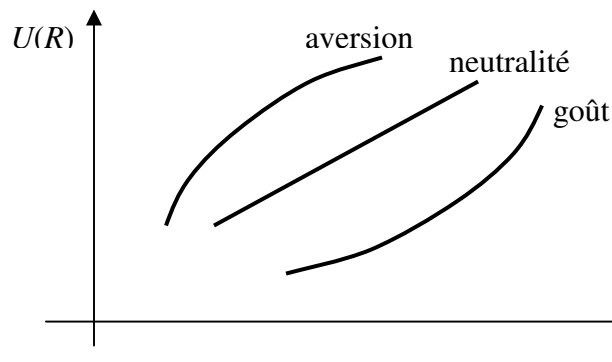


Figure 3.20 – Fonction d'utilité

Précisons quelque peu ce concept. Considérons un investisseur ayant le choix entre, d'une part, un rendement certain égal à  $R$  et, d'autre part, une loterie lui donnant un rendement aléatoire pouvant prendre deux valeurs  $(R - r)$  et  $(R + r)$  avec probabilité  $\frac{1}{2}$  chacune. S'il a de l'aversion pour le risque, l'utilité du return certain sera supérieure à l'utilité espérée du return de la loterie :

$$U(R) > \frac{1}{2}[U(R-r) + U(R+r)],$$

ce qui se traduit graphiquement par la figure 3.21.

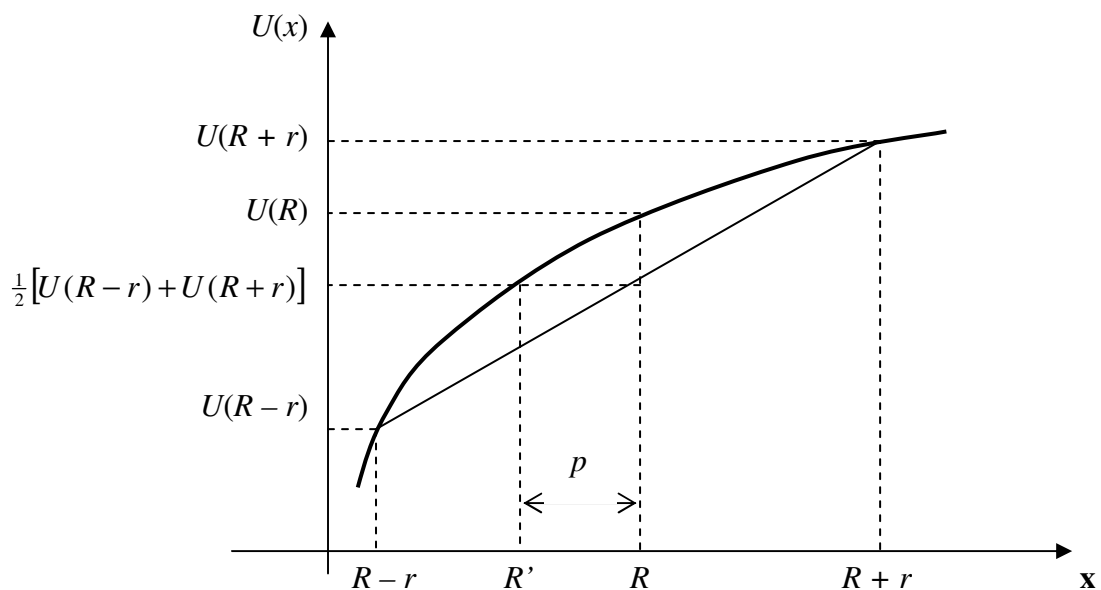


Figure 3.21 – Aversion pour le risque

Sur cette figure, on a noté  $R'$  le rendement certain dont l'utilité est égale au rendement espéré de la loterie. La différence

$$p = R - R',$$

représente le prix que l'investisseur est prêt à payer pour éviter de devoir participer à la loterie : c'est la *prime de risque*.

Des développements de Taylor pour  $U(R + r)$ ,  $U(R - r)$  et  $U(R') = U(R - p)$  conduisent aisément à la relation

$$p = -\frac{U''(R)}{U'(R)} \cdot \frac{r^2}{2}.$$

Le premier facteur de cette expression est le *coefficient d'aversion absolue au risque* :

$$\alpha(R) = -\frac{U''(R)}{U'(R)}.$$

Les deux exemples les plus utilisés de fonction d'utilité correspondant à l'aversion pour le risque sont la fonction exponentielle et la fonction quadratique.

Si  $U(R) = a.e^{bR}$ , avec  $a$  et  $b < 0$ , on a  $\alpha(R) = -b$ .

Si  $U(R) = aR^2 + bR + c$ , avec  $a < 0$  et  $b > 0$ , on doit bien sûr se limiter à des valeurs de  $R$  ne dépassant pas  $-b/2a$  pour que la fonction d'utilité reste croissante. Le coefficient d'aversion absolue au risque est ici donné par

$$\alpha(R) = \frac{1}{-\frac{b}{2a} - R}.$$

Lorsque cette dernière forme pour la fonction d'utilité peut être admise, on dispose d'une autre justification de la caractérisation de la distribution des returns par les deux seuls paramètres moyenne et variance, sans ajouter d'hypothèse de normalité (voir §3.1.1). En effet, dans ce cas, l'utilité espérée du return d'un portefeuille (quantité que l'investisseur désire optimiser) est donnée par

$$\begin{aligned} E[U(R_p)] &= E[aR_p^2 + bR_p + c] \\ &= aE(R_p^2) + bE(R_p) + c \\ &= a(\sigma_p^2 + E_p^2) + bE_p + c \end{aligned}$$

et cette quantité ne dépend que des deux premiers moments.

### ***Sélection du portefeuille optimal***

Considérons maintenant l'investisseur présentant de l'aversion pour le risque et qui est amené à choisir un portefeuille parmi ceux de la frontière efficiente.

On commence par construire les *courbes d'indifférence* relatives à sa fonction d'utilité, c'est-à-dire les courbes correspondant aux couples (espérance ; écart type) pour lesquels l'utilité espérée du return est égale à une valeur donnée (voir figure 3.22). Ces courbes d'indifférence ont une forme convexe emboîtée, l'utilité croissant au fur et à mesure que la courbe d'indifférence se déplace vers le haut et la gauche.



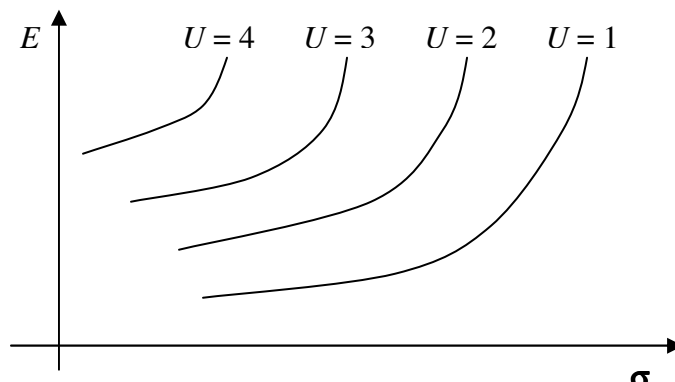


Figure 3.22 – Courbes d'indifférence

En superposant les courbes d'indifférence et la frontière efficiente, on détermine aisément, comme indiqué sur la figure 3.23, le portefeuille efficient  $P$  qui correspond à l'utilité espérée maximale.

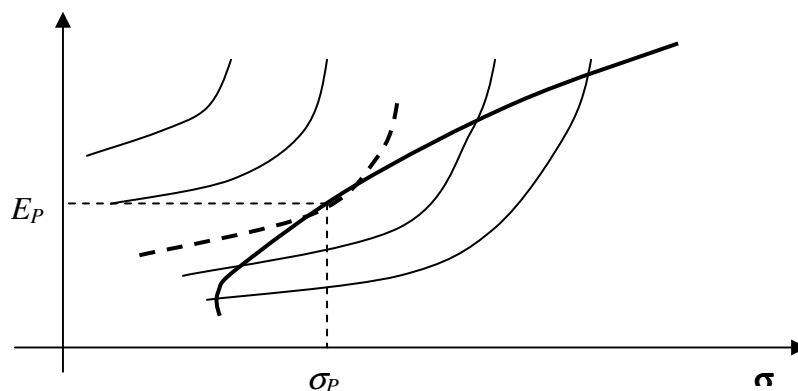


Figure 3.23 – Sélection du portefeuille optimal

### Autres points de vue

A côté du portefeuille efficient sur base du système de préférences de l'investisseur, modélisé par une fonction d'utilité, d'autres objectifs ou contraintes peuvent être pris en considération.

Examinons par exemple le cas de la *contrainte de déficit*. En plus de l'optimisation du couple  $(E, \sigma)$ , c'est-à-dire de la détermination de la frontière efficiente, et avant la sélection du portefeuille (par la fonction d'utilité), on impose ici que le return du portefeuille ne soit inférieur à un seuil<sup>58</sup> fixé  $u$  qu'avec une très faible probabilité,  $p$  disons :

$$\Pr[R_p \leq u] \leq p.$$

Dans le cas où l'hypothèse de normalité des returns est admise, on a

<sup>58</sup> Si  $u = 0$ , cette contrainte signifie que, sauf événement de faible probabilité, le capital investi doit être au moins maintenu.

$$\Pr \left[ \frac{R_p - E_p}{\sigma_p} \leq \frac{u - E_p}{\sigma_p} \right] \leq p ,$$

c'est-à-dire

$$\frac{u - E_p}{\sigma_p} \leq z_p ,$$

où  $z_p$  est le  $p$ -quantile de la distribution normale réduite ( $z_p < 0$  car  $p$  est inférieur à  $1/2$ ). La condition peut donc s'écrire

$$E_p \geq u - z_p \cdot \sigma_p .$$

Les portefeuilles obéissant à la contrainte de déficit sont situés au-dessus de la droite d'équation  $E_p = u - z_p \cdot \sigma_p$  (voir figure 3.24).

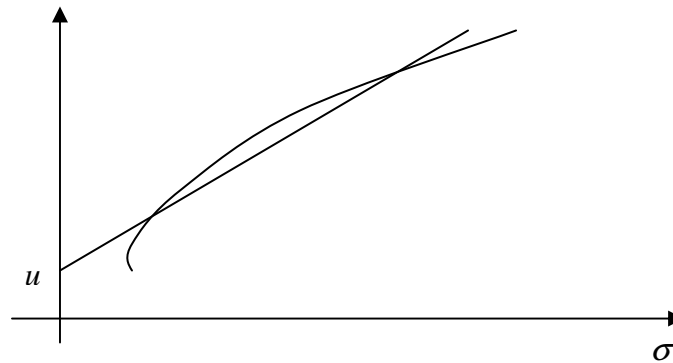


Figure 3.24 – Contrainte de déficit

La portion de la frontière efficiente délimitée par cette droite de contrainte constitue l'ensemble des portefeuilles parmi lesquels la sélection de l'investisseur s'effectuera.

Pour  $p$  fixé, une augmentation de  $u$  (rendement exigé plus élevé) provoque un déplacement de la droite de contrainte vers le haut ; de même, pour  $u$  fixé, une diminution de  $p$  (plus grande sécurité de respecter la contrainte) provoque un déplacement de cette droite vers le haut en pivotant autour du point  $(0, u)$ . Dans les deux cas, le tronçon de frontière efficiente obéissant à la contrainte se restreint.

D'ailleurs, profitant de ces propriétés, on peut déterminer le portefeuille optimal suivant l'un des deux critères en utilisant la droite tangente à la frontière efficiente.

### 3.2.8. Le modèle de marché

Certains développements dans le modèle de marché reprennent des raisonnements tenus dans la construction du modèle de Sharpe, où l'indice est remplacé par le marché dans sa totalité.

Ce modèle constitue cependant davantage une réflexion macro-économique plutôt qu'une recherche de portefeuilles efficients.

### ***Risque systématique et risque spécifique***

On a déjà rencontré (§ 3.1.1) la notion de risque systématique d'un titre

$$\beta_j = \frac{\sigma_{jM}}{\sigma_M^2},$$

qui mesure l'importance du risque du titre ( $j$ ) par rapport au risque du titre moyen du marché. Il apparaît comme coefficient de régression lorsque l'on exprime le return de ce titre en fonction linéaire du return du marché :

$$R_{jt} = \alpha_j + \beta_j R_{Mt} + \varepsilon_{jt}.$$

On suppose bien entendu que les résidus vérifient les hypothèses classiques de la régression linéaire, entre autres que les résidus sont d'espérance nulle, de variance constante et qu'ils sont non corrélés avec la variable explicative  $R_{Mt}$ .

A côté du *risque systématique*  $\beta_j$ , qui est le même pour toute époque, une autre source de fluctuation de  $R_j$  est le résidu  $\varepsilon_{jt}$  qui est spécifique à la période  $t$ . On appellera *risque spécifique* la variance des résidus

$$\sigma_{\varepsilon_j}^2 = \text{var}(\varepsilon_{jt}).$$

### ***Remarque***

En pratique, les coefficients  $\alpha_j$  et  $\beta_j$  de la régression sont estimés par la méthode des moindres carrés ; par exemple,  $\hat{\beta}_j = s_{jM} / s_M^2$ . On estime alors les résidus par

$$\hat{\varepsilon}_{jt} = R_{jt} - (\hat{\alpha}_j + \hat{\beta}_j R_{Mt})$$

et le risque spécifique est estimé par son estimateur ergodique

$$\frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_{jt}^2.$$

Dans la suite de ce paragraphe, nous omettrons l'indice  $t$  relatif au temps. ♣

Nous allons voir comment le risque  $\sigma_j^2$  d'un titre est constitué d'une composante systématique et d'une composante spécifique. On a

$$\begin{aligned}
\sigma_j^2 &= \text{var}(R_j) \\
&= E[(\alpha_j + \beta_j R_M + \varepsilon_j - E(\alpha_j + \beta_j R_M + \varepsilon_j))^2] \\
&= E[(\beta_j (R_M - E_M) + \varepsilon_j)^2] \\
&= \beta_j^2 E[(R_M - E_M)^2] + E(\varepsilon_j^2) + 2\beta_j E[(R_M - E_M)\varepsilon_j] \\
&= \beta_j^2 \text{var}(R_M) + \text{var}(\varepsilon_j)
\end{aligned}$$

d'où la relation de décomposition annoncée :

$$\sigma_j^2 = \beta_j^2 \sigma_M^2 + \sigma_{\varepsilon_j}^2.$$

### ***Bêta d'un portefeuille***

En utilisant l'expression de régression pour  $R_j$ , on développe aisément  $R_P$  :

$$\begin{aligned}
R_P &= \sum_{j=1}^N X_j R_j \\
&= \sum_{j=1}^N X_j (\alpha_j + \beta_j R_M + \varepsilon_j) \\
&= \sum_{j=1}^N X_j \alpha_j + \left( \sum_{j=1}^N X_j \beta_j \right) R_M + \sum_{j=1}^N X_j \varepsilon_j
\end{aligned}$$

ce qui montre que, comme pour le return, le bêta du portefeuille est la moyenne des bêtas des titres constitutifs, pondérée par les proportions exprimées en terme de capitalisation boursière :

$$\beta_P = \sum_{j=1}^N X_j \beta_j.$$

### ***Lien entre le modèle de marché et la diversification de portefeuille***

Comme pour le modèle à indice simple, on suppose ici que les résidus des régressions relatives à des titres différents sont non corrélés :  $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$  pour  $i \neq j$ .

Le risque du portefeuille s'écrit

$$\begin{aligned}
\sigma_P^2 &= \text{var} \left( \sum_{j=1}^N X_j \alpha_j + \beta_P R_M + \sum_{j=1}^N X_j \varepsilon_j \right) \\
&= \beta_P^2 \sigma_M^2 + \sum_{j=1}^N X_j^2 \sigma_{\varepsilon_j}^2
\end{aligned}$$

Si pour simplifier on considère un portefeuille constitué de  $N$  titres en égales proportions

$$X_j = \frac{1}{N} \quad j = 1, \dots, N,$$

le risque du portefeuille peut se développer en

$$\begin{aligned} \sigma_p^2 &= \text{var} \left( \sum_{j=1}^N X_j \alpha_j + \beta_p R_M + \sum_{j=1}^N X_j \varepsilon_j \right) \\ &= \beta_p^2 \sigma_M^2 + \frac{1}{N^2} \sum_{j=1}^N \sigma_{\varepsilon_j}^2 \\ &= \beta_p^2 \sigma_M^2 + \frac{1}{N} \overline{\sigma_{\varepsilon}^2} \end{aligned}$$

où on a introduit la variance résiduelle moyenne

$$\overline{\sigma_{\varepsilon}^2} = \frac{1}{N} \sum_{j=1}^N \sigma_{\varepsilon_j}^2.$$

Le premier terme de la décomposition est indépendant de  $N$  alors que le second tend vers 0 lorsque  $N$  devient très grand. Cette analyse montre donc que le risque  $\sigma_p^2$  du portefeuille peut se décomposer en deux termes :

- la composante systématique  $\beta_p^2 \sigma_M^2$  (*risque non diversifiable*) ;
- la composante spécifique  $\sum X_j^2 \sigma_{\varepsilon_j}^2$  (*risque diversifiable*).

### 3.3. MODELE D'EQUILIBRE DES ACTIFS FINANCIERS ET APPLICATIONS

#### 3.3.1. Capital asset pricing model

Ce modèle, développé indépendamment par W. Sharpe<sup>59</sup> et J. Lintner<sup>60</sup> et dont l'acronyme est MEDAF, ou C.A.P.M. en anglais, ne s'intéresse pas, contrairement à la section précédente, au choix d'un portefeuille pour un investisseur individuel, mais au comportement d'ensemble d'un marché lorsque les investisseurs agissent rationnellement<sup>61</sup> et présentent de l'aversion pour le risque. L'objectif, dans ces conditions, est de déterminer la juste valeur d'une action.

<sup>59</sup> SHARPE W., Capital assets prices, *Journal of Finance*, Vol. 19, 1964, pp. 425-442.

<sup>60</sup> LINTNER J., The valuation of risky assets and the selection of risky investments, *Review of Economics and Statistics*, Vol. 47, 1965, pp. 13-37.

<sup>61</sup> C'est-à-dire suivant la théorie du portefeuille basée sur l'analyse moyenne-variance.

## ***Hypothèses***

Le modèle en question repose sur un certain nombre d'hypothèses.

Les hypothèses relatives au comportement des investisseurs sont :

- ils composent leur portefeuille en utilisant la théorie du portefeuille de Markowitz, c'est-à-dire en s'appuyant sur le couple moyenne-variance ;
- ils ont tous les mêmes anticipation, c'est-à-dire qu'aucun n'a d'information privilégiée et qu'ils sont d'accord sur les valeurs des paramètres  $E_i$ ,  $\sigma_i$  et  $\sigma_{ij}$  à utiliser.

On formule également des hypothèses relatives aux transactions :

- elles se réalisent sans coût ;
- les époques d'achat et de vente ainsi que les durées de détention sont les mêmes pour tous les investisseurs.

On suppose enfin que les conditions suivantes sont vérifiées relativement au marché :

- il n'y a pas de taxation, que ce soit sur les plus-values, les dividendes ou les revenus d'intérêt ;
- les acheteurs et les vendeurs sont très nombreux sur le marché et ils n'ont d'autre influence sur les prix que celle résultant de la loi de l'offre et de la demande ;
- il existe un taux d'intérêt sans risque,  $R_F$ , utilisé tant pour l'emprunt que pour le placement ;
- les possibilités d'emprunt et de placement à ce taux n'ont pas de limite de volume.

Ces hypothèses ne sont bien évidemment pas réalistes. Il existe d'ailleurs des extensions par rapport au modèle que nous présentons ici et qui assouplissent certaines des hypothèses formulées. Par ailleurs, même le modèle de base donne de bons résultats, ainsi que les applications qui en découlent (voir §§ 3.3.3, 3.3.4 et 3.3.5).

## ***Théorème de séparation***

Ce théorème exprime que, dans les conditions précisées ci-dessus, tous les portefeuilles détenus par les investisseurs sont, à l'équilibre, des combinaisons de l'actif sans risque et du portefeuille de marché.

D'après les hypothèses, tous les investisseurs ont la même frontière efficiente pour les actions et le même taux sans risque  $R_F$ . Par conséquent, d'après l'étude du modèle de Markowitz avec titre non risqué (§ 3.2.5), le portefeuille de chaque investisseur se trouve sur la droite issue du point  $(0, R_F)$  et tangente à la frontière efficiente. Ce portefeuille est composé (voir figure 3.25)

- en proportion  $X$ , de l'actif sans risque
- en proportion  $1 - X$ , du portefeuille  $A$  correspondant au point de contact de la tangente.

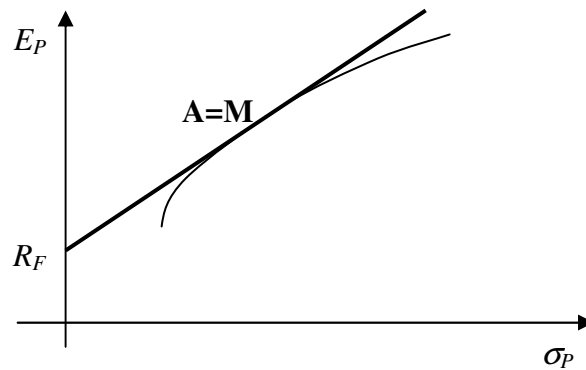


Figure 3.25 – Théorème de séparation et droite de marché

Le portefeuille risqué  $A$  est donc le même pour tous les investisseurs. Le marché va par conséquent, par le principe de l'offre et de la demande, adapter les prix de manière à ce que les proportions dans ce portefeuille soient celles du marché tout entier ( $A \equiv M$ ) et les portefeuilles détenus par les investisseurs sont parfaitement diversifiés.

Le choix de l'investisseur se fera donc seulement sur la proportion  $X$  du portefeuille de marché (et donc celle  $1 - X$  de l'actif non risqué). Si le portefeuille choisi est situé à gauche du point  $M$  ( $0 < X < 1$ ), il s'agit réellement d'une combinaison des deux investissements ; s'il se trouve à la droite de  $M$  ( $X > 1$ ), l'investisseur emprunte au taux  $R_F$  pour pouvoir acquérir plus de 100 % du portefeuille de marché. La droite en question porte le nom de *droite de marché*.

Le théorème de séparation a une interprétation simple. La droite de marché passe par les points  $(0, R_F)$  et  $(\sigma_M, E_M)$ . Son équation est par conséquent donnée par

$$E_P = R_F + \frac{E_M - R_F}{\sigma_M} \cdot \sigma_P.$$

Le return espéré  $E_P$  d'un portefeuille est égal au taux sans risque  $R_F$ , augmenté d'une prime de risque perçue par l'investisseur qui accepte de prendre un risque  $\sigma_P$ . Le coefficient de  $\sigma_P$  (la pente de la droite de marché) est donc l'accroissement de return espéré obtenu pour supporter une unité de risque : c'est le prix unitaire du risque sur le marché.

### **Equation du C.A.P.M.**

Nous allons déterminer maintenant une relation assez similaire à la précédente – c'est-à-dire une relation entre le return espéré et le risque – mais à propos d'un titre au lieu d'un portefeuille.

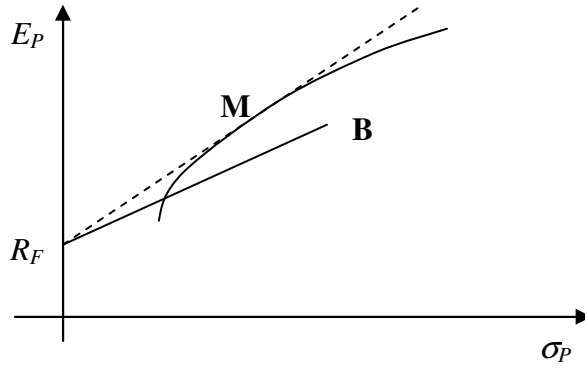


Figure 3.26 – C.A.P.M.

Pour un portefeuille d'actions quelconque  $B$ , la droite reliant les points  $(0, R_F)$  et  $(\sigma_B, E_B)$  a pour pente

$$\Psi_B = \frac{E_B - R_F}{\sigma_B}.$$

Celle-ci est clairement maximale lorsque  $B = M$  et, de la même manière, la valeur maximale de  $\Psi_B^2$  est  $\Psi_M^2$ . Par conséquent, si on appelle  $X_1, X_2, \dots, X_N$  les proportions des différents titres dans le portefeuille de marché ( $\sum X_i = 1$ ), on aura

$$(\Psi_M^2)_{X_k}' = 0 \quad k = 1, \dots, N$$

Comme

$$\begin{cases} E_M - R_F = \sum_{j=1}^N X_j E_j - \left( \sum_{j=1}^N X_j \right) R_F = \sum_{j=1}^N X_j (E_j - R_F) \\ \sigma_M^2 = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} \end{cases}$$

la dérivée de

$$\Psi_M^2 = \frac{(E_M - R_F)^2}{\sigma_M^2} = \frac{\left( \sum_{j=1}^N X_j (E_j - R_F) \right)^2}{\sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij}},$$

par rapport à  $X_k$  est donnée par



$$\begin{aligned}
(\Psi_M^2)_{X_k} &= \frac{2 \left( \sum_{j=1}^N X_j (E_j - R_F) \right) (E_k - R_F) \cdot \sigma_M^2 - \left( \sum_{j=1}^N X_j (E_j - R_F) \right)^2 \cdot 2 \sum_{j=1}^N X_j \sigma_{kj}}{\sigma_M^4} \\
&= \frac{2(E_M - R_F)(E_k - R_F)\sigma_M^2 - 2(E_M - R_F)^2 \sum_{j=1}^N X_j \sigma_{kj}}{\sigma_M^4} \\
&= \frac{2(E_M - R_F) \cdot ((E_k - R_F)\sigma_M^2 - (E_M - R_F)\sigma_{kM})}{\sigma_M^4}
\end{aligned}$$

Celle-ci est nulle si

$$E_k - R_F = (E_M - R_F) \frac{\sigma_{kM}}{\sigma_M^2}$$

ou encore

$$E_k = R_F + \beta_k (E_M - R_F).$$

Il s'agit là de l'équation du C.A.P.M., dont l'interprétation est similaire à la relation du paragraphe précédent. Le return espéré  $E_k$  du titre ( $k$ ) est égal au taux sans risque  $R_F$ , augmenté d'une prime de risque perçue par l'investisseur qui accepte de prendre un risque. Cette prime de risque est l'accroissement de return espéré, auquel on donne d'autant plus d'importance que ce titre est risqué au sein de ce marché ( $\beta_k$ ).

### Remarque

Comme nous l'avons dit, les hypothèses à la base du modèle qui vient d'être développé sont passablement peu réalistes. Des études empiriques ont été réalisées afin de tester si les résultats obtenus à partir de l'application du modèle du C.A.P.M. sont valides. Parmi les analyses les plus fouillées, on peut citer celle de Fama et MacBeth<sup>62</sup>, qui, considérant la relation

$$E_k = R_F + \beta_k (E_M - R_F),$$

comme une expression de  $E_k$  en fonction de  $\beta_k$ , ont testé sur la bourse de New-York les hypothèses suivantes

- la relation  $E_k = f(\beta_k)$  est linéaire et croissante ;
- $\beta_k$  est une mesure complète du risque de l'action ( $k$ ) sur le marché ; c'est-à-dire que le risque spécifique  $\sigma_{\varepsilon_k}^2$  n'est pas significativement explicatif de  $E_k$ .

<sup>62</sup> FAMA E. et MACBETH J., Risk, return and equilibrium : empirical tests, *Journal of Political Economy*, Vol. 71, n°1, 1974, pp. 607-636.

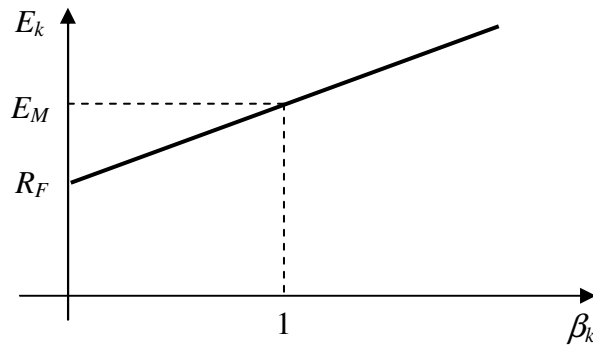


Figure 3.27 – Test du C.A.P.M.

Pour ce faire, ils ont utilisé des généralisations de l'équation  $E_k = f(\beta_k)$ , comprenant des puissances de  $\beta_k$  de degré supérieur à 1 et un terme prenant en considération le risque spécifique. Leur conclusion est que le modèle du C.A.P.M. est, dans la plupart des cas, acceptable.♣

### 3.3.2. Arbitrage pricing theory

Dans le modèle du C.A.P.M., la prime de risque  $E_k - R_F$  d'un titre est exprimée comme un multiple de la prime de risque  $E_M - R_F$  du marché :

$$E_k - R_F = \beta_k(E_M - R_F),$$

le coefficient de proportionnalité étant le  $\beta$  du titre. On peut donc considérer que cette approche permet d'exprimer la prime de risque d'un titre à partir de la prime de risque d'un seul facteur macroéconomique explicatif ou, ce qui revient au même, à partir d'un agrégat incluant tous les facteurs macroéconomiques interagissant avec le marché.

Le modèle de l'*Arbitrage Pricing Theory*<sup>63</sup> (A.P.T.) permet une analyse plus fine du portefeuille que le modèle du C.A.P.M., la décomposition du risque selon le seul facteur de marché, le bêta, pouvant se révéler insuffisante pour décrire l'ensemble des risques d'un portefeuille d'actions ; d'où l'intérêt de recourir à des décompositions du risque en fonction de plusieurs facteurs  $F_1, F_2, \dots, F_p$  :

$$E_k - R_F = \sum_{j=1}^p \alpha_{kj}(E_{F_j} - R_F).$$

La théorie de l'A.T.P. indique que, dans un marché efficient, les prix des actifs cotés s'équilibrent par arbitrages successifs grâce à l'intervention des acteurs sur le marché. Si l'on s'attache à regarder l'évolution des prix relatifs, on peut extraire du marché un petit nombre de facteurs d'arbitrage qui ont permis aux prix de s'équilibrer. C'est précisément ce que réalise le modèle A.P.T.

<sup>63</sup> ROSS S.A., The arbitrage theory of capital asset pricing, *Journal of economic theory*, 1976, pp. 343-362.

Les premières versions<sup>64</sup> du modèle A.P.T. s'appuyaient sur une liste prédéfinie de facteurs fondamentaux, comme un indice d'activité industrielle, le taux d'inflation, l'écart entre taux d'intérêt à court et à long terme, l'écart des rendements d'obligations aux ratings très différents (voir § 4.2.1), ... Les coefficients  $\alpha_{k1}, \dots, \alpha_{kp}$  se déterminent alors par une technique de régression à partir des observations historiques

$$R_{kt} \text{ et } R_{F_j,t} \quad (j = 1, \dots, p)$$

Les versions plus récentes constituent des méthodes davantage empiriques, dégagant des facteurs explicatifs non corrélés par un technique statistique<sup>65</sup> (*analyse factorielle*), sans que le nombre de ces facteurs soit connu a priori, et même sans que ceux-ci aient la moindre interprétation économique.

Les facteurs extraits<sup>66</sup> des séries temporelles des rendements des prix des actifs ont une existence purement statistique. Ce ne sont pas, pris individuellement, des variables communément utilisées pour décrire un processus de construction de portefeuille ou une stratégie de gestion. Aucun d'eux n'est le représentant d'un taux d'intérêt, de l'inflation, ou encore d'un taux de change. Ils constituent l'équivalent d'un système d'axes orthogonaux en géométrie.

Le seul but qui a été poursuivi est de disposer d'un référentiel qui permette de décrire l'ensemble des interrelations des actifs étudiés de façon stable dans le temps. Le référentiel étant établi, le risque de tout actif côté (actions, obligations, fonds d'investissement, ...) est décomposé en une partie systématique (commune à tous les actifs du marché) représentable dans l'espace des facteurs et une partie spécifique (propre à l'actif). La partie systématique est ensuite explicitée par des coefficients de sensibilité (les  $\alpha_{kj}$ ) aux différents facteurs statistiques.

Le pouvoir explicatif du modèle est dû au fait que les différentes variables usuelles (variables économiques, sectorielles, fondamentales, etc.) utilisées pour comprendre le comportement de celui-ci sont également représentées dans le référentiel des facteurs, dès lors qu'un support coté associé existe (historique de prix).

La relation liant le rendement d'un titre aux différents facteurs permet de décomposer sa variance en une partie liée aux facteurs de risque systématique (les facteurs statistiques explicatifs) et une partie spécifique aux titres et diversifiable (les résidus de la régression), soit

$$\sigma_k^2 = \sum_{j=1}^p \alpha_{kj}^2 \text{var}(R_{F_j}) + \sigma_{\varepsilon_k}^2 .$$

<sup>64</sup> DHRYMES P.J., FRIEND I. et GULTEKIN N.B., A critical reexamination of the empirical evidence on the arbitrage pricing theory, *Journal of Finance*, n° 39, 1984, pp. 323-346.

CHEN N.F., ROLL R. et ROSS S.A., Economic forces of the stock market, *Journal of Business*, n° 59, 1986, pp. 383-403.

On consultera plus généralement GRINOLD C. et KAHN N., *Active Portfolio Management*, Mc Graw Hill, 1998.

<sup>65</sup> Voir par exemple SAPORTA G., *Probabilités, analyse des données et statistique*, Technip, 1990, ou MORRISON D., *Multivariate statistical methods*, 1976.

<sup>66</sup> Le lecteur intéressé par les développements mathématiques sous-jacents à l'extraction des facteurs statistiques des séries historiques de rendements croisés d'actifs consultera utilement l'ouvrage de MEHTA M. L., *Random matrices*, Academic Press, 1996. Cet ouvrage traite en détail des problèmes de calculs de valeurs propres et vecteurs propres de matrices possédant un très grand nombre d'éléments qui sont générés de façon aléatoire.

### Illustration

Un développement technique de cette méthode – accompagné d’un logiciel est dû à la société Advanced Portfolio Technologies Inc.<sup>67</sup> Il extrait du marché (représenté par les séries temporelles de rendements croisés d’actifs) des facteurs statistiques à l’aide d’un algorithme de recherche de forme.

Ainsi, si le haut de la figure 3.28 représente les séries observées des rendements de quatre titres, le bas de la même figure illustre les trois facteurs primaires qui permettent, par combinaison linéaire, de reconstruire les quatre séries précédentes.

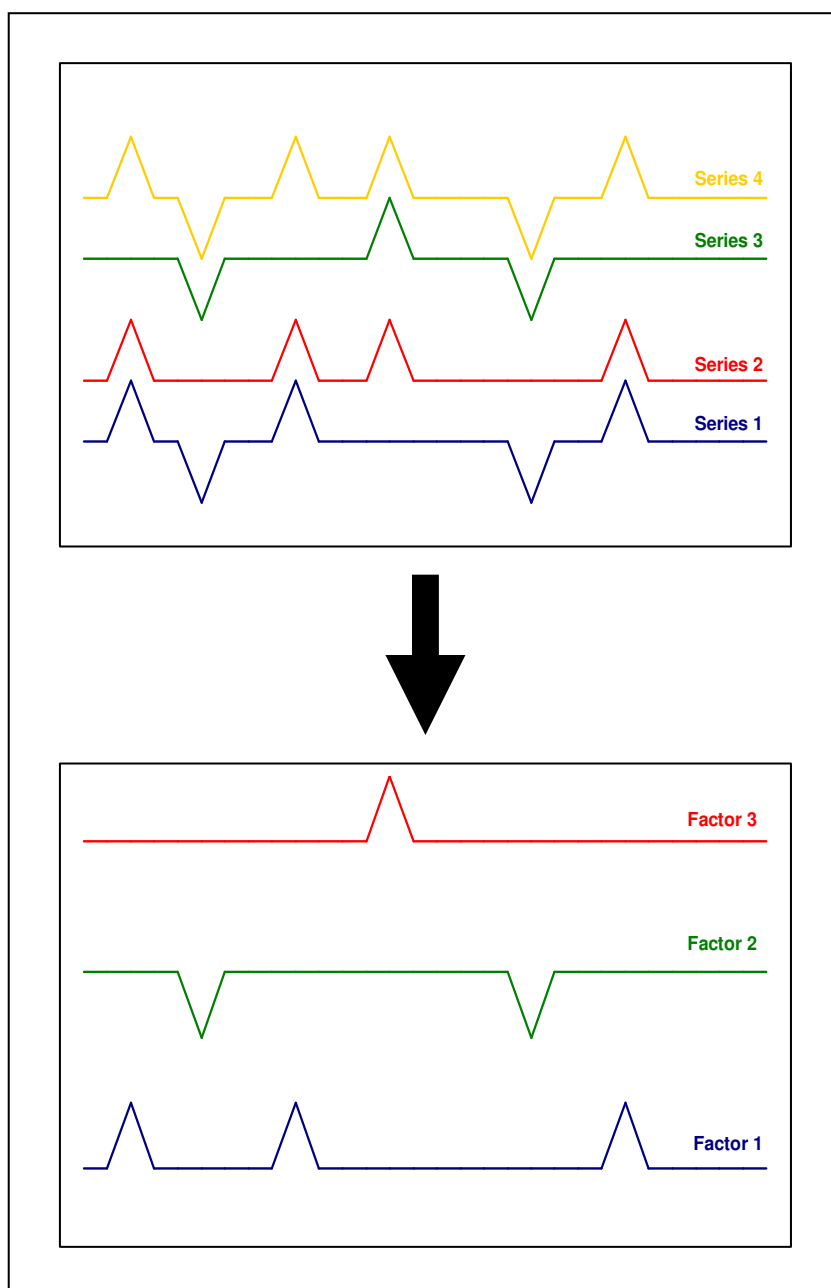


Figure 3.28 – Arbitrage pricing theory

<sup>67</sup> On consultera à ce sujet le site Internet <http://www.aptltd.com>.

Par exemple, la première série se décompose en

$$R_1 - R_F = 1 \cdot (R_{F_1} - R_F) + 1 \cdot (R_{F_2} - R_F) + 0 \cdot (R_{F_3} - R_F) + \varepsilon_1 \cdot \clubsuit$$

### 3.3.3. Evaluation des performances

#### *Principe*

Le gestionnaire de portefeuille<sup>68</sup> est bien entendu intéressé par la performance du produit qu'il gère. Il comparera pour ce faire le rendement de son portefeuille à celui du marché sur lequel il investit ; d'un point de vue pratique, cette comparaison se réalisera par rapport à un indice représentatif du marché, du secteur, ... considéré.

#### *Remarque*

Le rendement d'un portefeuille réel entre les instants  $s$  et  $t$  se calcule simplement par la relation

$$R_{P,[s;t]} = \frac{V_t - V_s}{V_s},$$

à condition qu'il n'y ait pas de flux de trésorerie au sein de ce portefeuille durant l'intervalle de temps en question. Cependant, ces flux existent en général (nouveaux titres acquis, titres vendus, ...). Il convient donc d'évaluer le return en éliminant l'effet de ces mouvements.

Notons  $t_1 < \dots < t_n$  les époques où ces flux interviennent et posons  $t_0 = s$  et  $t_{n+1} = t$ . Le return à prendre en considération est alors donné par

$$R_{P,[s;t]} = \prod_{k=0}^n (1 + R_{P,[t_k;t_{k+1}]}) - 1,$$

où on a posé

$$R_{P,[t_k;t_{k+1}]} = \frac{V_{t_{k+1}}^{(-)} - V_{t_k}^{(+)}}{V_{t_k}^{(+)}} ,$$

$V_{t_j}^{(-)}$  et  $V_{t_j}^{(+)}$  représentant la valeur du portefeuille respectivement juste avant et juste après le mouvement de trésorerie de l'instant  $t_j \cdot \clubsuit$

<sup>68</sup> Les stratégies de gestion, passive et active, seront abordées au paragraphe suivant.

Le paragraphe 3.2 l'a montré à l'évidence, la qualité d'un titre ou d'un portefeuille n'est pas seulement mesurée par son return. Que penser en effet des portefeuilles  $A$  et  $B$  dont les returns pour une période donnée sont respectivement de 6,2 % et 6,3 % mais pour lesquels le risque de  $B$  est double de celui de  $A$  ? Les indices de mesure de performance présentés ci-après prendront donc en considération non seulement le rendement, mais également le risque du titre ou du portefeuille.

Les indicateurs que nous signalons ici s'appuient tous sur des relations issues du modèle d'évaluation des actifs financiers et plus particulièrement sur l'équation du C.A.P.M. Ils supposent donc que les hypothèses sous-jacente à ce modèle sont satisfaites.

Les deux premiers indicateurs s'articulent respectivement autour de l'équation de la droite de marché et de l'équation du C.A.P.M. ; le troisième est une variante du deuxième.

### *Indice de Sharpe*

L'équation de la droite de marché est :

$$E_P = R_F + \frac{E_M - R_F}{\sigma_M} \cdot \sigma_P,$$

qui peut se réécrire sous la forme suivante :

$$\frac{E_P - R_F}{\sigma_P} = \frac{E_M - R_F}{\sigma_M}.$$

Cette relation exprime que l'excès de return espéré (par rapport au taux sans risque), rapporté à l'écart type, est, à l'équilibre, identique pour un portefeuille bien diversifié et pour le marché. On appelle *indice de Sharpe* l'expression

$$SI_P = \frac{E_P - R_F}{\sigma_P},$$

qu'on comparera en pratique à l'expression équivalente calculée pour un indice représentatif du marché.

### *Illustration*

Reprenons les données utilisées pour le modèle à indice simple de Sharpe (§ 3.2.4) :

$$\begin{array}{lll} E_1 = 0,05 & E_2 = 0,08 & E_3 = 0,10 \\ \sigma_1 = 0,10 & \sigma_2 = 0,12 & \sigma_3 = 0,15 \\ \rho_{12} = 0,3 & \rho_{13} = 0,1 & \rho_{23} = 0,4 \end{array}$$

et considérons le portefeuille particulier relatif à la valeur  $\lambda = 0,010$  du paramètre de risque. On a, dans ce cas,

$$X_1 = 0,4387 \quad X_2 = 0,1118 \quad X_3 = 0,4496$$

et donc  $E_P = 0,0758$  et  $\sigma_P = 0,0912$ . On a également  $E_I = 0,04$  et  $\sigma_I = 0,0671$  et on prend, comme au § 3.2.5,  $R_F = 0,03$ .

L'indice de Sharpe du portefeuille est donc donné par

$$SI_P = \frac{0,0758 - 0,03}{0,0912} = 0,7982,$$

alors que celui relatif à l'indice est égal à

$$SI_I = \frac{0,04 - 0,03}{0,0671} = 0,1490,$$

ce qui montre que le portefeuille en question est plus performant que le marché. ♣

### ***Indice de Treynor***

L'équation du C.A.P.M. pour le  $k$ -ème titre du portefeuille

$$E_k = R_F + \beta_k (E_M - R_F),$$

permet d'écrire

$$\sum_{k=1}^N X_k E_k = \left( \sum_{k=1}^N X_k \right) \cdot R_F + \left( \sum_{k=1}^N X_k \beta_k \right) \cdot (E_M - R_F)$$

ou encore,

$$E_P = R_F + \beta_P (E_M - R_F).$$

En tenant compte du fait que  $\beta_M = 1$ , cette dernière relation peut s'écrire

$$\frac{E_P - R_F}{\beta_P} = \frac{E_M - R_F}{\beta_M},$$

avec une interprétation similaire à celle de l'indice de Sharpe. L'*indice de Treynor* est donc défini par

$$TI_P = \frac{E_P - R_F}{\beta_P},$$

qui sera comparé à l'expression similaire pour un indice.

### ***Illustration***

Reprenons les données ci-dessus avec, en plus (voir § 3.2.4),

$$\beta_1 = 0,60 \quad \beta_2 = 1,08 \quad \beta_3 = 1,32$$

ce qui donne  $\beta_P = 0,9774$ .

L'indice de Treynor de ce portefeuille est donc donné par

$$TI_P = \frac{0,0758 - 0,03}{0,9774} = 0,0469,$$

alors que celui relatif à l'indice vaut

$$TI_I = \frac{0,04 - 0,03}{1} = 0,0100,$$

ce qui conduit à la même conclusion.♣

## ***Indice de Jensen***

D'après le raisonnement consacré à l'indice de Treynor, on a

$$E_P - R_F = \beta_P(E_M - R_F).$$

Cette relation étant relative, à l'équilibre, pour un portefeuille bien diversifié, un portefeuille  $P$  présentera un excédent de return par rapport au marché s'il existe un nombre  $\alpha_P > 0$  tel que

$$E_P - R_F = \alpha_P + \beta_P(E_M - R_F).$$

L'indice de Jensen

$$JI_P = \hat{\alpha},$$

est l'estimateur du terme indépendant de la régression

$$E_{P,t} - R_{F,t} = \alpha + \beta(E_{I,t} - R_{F,t}),$$

dont la variable à expliquer (resp. explicative) est l'excès de return du portefeuille par rapport au taux sans risque (resp. l'excès de return d'un indice représentatif du marché). On compare bien sûr sa valeur à 0.

## ***Illustration***

On vérifie aisément qu'avec les données précédentes, on a



$$JI_P = (0,0758 - 0,03) - 0,9774 \cdot (0,04 - 0,03) = 0,0360,$$

qui est bien strictement positif.♣

### 3.3.4. Stratégies de gestion d'un portefeuille d'actions

#### *Gestion passive*

L'objectif d'une gestion passive est d'obtenir un return égal à celui du marché.

Par définition du marché, les gains (rendements supérieurs à celui du marché) réalisés par certains investisseurs sont compensés par des pertes (rendements inférieurs à celui du marché) subies par d'autres investisseurs<sup>69</sup> : le return moyen obtenu par l'ensemble des investisseurs est celui du marché. La réalité est quelque peu différente : en raison des coûts de transaction, le return moyen des investisseurs est légèrement inférieur à celui du marché.

La stratégie passive consiste donc à

- constituer un portefeuille de composition identique (ou très similaire) au marché, qui correspond à une diversification optimale ;
- limiter autant que faire se peut le volume des transactions.

Cette manière de faire pose un certain nombre de problèmes. Par exemple, pour la gestion de certains types de portefeuilles, des réglementations imposent que chaque titre ne soit présent qu'à une hauteur maximale fixée, ce qui est incompatible avec une gestion passive si un titre représente sur le marché une capitalisation boursière particulièrement importante. Autre problème : la présence de certains titres, dont le cours est très élevé (et qui sont indivisibles !), pourraient conduire à la construction de portefeuilles dont la valeur serait importante au point qu'ils deviennent inutilisables en pratique.

Ces problèmes ont conduit à la réalisation d'organismes de placement collectifs qui « imitent » le marché : les fonds d'indice. Après avoir choisi un indice représentatif du marché sur lequel on désire investir, on construit un portefeuille constitué des mêmes titres que ceux de l'indice (ou parfois simplement des plus importants), dans les mêmes proportions.

Bien entendu, au fur et à mesure de l'évolution des cours des titres constitutifs, la composition du portefeuille doit être adaptée, ce qui est à l'origine de certaines difficultés. Ainsi, les temps de réaction provoquent des écarts inévitables entre le rendement du portefeuille et celui du marché : ce sont les *erreurs de poursuite*<sup>70</sup>. De plus, ce type de gestion entraîne un certain nombre de frais de transaction : pour l'adaptation du portefeuille à l'indice, pour le réinvestissement des dividendes, ... Pour ces raisons, le rendement d'un tel portefeuille sera en général légèrement inférieur à celui de l'indice.

<sup>69</sup> Une telle situation porte le nom, dans la théorie des jeux, de *jeu à somme nulle*. On consultera par exemple BINMORE K., *Jeux et théorie des jeux*, De Boeck & Larcier, 1999.

<sup>70</sup> En anglais, *tracking errors*.

## ***Gestion active***

L'objectif de la gestion active est d'obtenir un return supérieur à celui du marché.

Un marché parfaitement efficient ne peut être battu que momentanément et par chance : à long terme, le rendement ne peut dépasser celui du marché. Une gestion active suppose donc le marché non parfaitement efficient.

Deux grands principes permettent d'atteindre l'objectif fixé.

- a) L'*allocation d'actifs*<sup>71</sup> évolutive dans le temps – encore appelée *market timing* – consiste en la construction d'un portefeuille constitué en partie du portefeuille de marché ou d'un portefeuille indiciel et en partie dans un actif sans risque (ou significativement moins risqué que des actions, comme les obligations). On fait alors évoluer les proportions respectives de ces deux composantes au cours du temps suivant que l'on anticipe une hausse ou une baisse de l'indice.
- b) La *sélection de titres*<sup>72</sup> consiste en la construction d'un portefeuille d'actions en choisissant les titres considérés comme sous-évalués et qui devraient donc procurer dans un avenir plus ou moins proche (réaction du marché) un rendement supérieur à celui du marché.

Dans la pratique, les professionnels mettent en œuvre des stratégies basées sur l'une des deux approches ou sur un mélange des deux.

Afin d'évaluer la qualité de la gestion active, il convient de comparer le portefeuille construit et le portefeuille de marché tant du point de vue du return espéré que du risque encouru. Ces indices de performance de portefeuilles ont été étudiés au § 3.3.3.

Examinons maintenant quelques méthodes de mise en œuvre du market timing ainsi qu'une technique de stock picking : l'application du *dividend discount model*.

## ***Market timing***

Cette technique consiste donc en la gestion d'un portefeuille composé du portefeuille de marché ( $M$ ) pour les actions et d'un taux obligataire ( $O$ ), dans des proportions respectives  $X$  et  $1 - X$ , et à l'adaptation de  $X$  en fonction des performances attendues de ces deux composantes.

Ces performances, qui déterminent une politique de market timing, peuvent être évaluées par différents critères :

- le « *price-earning* » ratio, présenté au § 3.1.3 :  $PER = \text{cours/bénéfice}$  ;
- le « *yield gap* », qui est le rapport entre le rendement obligataire et le rendement des actions (dividende/cours) ;
- le « *earning yield* », qui est le produit du  $PER$  par le taux obligataire ;
- la *prime de risque*, qui est la différence entre le rendement du portefeuille de marché et le rendement obligataire :  $PR = E_M - E_O$  ; elle peut être estimée à l'aide d'un

<sup>71</sup> En anglais *asset allocation*.

<sup>72</sup> En anglais *stock picking*.

historique, mais il est préférable d'utiliser une estimation ex ante réalisée par un analyste financier, par exemple à l'aide du D.D.M. (voir ci-après). Bien sûr, de petites valeurs des trois premiers critères sont favorables à l'investissement en actions ; la situation est inversée pour la prime de risque.

Une première méthode pour mettre en œuvre une politique de market timing est le recours à des *canaux décisionnels*. Si on note  $c$  un des quatre critères cités ci-dessus, pour lequel on dispose d'observations historiques (et donc une estimation  $\bar{c}$  de sa moyenne et  $s_c$  de son écart type), on choisit – de manière quelque peu arbitraire – d'investir un certain pourcentage en actions en fonction de la valeur observée de  $c$ , comparée à  $\bar{c}$ , l'écart entre les deux étant modulé par  $s_c$ . On choisira par exemple d'investir 70 %, 60 %, 50 %, 40 % ou 30 % en actions d'après la position de  $c$  par rapport aux limites<sup>73</sup>  $\bar{c} - \frac{3}{2}s_c$ ,  $\bar{c} - \frac{1}{2}s_c$ ,  $\bar{c} + \frac{1}{2}s_c$  et  $\bar{c} + \frac{3}{2}s_c$ .

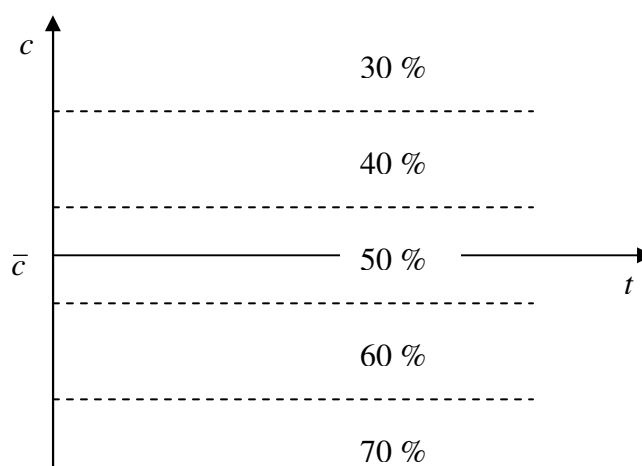


Figure 3.29 – Canaux décisionnels figés

Cette manière de procéder ne tient pas compte de l'évolution du paramètre  $c$  dans le temps. On peut par conséquent calculer les paramètres  $\bar{c}$  et  $s_c$  sur un historique glissant (un an par exemple).

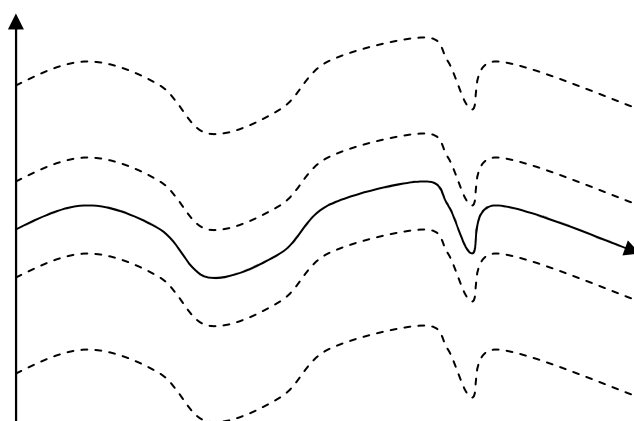


Figure 3.30 – Canaux décisionnels mobiles

Une autre manière de procéder, plus rigoureuse, est utilisable seulement avec la prime de risque. Dans la recherche de la frontière efficiente, on a chaque fois recherché le minimum par

<sup>73</sup> L'ordre des canaux doit être inversé pour la prime de risque.

rapport aux proportions de l'expression  $\sigma_p^2 - \lambda E_p$  où le paramètre  $\lambda$  correspond au risque ( $\lambda = 0$  pour un portefeuille prudent,  $\lambda = +\infty$  pour un portefeuille spéculatif). Ce paramètre est égal à la pente de la droite, dans le plan  $(E, \sigma^2)$ , tangente à la frontière efficiente et issue du point  $(R_F, 0)$ . D'après le théorème de séparation (voir § 3.3.1), le point de contact de cette tangente correspond au portefeuille de marché (voir figure 3.31) et on a par conséquent

$$\lambda = \frac{\sigma_M^2}{E_M - R_F}.$$

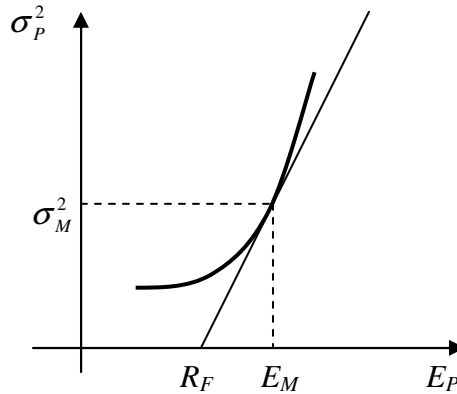


Figure 3.31 – Théorème de séparation

Par ailleurs, le return du portefeuille composé d'une proportion  $X$  de portefeuille de marché et d'une proportion  $1 - X$  de taux obligataire est donné par

$$R_p = XR_M + (1 - X)R_o,$$

ce qui permet de déterminer

$$\begin{aligned} E_p &= XE_M + (1 - X)E_o \\ \sigma_p^2 &= X^2\sigma_M^2 + 2X(1 - X)\sigma_{MO} + (1 - X)^2\sigma_o^2 \end{aligned}$$

Le problème consiste donc à déterminer la valeur de  $X$  qui minimise l'expression

$$Z(X) = \sigma_p^2 - \lambda E_p = X^2\sigma_M^2 + 2X(1 - X)\sigma_{MO} + (1 - X)^2\sigma_o^2 - \lambda[XE_M + (1 - X)E_o].$$

L'annulation de la dérivée de cette fonction

$$\begin{aligned} Z'(X) &= 2X\sigma_M^2 + 2(1 - 2X)\sigma_{MO} - 2(1 - X)\sigma_o^2 - \lambda(E_M - E_o) \\ &= 2X(\sigma_M^2 - 2\sigma_{MO} + \sigma_o^2) + 2\sigma_{MO} - 2\sigma_o^2 - \lambda \cdot PR \end{aligned}$$

fournit la proportion recherchée :

$$X = \frac{\lambda \cdot PR - 2(\sigma_{MO} - \sigma_o^2)}{2(\sigma_M^2 - 2\sigma_{MO} + \sigma_o^2)}$$

ou, de manière équivalente, en remplaçant  $\lambda$  et  $PR$  par leur valeur,

$$X = \frac{\frac{E_M - E_0}{E_M - R_F} \cdot \sigma_M^2 - 2(\sigma_{MO} - \sigma_O^2)}{2(\sigma_M^2 - 2\sigma_{MO} + \sigma_O^2)}.$$

### *Illustration*

Si on dispose des données suivantes

$$\begin{array}{ll} E_M = 0,08 & \sigma_M = 0,10 \\ E_O = 0,06 & \sigma_O = 0,02 \\ R_F = 0,04 & \rho_{MO} = 0,6 \end{array}$$

On peut calculer successivement

$$\begin{aligned} \sigma_{MO} &= 0,10 \cdot 0,02 \cdot 0,6 = 0,0012 \\ \lambda &= \frac{0,10^2}{0,08 - 0,04} = 0,25 \\ PR &= 0,08 - 0,06 = 0,02 \end{aligned}$$

et par conséquent

$$X = \frac{0,25 \cdot 0,02 - 2 \cdot (0,0012 - 0,02^2)}{2 \cdot (0,10^2 - 2 \cdot 0,0012 + 0,02^2)} = 0,2125.$$

Il conviendra donc d'investir, dans ces conditions, 21,25 % en actions (portefeuille de marché) et 78,75 % en obligations. ♣

### ***Dividend discount model***

La méthode dite « *dividend discount model* » (D.D.M.) a pour but de comparer, pour une action, son return attendu et son return à l'équilibre, ce qui permettra de déterminer si elle est sur-évaluée ou sous-évaluée.

La détermination du return attendu,  $\tilde{R}_k$ , se réalise à l'aide d'un modèle d'actualisation des dividendes futurs. On peut utiliser un raisonnement du type de celui de la formule de Gordon-Shapiro (§ 3.1.3) ou sur une généralisation de celui-ci. Alors que la relation de Gordon-Shapiro suppose un taux de croissance constant pour les dividendes, des modèles plus évolués utilisent par exemple (modèle à deux taux) un taux de croissance constant pendant quelques années, puis un autre taux constant, plus faible, pour les années suivantes, ou encore (modèle à trois taux), entre ces deux périodes à taux constants, une période de quelques années où le taux de croissance diminue linéairement afin de réaliser un raccordement continu.

Le return à l'équilibre  $E_k$  se détermine grâce à l'équation du C.A.P.M. (§ 3.3.1). Cette équation s'écrit

$$E_k = R_F + \beta_k (E_M - R_F).$$

Si on considère qu'elle exprime  $E_k$  en fonction de  $\beta_k$ , il s'agit là de l'équation d'une droite ; celle-ci passe par le point  $(0, R_F)$  et, puisque  $\beta_M = 1$ , par le point  $(1, E_M)$ . Cette droite porte le nom de *ligne d'évaluation des actifs financiers* ou encore « *security market line* ».

Si le return attendu  $\tilde{R}_k$  de chaque titre était égal à son return à l'équilibre  $E_k$ , tous les points  $(\beta_k, \tilde{R}_k)$  seraient situés sur la *security market line*. Ce n'est en pratique pas le cas, à cause de certaines inefficiences du marché (voir figure 3.32).

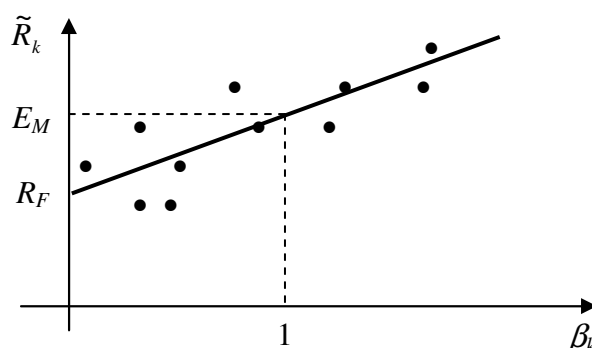


Figure 3.32 – *Security market line*

On considère, dans cette technique, que l'évaluation  $\tilde{R}_k$  par les analystes est correcte et que les écarts constatés sont dus à l'inefficience du marché. Par conséquent, les titres dont le point représentatif se situe au-dessus de la *security market line* sont considérés comme sous-évalués et le marché devrait, tôt ou tard, rectifier la situation et procurer un return supplémentaire à l'investisseur ayant acquis ces titres.

### 3.4. MODELES D'EVOLUTION DES ACTIONS

Les paragraphes qui précèdent traitent des aspects statiques : on considère une « photographie » de la situation à un moment donné. Nous abordons ici la modélisation de l'évolution du return ou du cours d'une action dans le temps.

La notation utilisée est ici quelque peu différente : le cours du titre à l'instant  $t$  est noté  $S_t$ . Il s'agit d'une notation classique (pour *stock* en anglais) ; par ailleurs, les présents modèles servent entre autres de support au développement des modèles d'évaluation des options sur action (voir § 5.3), pour lesquels la notation  $C_t$  est réservée aux options d'achat (*call* en anglais).

Signalons enfin que ce qui va suivre, sauf mention explicite du contraire, est uniquement valable pour des titres ne donnant pas lieu à distribution de dividende.

### 3.4.1. Modèles déterministes

#### *Modèle discret*

On évalue ici le titre aux instants  $t = 0, 1, \dots$ . Si on suppose que le rendement du titre entre les instants  $t$  et  $t + 1$  est  $i$ , on peut écrire

$$i = \frac{S_{t+1} - S_t}{S_t},$$

ce qui conduit à l'équation d'évolution

$$S_{t+1} = S_t \cdot (1 + i).$$

Si le taux de rendement  $i$  est constant et que l'on prend la valeur initiale  $S_0$  en compte, l'équation (aux différences) ci-dessus a pour solution

$$S_t = S_0 \cdot (1 + i)^t.$$

Si le taux varie de période en période ( $i_k$  pour la période  $]k - 1 ; k]$ ), la relation précédente devient

$$S_t = S_0 \cdot (1 + i_1)(1 + i_2) \dots (1 + i_t).$$

#### *Modèle continu*

On considère ici une évolution infinitésimale de la valeur du titre. Si on suppose le rendement entre les instants  $t$  et  $t + \Delta t$  (avec  $\Delta t \ll \text{petit}$ ) proportionnel à la durée  $\Delta t$  avec un facteur de proportionnalité  $\delta$ :

$$\delta \cdot \Delta t = \frac{S_{t+\Delta t} - S_t}{S_t},$$

l'équation d'évolution est ici une équation différentielle<sup>74</sup>:

$$S'_t = S_t \cdot \delta.$$

La solution de cette équation est donnée par

<sup>74</sup> Obtenue en faisant tendre  $\Delta t$  vers 0.

$$S_t = S_0 \cdot e^{\delta t}.$$

On remarque le lien évident entre cette relation et celle qui lui correspond pour le cas discret, à condition de poser  $\delta = \ln(1 + i)$ .

Dans le cas où le taux de rendement  $\delta$  n'est pas constant, l'équation différentielle d'évolution prend la forme

$$S'_t = S_t \cdot \delta(t),$$

ce qui conduit à la solution plus complexe

$$S_t = S_0 \cdot e^{\int_0^t \delta(t) dt}.$$

### Remarque

Les paramètres apparaissant dans les modèles ci-dessus (les taux constants  $i$  et  $\delta$ , ou les taux variables  $i_1, i_2, \dots$  et  $\delta(t)$ ) doivent bien sûr, pour une utilisation pratique, être estimés à partir des observations historiques par exemple.♣

### Généralisation

On peut bien entendu superposer ces deux aspects, discret et continu. On considère donc

- un processus d'évolution continue du taux de rendement, représenté par la fonction  $\delta(t)$  ; auquel se superpose
- un ensemble de variations discrètes intervenant aux époques  $\tau_1, \tau_2, \dots, \tau_k, \dots$  tel que le taux de rendement entre  $\tau_{k-1}$  et  $\tau_k$  est égal à  $i_k$ .

Si on note  $n$  le plus grand entier tel que  $\tau_n \leq t$ , l'évolution de la valeur du titre est donnée par

$$S_t = S_0 \cdot (1 + i_1)^{\tau_1} (1 + i_2)^{\tau_2 - \tau_1} \dots (1 + i_n)^{\tau_n - \tau_{n-1}} (1 + i_{n+1})^{t - \tau_n} \cdot e^{\int_0^t \delta(t) dt}.$$

Cette présentation permet par exemple de prendre en considération, dans une modélisation discrète ou continue, le processus de versement de dividendes. Ainsi, dans le cas où le modèle ne comporte que le volet continu représenté par  $\delta(t)$ , la relation ci-dessus représente l'évolution de la valeur d'un titre versant des dividendes aux époques  $\tau_1, \tau_2, \dots$  dont le montant  $D_k$  versé en  $\tau_k$  est lié à  $i_k$  par la relation

$$i_k = -\frac{D_k}{S_k^{(-)}},$$

où  $S_k^{(-)}$  est la valeur du titre juste avant le versement du  $k$ -ème dividende.



### 3.4.2. Modèles stochastiques

#### *Modèle discret*

On suppose que l'évolution, d'époque en époque se fait de la manière suivante : l'action  $a$ , à l'instant  $t$ , la valeur (aléatoire)  $S_t$  et aura à l'instant suivant  $t + 1$  une des deux valeurs  $S_t \cdot u$  (supérieure à  $S_t$ ) ou  $S_t \cdot d$  (inférieure à  $S_t$ ) avec les probabilités respectives  $\alpha$  et  $(1 - \alpha)$ .

On a donc  $d \leq 1 \leq u$ , mais on suppose même que  $d \leq 1 < 1 + R_F \leq u$ , sans quoi des opportunités d'arbitrage sont évidemment possibles. En pratique, les paramètres  $u$ ,  $d$  et  $\alpha$  sont à estimer à partir des observations.

On utilise généralement la représentation graphique suivante de l'évolution du prix de l'action :

$$\begin{array}{lcl} & \nearrow & S_{t+1} = S_t \cdot u \quad (\alpha) \\ S_t & & \\ & \searrow & S_{t+1} = S_t \cdot d \quad (1 - \alpha) \end{array}$$

On suppose que les paramètres  $u$ ,  $d$  et  $\alpha$  demeurent constants au cours du temps et nous n'explicitons plus la probabilité  $\alpha$  dans les graphiques qui suivent, les branches montantes par exemple correspondant toujours à l'augmentation (au taux  $u$ ) de la valeur du titre, avec probabilité  $\alpha$ .

Remarquons que le return de l'action entre les époques  $t$  et  $(t + 1)$  est donné par

$$\frac{S_{t+1} - S_t}{S_t} = \begin{cases} u - 1 & (\alpha) \\ d - 1 & (1 - \alpha) \end{cases}$$

Entre les instants  $t + 1$  et  $t + 2$ , on aura de la même manière, suivant la branche atteinte à la fin de la période précédente,

$$\begin{array}{lcl} & \nearrow & S_{t+2} = S_{t+1} \cdot u = S_t \cdot u^2 \\ S_{t+1} & & \\ & \searrow & S_{t+2} = S_{t+1} \cdot d = S_t \cdot ud \end{array}$$

ou

$$\begin{array}{lcl} & \nearrow & S_{t+2} = S_{t+1} \cdot u = S_t \cdot ud \\ S_{t+1} & & \\ & \searrow & S_{t+2} = S_{t+1} \cdot d = S_t \cdot d^2 \end{array}$$

On constate donc qu'une hausse suivie d'une baisse conduit au même résultat qu'une baisse suivie d'une hausse. D'une manière plus générale, on peut construire le graphe (voir figure

3.33), appelé *treillis binomial*, qui va de l'époque 0 (où l'action a la valeur certaine  $S_0$ ) à l'époque  $t$  :

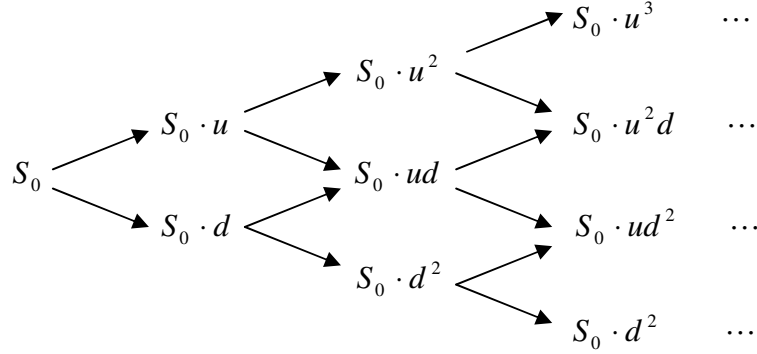


Figure 3.33 – Treillis binomial

On voit donc que la valeur (aléatoire) de l'action à l'instant  $t$  est donnée par

$$S_t = S_0 \cdot u^N d^{t-N},$$

où le nombre  $N$  de hausses est bien entendu une variable aléatoire binomiale<sup>75</sup> de paramètres  $(t; \alpha)$  :

$$\Pr[N = k] = \binom{t}{k} \alpha^k (1 - \alpha)^{t-k}.$$

On peut démontrer la propriété suivante : la valeur moyenne de l'action à l'époque  $t$  est donnée par

$$E(S_t) = S_0 \cdot (\alpha u + (1 - \alpha) d)^t.$$

En effet, on a

$$\begin{aligned} E(S_t) &= \sum_{k=0}^t S_0 \cdot u^k d^{t-k} \cdot \binom{t}{k} \alpha^k (1 - \alpha)^{t-k} \\ &= S_0 \cdot \sum_{k=0}^t \binom{t}{k} (\alpha u)^k ((1 - \alpha) d)^{t-k} \end{aligned}$$

qui conduit à la relation annoncée grâce à la formule du binôme de Newton.

Remarquons que cette propriété est la généralisation, en moyenne, pour le cas aléatoire de la formule déterministe  $S_t = S_0 \cdot (1 + i)^t$ .

<sup>75</sup> On consultera l'annexe 2 pour le développement de cette notion et pour les propriétés de cette variable aléatoire.

## Modèle continu

Le mode d'évolution de la valeur de l'action dans le modèle binomial est du type *promenade aléatoire* : à chaque transition, deux mouvements sont possibles (baisse et hausse) avec une probabilité inchangée. Lorsque la durée entre chaque transition tend vers 0, ce type de suite aléatoire converge vers un *mouvement brownien standard*<sup>76</sup> (MBS). Rappelons qu'il s'agit là d'un processus stochastique  $w_t$  (variable aléatoire fonction du temps) obéissant aux hypothèses suivantes :

- $w_0 = 0$  ;
- $w_t$  est un processus à accroissements indépendants : si  $s < t < u$ , alors  $w_u - w_t$  est indépendant de  $w_t - w_s$  ;
- $w_t$  est un processus à accroissements stationnaires : les variables aléatoires  $w_{t+h} - w_t$  et  $w_h$  sont identiquement distribuées ;
- quel que soit  $t$ , la variable aléatoire  $w_t$  est distribuée suivant une loi normale de moyenne nulle et d'écart type  $\sqrt{t}$  :

$$f_{w_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}.$$

La première utilisation de ce processus pour modéliser l'évolution de la valeur d'un actif financier est due à L. Bachelier<sup>77</sup>. Il suppose que la valeur du titre à l'instant  $t$  est une fonction du premier degré du MBS :  $S_t = a + bw_t$ . D'après la définition ci-dessus,  $a$  est la valeur du titre en  $t = 0$  et  $b$  est une mesure de la volatilité  $\sigma$  du titre par unité de temps. La relation utilisée était donc :

$$S_t = S_0 + \sigma \cdot w_t.$$

Les inconvénients de cette approche sont de deux ordres :

- une même variation absolue (10 EUR par exemple) correspond à des variations de rendement (et donc relatives) très différentes suivant le niveau du cours (20 % pour une cotation de 50 EUR et 5 % pour une valeur de 200 EUR) ;
- la variable aléatoire  $S_t$  suit une loi normale de moyenne  $S_0$  et d'écart type  $\sigma\sqrt{t}$  ; cette modélisation permet donc des prix négatifs.

C'est pourquoi P. Samuelson<sup>78</sup> a proposé la modélisation suivante. Durant le petit intervalle de temps  $[t ; t + dt]$ , le return (et non le prix) de l'action évolue suivant un *processus d'Itô*

$$\frac{S_{t+dt} - S_t}{S_t} = \frac{dS_t}{S_t} = E_R \cdot dt + \sigma_R \cdot dw_t,$$

où le terme non aléatoire (tendance) est proportionnel au return espéré et le terme stochastique fait intervenir la volatilité par unité de temps de ce return. Ce modèle est un *mouvement brownien géométrique*.

<sup>76</sup> L'annexe 2 détaille les résultats, raisonnements et propriétés de ces processus stochastiques.

<sup>77</sup> BACHELIER L., *Théorie de la spéculation*, Gauthier-Villars, 1900. Il fallut attendre plusieurs dizaines d'années pour que son raisonnement soit admis et amélioré.

<sup>78</sup> SAMUELSON P., Mathematics on speculative price, *SIAM Review*, Vol. 15, n° 1, 1973.

### Illustration

La figure 3.34 représente une trajectoire (évolution au cours du temps) simulée d'un tel processus sur 1000 très courtes périodes avec les valeurs  $E_R = 0,1$  et  $\sigma_R = 0,02$ , à partir d'une valeur de départ  $S_0 = 100$ .



Figure 3.34 – Mouvement brownien géométrique ♣

On peut établir une première propriété dans le contexte de cette modélisation : le processus stochastique  $S_t$  modélisant l'évolution de la valeur de l'action peut s'écrire

$$S_t = S_0 \cdot \exp \left[ \left( E_R - \frac{\sigma_R^2}{2} \right) \cdot t + \sigma_R \cdot w_t \right],$$

ce qui exprime que  $S_t$  suit une *distribution log-normale* (ne pouvant prendre que des valeurs positives).

En effet, en appliquant la formule d'Itô<sup>79</sup> à la fonction  $f(x, t) = \ln x$  avec  $x = S_t$ , on obtient

$$\begin{aligned} d(\ln S_t) &= \left[ 0 + \frac{1}{S_t} E_R S_t - \frac{1}{2 S_t^2} \sigma_R^2 S_t^2 \right] \cdot dt + \frac{1}{S_t} \sigma_R S_t \cdot dw_t \\ &= \left( E_R - \frac{\sigma_R^2}{2} \right) \cdot dt + \sigma_R \cdot dw_t \end{aligned}$$

Cette équation se résout en

<sup>79</sup> Voir l'annexe 2.

$$\ln S_t = C^* + \left( E_R - \frac{\sigma_R^2}{2} \right) \cdot t + \sigma_R \cdot w_t.$$

La constante d'intégration  $C^*$  vaut bien sûr  $\ln S_0$  et le passage à l'exponentielle donne la formule annoncée.

On déduit aisément de là les moments de la variable aléatoire  $S_t$  :

$$\begin{aligned} E(S_t) &= S_0 \cdot e^{E_R t} \\ \text{var}(S_t) &= S_0^2 \cdot e^{2E_R t} (e^{\sigma_R^2 t} - 1) \end{aligned}$$

la première de ces relations montrant que le rendement moyen  $E(S_t/S_0)$  de cet actif sur l'intervalle  $[0 ; t]$  est équivalent à une capitalisation au taux instantané  $E_R$ .

On peut établir une seconde propriété, relative au rendement instantané du titre sur l'intervalle  $[0 ; t]$  : celui-ci obéit à une loi normale de moyenne et d'écart type donnés par

$$\left( E_R - \frac{\sigma_R^2}{2} ; \frac{\sigma_R}{\sqrt{t}} \right).$$

Ce résultat peut paraître paradoxal, la moyenne du rendement n'étant pas égal à  $E_R$  ; ceci provient de la structure de processus stochastique et n'est pas incompatible avec l'intuition puisqu'on a  $E(S_t) = S_0 \cdot e^{E_R t}$ .

Pour établir cette propriété, en notant  $\delta_t$  le processus stochastique du rendement instantané, on peut écrire

$$S_t = S_0 \cdot e^{\delta_t \cdot t},$$

c'est-à-dire, d'après la propriété précédente,

$$\begin{aligned} \delta_t &= \frac{1}{t} \cdot \ln \left( \frac{S_t}{S_0} \right) \\ &= \left( E_R - \frac{\sigma_R^2}{2} \right) + \sigma_R \cdot \frac{w_t}{t} \end{aligned}$$

ce qui établit la propriété.

## CHAPITRE 4

### OBLIGATIONS

#### 4.1. CARACTERISTIQUES ET EVALUATION

Pour un investisseur, une *obligation* est un actif financier émis par une institution publique ou une entreprise privée, correspondant à un prêt qui donne droit, en retour, à la perception d'intérêts (appelés coupons) et au remboursement à l'échéance du prêt. Ce titre est négociable et son prix à l'émission, sa valeur de remboursement, le montant de ses coupons et sa durée de vie sont généralement connus et fixés a priori.

##### 4.1.1. Définitions

Une obligation est caractérisée par différents éléments :

- a) La valeur *nominale*  $VN$  d'une obligation est le montant indiqué sur le titre et qui permet, avec le *taux nominal* de l'obligation, de déterminer le montant des coupons.
- b) Le *prix* de l'obligation sera noté  $P$ . Il peut s'agir du *prix à l'émission* (en  $t = 0$ ) ou à un instant  $t$  ultérieur quelconque. Le *prix à l'échéance* est bien sûr identique à la *valeur de remboursement*  $R$  initialement prévu.
- c) Les *coupons*  $C_t$  constituent l'intérêt versé par l'émetteur. Ils sont payés à différentes époques que nous supposerons régulièrement espacées et annuelles ( $t = 1, 2, \dots, T$ ).
- d) La *durée* ou *maturité*  $T$  représente l'espace de temps séparant l'instant d'émission et l'époque de remboursement du titre.

Les flux financiers associés à une obligation sont donc

- de la part de l'acquéreur, le paiement de son prix ; il s'agit soit du prix à l'émission versé à l'émetteur, soit du cours de l'obligation versé à un vendeur quelconque à une époque postérieure à l'émission ;
- de la part de l'émetteur, le paiement des coupons à partir de l'époque d'acquisition ainsi que la valeur de remboursement à l'échéance.

Le prix à l'émission, la valeur nominale et la valeur de remboursement ne sont pas nécessairement égaux : il peut exister des primes – positives ou négatives – à l'émission et/ou au remboursement.

Les obligations décrites ci-dessus sont celles que nous étudierons dans ce chapitre ; on les appelle *obligations à taux fixe*. Il existe un grand nombre de variations par rapport à ce modèle simple d'obligation.

Ainsi, il se peut qu'aucun coupon ne soit versé pendant la durée de l'obligation, le rendement étant comme seule origine la différence entre le prix à l'émission et la valeur de remboursement. On parle alors d'*obligation zéro-coupon*<sup>80</sup>. Un tel titre est équivalent à un placement à un taux fixe.

Il existe également des obligations plus compliquées que celles décrites ci-dessus, par exemple<sup>81</sup>

- les *obligations à taux variable*, pour lesquelles le montant de chaque coupon est périodiquement déterminé en fonction d'un paramètre tel qu'un indice ;
- les *obligations à charnière*, qui autorisent le remboursement avant la date d'échéance ;
- les *obligations à lots*, pour lesquels l'émetteur (public) rembourse certaines obligations chaque années par tirage au sort ;
- les *obligations convertibles* en actions ;
- ...

#### 4.1.2. Rendement d'une obligation

Le rendement d'une obligation peut bien sûr se calculer par le taux nominal  $r_n$  défini comme étant le rapport du montant de coupon à la valeur nominale

$$r_n = \frac{C}{VN},$$

définition qui n'a guère de sens que si les différents coupons ont tous la même valeur. On peut également adapter cette définition en remplaçant le dénominateur par le prix de l'obligation à un instant particulier. Le taux nominal est d'un intérêt limité puisqu'il ne fait aucunement intervenir la durée de vie de l'obligation ; ainsi son utilisation pour comparer deux obligations n'a guère de sens.

Pour un intervalle de temps fixé (une année par exemple), on pourrait envisager d'utiliser un taux de rendement équivalent au return d'une action

$$\frac{P_t + C_t - P_{t-1}}{P_{t-1}}.$$

Cette notion n'est guère utilisée en pratique.

<sup>80</sup> Une obligation peut donc être d'une certaine manière considérée comme une superposition d'obligations zéro-coupon.

<sup>81</sup> On consultera par exemple

COLMANT B., DELFOSSE V. et ESCH L., *Obligations, les notions financières essentielles*, Larcier, 2002.  
FABOZZI J.F., *Bond markets, analysis and strategies*, Prentice-Hall, 2000.

## Taux de rendement actuariel à l'émission

Le *taux de rendement actuariel à l'émission* – ou plus simplement le *taux actuariel* –  $r$  d'une obligation est celui pour lequel on a égalité entre la valeur actualisée des coupons et de la valeur de remboursement d'une part et le prix à l'émission d'autre part :

$$P = \sum_{t=1}^T C_t(1+r)^{-t} + R(1+r)^{-T}.$$

### Illustration

Considérons par exemple une obligation de durée égale à 6 ans, de valeur nominale 100, émise à 98 et remboursée à 105 (primes à l'émission et au remboursement respectivement égales à 2 et 5) et ayant un taux nominal égal à 10 %. L'équation définissant son taux actuariel est donc

$$98 = \frac{10}{1+r} + \frac{10}{(1+r)^2} + \frac{10}{(1+r)^3} + \frac{10}{(1+r)^4} + \frac{10}{(1+r)^5} + \frac{10+105}{(1+r)^6}.$$

Cette équation (du 6<sup>ème</sup> degré en la variable  $r$ ) se résout numériquement et donne  $r = 0,111044$ , c'est-à-dire approximativement  $r = 11,1 \%$ . ♣

Le taux actuariel d'une obligation zéro-coupon est bien entendu celui d'un placement sans risque ; il est défini par

$$P = R(1+r)^{-T}.$$

Celui d'une obligation émise et remboursable au pair ( $P = VN = R$ ) et dont les coupons sont égaux ( $C_t = C$  pour tout  $t$ ) est égal au taux nominal :  $r = r_n$ . En effet, pour ce type particulier d'obligation, on a

$$\begin{aligned} P &= \sum_{t=1}^T C(1+r)^{-t} + P(1+r)^{-T} \\ &= C \frac{(1+r)^{-1} - (1+r)^{-T-1}}{1 - (1+r)^{-1}} + P(1+r)^{-T} \\ &= C \frac{1 - (1+r)^{-T}}{r} + P(1+r)^{-T} \end{aligned}$$

d'où on déduit aisément  $r = C/P = r_n$ .

## Taux de rendement actuariel à un instant quelconque

Le taux de rendement actuariel tel que défini ci-dessus est calculé à l'émission de l'obligation ; on l'appelle parfois *taux ex ante*. Ce faisant, on suppose donc que ce taux va rester constant pendant toute la vie du titre (et indépendant de l'échéance).



Un grand principe de la mathématique financière (le *principe d'équivalence*) exprime que ce taux est indépendant de l'époque où sont « ramenés » les différents mouvements financiers.

### Illustration

Pour l'exemple du paragraphe précédent (obligation de valeur nominale 100, émise à 98 et remboursée à 105, versant un coupon annuel de 10 à la fin de chacune des 6 années de la vie du titre) dont le taux de rendement actuariel est de 11,1 %, si on examine la valeur acquise par exemple à l'échéance, on a

- a) pour l'investissement,  $98 \cdot (1 + r)^6$  ;
- b) pour les flux financiers générés

$$10 \cdot [(1 + r)^5 + (1 + r)^4 + (1 + r)^3 + (1 + r)^2 + (1 + r) + 1] + 105.$$

L'égalité de ces deux quantités se réalise également pour  $r = 11,1 \%$ . ♣

Si nous nous plaçons maintenant à un instant  $t$  quelconque entre 0 et  $T$  et que l'on connaît l'évolution du taux du marché entre 0 et  $t$ , le taux de rendement actuariel à l'instant<sup>82</sup>  $t$  – que nous noterons<sup>83</sup>  $r(t)$  – est celui pour lequel il y a égalité entre

- a) la valeur acquise en  $t$  de l'investissement, calculée à ce taux  $r(t)$  ;
- b) la somme de
  - la valeur acquise en  $t$  des coupons échus, réinvestis aux taux en vigueur observés entre 0 et  $t$  ;
  - la valeur actualisée en  $t$  des flux financiers générés, ultérieurs à  $t$ , calculée grâce au taux du marché à l'instant  $t$ .

### Illustration

Reprenons le même exemple que ci-dessus. Supposons que l'on se situe à l'instant juste postérieur au paiement du troisième coupon ( $t = 3$ ) et que le taux du marché est resté à 11,1 % pendant les deux premières années et est ensuite passé à 12 %. La définition ci-dessus donne l'équation définissant le taux de rendement actuariel à l'instant  $t = 3$  :

$$\begin{aligned} 98 \cdot (1 + r(3))^3 &= (10 \cdot 1,111 \cdot 1,12 + 10 \cdot 1,12 + 10) + \left( \frac{10}{1,12} + \frac{10}{1,12^2} + \frac{115}{1,12^3} \right) \\ &= 35,33 + 98,76 \\ &= 134,09 \end{aligned}$$

ce qui donne  $r(3) = 11,02 \%$ .

On remarquera évidemment que si le taux d'intérêt était resté constant (et égal à 11,1 %) pendant les trois premières années, le calcul ci-dessus aurait conduit à  $r(3) = 11,1 \%$ , résultat en cohérence avec le principe d'équivalence. ♣

<sup>82</sup> On l'appellera parfois taux *ex post*.

<sup>83</sup>  $r(0) = r$  est bien sûr le taux de rendement actuariel à l'émission.

Cette illustration met en évidence le phénomène de risque d'une obligation, lié à une fluctuation des taux d'intérêt. Il sera étudié plus en détail au § 4.2.1.

### Intérêts à bonifier

Lorsque l'acquisition d'une obligation a lieu entre deux époques de versement de coupon, l'acheteur payera, outre la valeur de l'obligation à cet instant, la partie du coupon à venir calculée au prorata de la durée écoulée depuis le paiement du dernier coupon. Le vendeur a en effet droit à l'intérêt partiel relatif à la période courant du dernier coupon échu à l'instant de la transaction. Ce principe s'appelle système des *intérêts à bonifier* et le prix effectivement payé est le *dirty price*, par opposition au *clean price* qui représente simplement le cours de l'obligation au moment de la transaction.

Considérons une obligation de durée  $T$  et un instant non entier  $t + \theta$  ( $t$  entier et  $0 \leq \theta < 1$ )

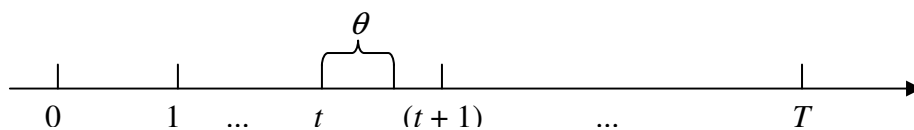


Figure 4.1 – Intérêts à bonifier

Le dirty price  $P^{(d)}$  de l'obligation est lié au clean price  $P$  par la relation

$$P_{t+\theta}^{(d)} = P_{t+\theta} + \theta \cdot C_{t+1}.$$

Le système des intérêts à bonifier influence le taux de rendement car l'acheteur préfinance (au moment de la transaction) une partie du prochain coupon. On aura donc une légère diminution du taux actuariel ; celle-ci sera bien entendu d'autant plus faible que l'instant de la transaction sera proche (avant ou après) d'une époque de versement de coupon. C'est pourquoi on parle d'*effet feston*.

Le taux de rendement actuariel se calcule classiquement : il est défini implicitement par la relation

$$P_{t+\theta} + \theta \cdot C_{t+1} = \sum_{k=t+1}^T C_k (1+r)^{-k-t+\theta} + R(1+r)^{-T-t+\theta}.$$

### Illustration

Considérons une obligation, émise et remboursable au pair, dont le prix reste constamment égal à 100 et présentant des coupons annuels de 10 pendant 5 ans, dont le taux actuariel à l'émission est évidemment égal à 10 %. On obtient le graphe de la figure 4.2.♣

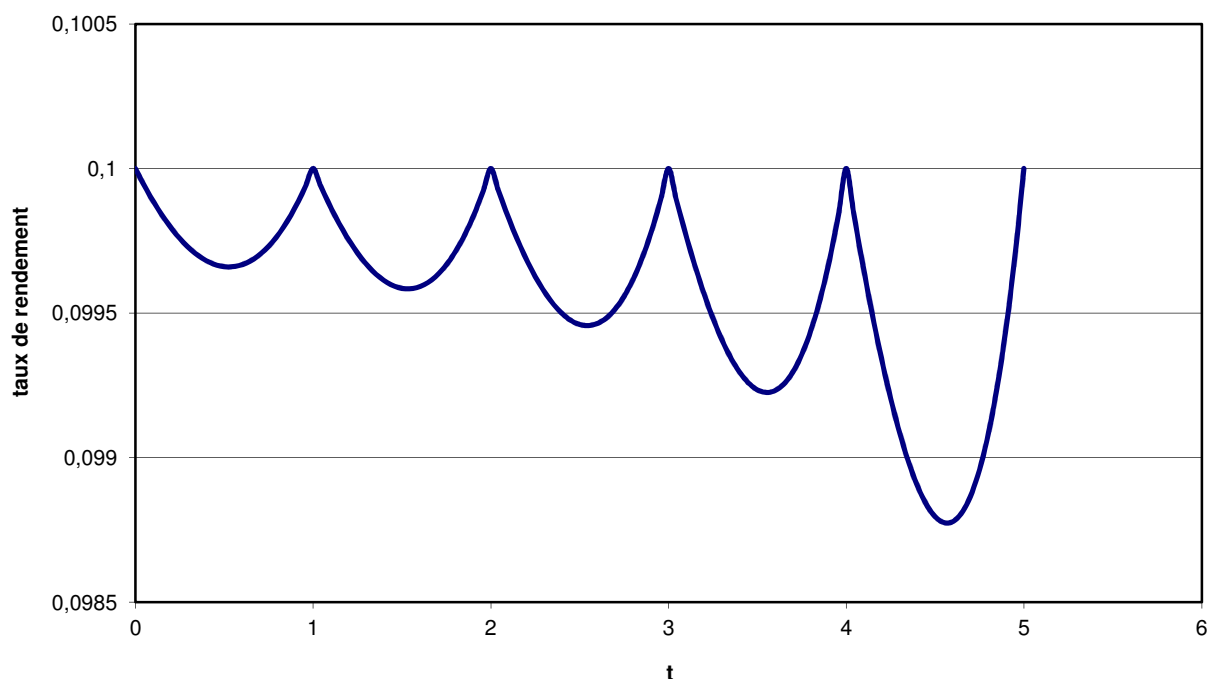


Figure 4.2 – Effet feston

#### 4.1.3. Evaluation d'une obligation

L'évaluation de la valeur d'une obligation à un instant quelconque (c'est-à-dire son cours) est la valeur actualisée des flux financiers futurs générés (coupons résiduels et remboursement final), le taux d'actualisation étant le taux exigé par le marché<sup>84</sup> pour les obligations de même nature et de même durée résiduelle (nous le noterons  $r_m$ ).

Au lendemain du paiement du  $t$ -ème coupon, l'évaluation se fait grâce à la relation

$$P_t = \sum_{k=t+1}^T C_k (1 + r_m)^{-k+t} + R(1 + r_m)^{-T+t}.$$

A un instant quelconque  $t + \theta$  ( $t$  entier et  $0 \leq \theta < 1$ ), on a, en prenant les intérêts à bonifier en considération,

$$P_{t+\theta} = -\theta \cdot C_{t+1} + \sum_{k=t+1}^T C_k (1 + r_m)^{-k+t+\theta} + R(1 + r_m)^{-T+t+\theta}.$$

<sup>84</sup> Qui résulte de la loi de l'offre et de la demande.

## 4.2. OBLIGATIONS ET RISQUE FINANCIER

### 4.2.1. Sources de risque

L'obligation est un actif financier généralement considéré comme peu risqué si on le compare aux actions ou, davantage encore, aux produits dérivés à fort effet de levier. Il n'en demeure pas moins vrai que la détention d'obligations présente certains risques.

#### *Risque de défaut*

On peut ainsi citer le *risque de défaut*, qui correspond essentiellement au cas où l'émetteur ne respecte pas ses engagements : paiement des coupons avec retard, faillite et donc non-paiement des coupons et non remboursement à l'échéance.

Ce risque de défaut est bien sûr non quantifiable. Notons néanmoins que les professionnels du marché obligataire attachent une importance considérable aux ratings attribués aux émetteurs par certaines agences (Moody's, Standard & Poors, ...), et ce en vue de « quantifier » la qualité de ces émetteurs et donc « d'estimer » leur propension à faire défaut. Ceci explique que pour une même maturité, dans une même classe, le coupon pour un émetteur de rating AAA (le meilleur rating) soit inférieur au coupon qu'un émetteur de rating B propose ; on comprend ici intuitivement la notion de prime de risque. Par ailleurs, le rating influencera directement un autre risque non quantifiable : la liquidité dudit titre.

#### *Risque de fluctuation des taux*

Un autre type de risque associé aux obligations – quantifiable celui-ci – est le risque lié aux fluctuations des taux d'intérêt.

Pour certaines obligations particulières (dites associées à un call), l'émetteur a le droit de rembourser l'obligation avant son échéance. Si le remboursement a lieu juste après une chute des taux d'intérêt, il ne sera pas possible pour l'investisseur de réinvestir le montant ainsi remboursé à un taux équivalent à celui du titre ; il en résulte donc un manque à gagner dans son chef.

Mais, comme nous l'avons dit en introduction (§ 4.1.1), nous n'envisagerons pas ici ce type de produit et nous limiterons donc à l'analyse du risque d'une obligation « classique ». Deux aspects, dans un certain sens contradictoires, de ce risque peuvent être considérés.

Il y a tout d'abord le *risque de réinvestissement* : en cas de variation des taux du marché, les coupons (voire le montant du remboursement lui-même) seront réinvestis à un taux différent. Une augmentation (resp. diminution) du taux d'intérêt sera donc ici favorable (resp. défavorable) à l'investisseur.

Par ailleurs, on se trouve également face à un *risque de réalisation* de l'obligation, en cas de vente de celle-ci avant son échéance. Le prix de vente s'évalue par la valeur actualisée (au

taux en vigueur sur le marché à cet instant) des coupons à venir et de la valeur de remboursement. Dans ce cas, une augmentation (resp. diminution) du taux d'intérêt sera bien sûr défavorable (resp. favorable) à l'investisseur.

### *Illustration*

Considérons une obligation de valeur nominale égale à 1000, sans prime à l'émission ni au remboursement ( $VN = P = R = 1000$ ), de durée 8 ans et de coupon constant égal à 100. On sait que, dans ce cas, le taux de rendement actuariel à l'émission  $r$  est égal au taux de rendement nominal, c'est-à-dire 10 %. Supposons par ailleurs que le taux du marché passe, à la fin de la deuxième année, de 10 à 12 %.

Que se passe-t-il si l'investisseur désire revendre cette obligation après 4 ans ? Pour analyser la situation, nous allons déterminer de taux de rendement actuariel  $r(4)$  à l'instant 4.

Le réinvestissement des coupons (à 10 % pendant les deux premières années et à 12 % ensuite) donne, en  $t = 4$ , une valeur acquise égale à<sup>85</sup>

$$A_4 = 100 \cdot 1,1 \cdot 1,12^2 + 100 \cdot 1,12^2 + 100 \cdot 1,12 + 100 = 475,424.$$

Par ailleurs, la valeur de revente en  $t = 4$  est donnée par la valeur actualisée (à 12 %) des coupons futurs et du remboursement<sup>86</sup> :

$$B_4 = \frac{100}{1,12} + \frac{100}{1,12^2} + \frac{100}{1,12^3} + \frac{1100}{1,12^4} = 939,253.$$

Le taux de rendement ex post est défini par la relation

$$P \cdot (1 + r(4))^4 = A_4 + B_4,$$

c'est-à-dire, pour notre exemple,

$$1000 \cdot (1 + r(4))^4 = 1414,677,$$

ce qui donne finalement  $r(4) = 9,06$  %. Ce taux qui est inférieur à celui que l'on trouverait en remplaçant  $A_4 + B_4$  par 1464,1 puisqu'il s'agirait de 10 %, taux de rendement actuariel de l'obligation à son émission.

En tenant les mêmes raisonnements pour un investisseur qui revend son obligation après 7 ans, on obtient pour valeur de réinvestissement  $A_7 = 1005,376$  et pour valeur de réalisation  $B_7 = 982,143$  (au lieu de respectivement de 948,717 et 1000 s'il n'y avait pas eu de variation de taux) et le taux de rendement ex post est alors donné par  $r(7) = 10,31$  % (au lieu de 10 %). Cette fois, la combinaison des deux effets s'inverse et devient favorable à l'investisseur. ♣

<sup>85</sup> Alors que sans modification de taux, on aurait obtenu  $A_4 = 464,1$ , ce qui met en évidence le caractère favorable d'une hausse des taux.

<sup>86</sup> Alors qu'avec le taux initial de 10 %, on aurait bien entendu obtenu  $B_4 = 1000$  ; cette fois, la hausse du taux d'intérêt est défavorable à l'investisseur.

#### 4.2.2. Duration

##### *Définition*

Compte tenu des constatations faites dans l'illustration ci-dessus, se pose la question d'évaluer l'époque avant laquelle une augmentation (resp. diminution) du taux d'intérêt est défavorable (resp. favorable) à l'investisseur et après laquelle elle lui devient favorable (resp. défavorable).

Cette question trouve sa réponse dans la notion de *duration*<sup>87</sup>. Considérons une obligation au moment de son émission<sup>88</sup>, de taux actuariel  $r$  ; sa duration, notée  $D$ , se définit par

$$D = \frac{\sum_{t=1}^T t \cdot C_t(1+r)^{-t} + T \cdot R(1+r)^{-T}}{\sum_{t=1}^T C_t(1+r)^{-t} + R(1+r)^{-T}}.$$

##### *Illustration*

Reprenons les données déjà utilisées précédemment : une obligation émise et remboursée au pair, de valeur égale à 1000, de durée 8 ans et de coupon constant égal à 100. Son taux actuariel à l'émission est de 10 %.

La duration de ce titre se calcule comme suit :

$$D = \frac{100 \cdot [1 \cdot 1,1^{-1} + 2 \cdot 1,1^{-2} + \dots + 8 \cdot 1,1^{-8}] + 1000 \cdot 8 \cdot 1,1^{-8}}{100 \cdot [1,1^{-1} + 1,1^{-2} + \dots + 1,1^{-8}] + 1000 \cdot 1,1^{-8}} = 5,868 \clubsuit$$

On introduit également la *duration modifiée*, ou encore le coefficient de *sensibilité*, qui est défini par

$$D_m = \frac{D}{1+r}.$$

##### *Remarque*

A côté de cette notion « simple » de duration, une notion plus évoluée de duration stochastique sera introduite au § 4.6.♣

<sup>87</sup> Cette notion a été introduite dans MACAULEY F., *Some theoretical problems suggested by the movements of interest rates, bond yields and stock prices in the United States since 1856*, New York : National Bureau of Economic Research, 1938, pp.44-53.

<sup>88</sup> On peut aisément généraliser cette notion à un moment quelconque, à condition de supposer que le taux  $r$  reste constant.

## Interprétations

La notion de duration possède plusieurs interprétations, dont voici les deux principales.

Elle peut tout d'abord être considérée comme la durée de vie moyenne de l'obligation, les coefficients de pondération de cette moyenne étant proportionnels aux valeurs actualisées au taux  $r$  des flux financiers (coupons et remboursement) aux époques correspondantes. De fait, en notant que le dénominateur n'est autre que le prix  $P$  à l'émission, on peut écrire

$$\begin{aligned} D &= \sum_{t=1}^T t \frac{C_t (1+r)^{-t}}{P} + T \frac{R(1+r)^{-T}}{P} \\ &= 1 \cdot \frac{C_1 (1+r)^{-1}}{P} + \dots + (T-1) \cdot \frac{C_{T-1} (1+r)^{-(T-1)}}{P} + T \cdot \frac{(C_T + R)(1+r)^{-T}}{P} \end{aligned}$$

On peut par ailleurs donner de la duration une interprétation en terme de sensibilité du prix au taux d'intérêt. Lorsque deux variables économiques  $x$  et  $y$  sont liées par une relation  $y = y(x)$ , une façon de décrire de quelle manière  $y$  dépend de  $x$  est d'utiliser la dérivée  $y'(x)$ .

La relation définissant le taux de rendement actuariel

$$P = \sum_{t=1}^T C_t (1+r)^{-t} + R(1+r)^{-T},$$

définit une fonction  $P(r)$ . La dérivée de cette dernière peut s'écrire

$$\begin{aligned} P'(r) &= -\sum_{t=1}^T t \cdot C_t (1+r)^{-t-1} - T \cdot R(1+r)^{-T-1} \\ &= -(1+r)^{-1} \left\{ \sum_{t=1}^T t \cdot C_t (1+r)^{-t} + T \cdot R(1+r)^{-T} \right\} \\ &= -(1+r)^{-1} D \cdot P(r) \end{aligned}$$

qui montre comment la duration intervient dans cette mesure de variation de  $P$  en fonction de  $r$ . Cette relation permet encore d'écrire

$$D = -(1+r) \frac{P'(r)}{P(r)} \quad \text{ou} \quad D_m = -\frac{P'(r)}{P(r)}$$

En considérant une petite variation de taux  $\Delta r$  au lieu d'une variation infinitésimale et en notant

$$\Delta P = P(r + \Delta r) - P(r),$$

la relation précédente s'écrit

$$\begin{aligned}
D &= - (1+r) \frac{\frac{\Delta P}{P}}{\frac{\Delta r}{1+r}} \\
&= - \frac{\frac{\Delta P}{P}}{\frac{\Delta(1+r)}{1+r}} \\
&= \varepsilon(P/1+r)
\end{aligned}$$

car la variation  $\Delta r$  est la même que  $\Delta(1+r)$ .

Il s'agit là d'une notion classique en économie, l'*élasticité*. De manière générale, l'élasticité  $\varepsilon(y/x)$  de  $y$  par rapport à  $x$

$$\varepsilon(y/x) = \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}}$$

s'interprète de la manière suivante : si  $x$  augmente de 1 %,  $y$  varie (augmente ou diminue suivant le signe de  $\Delta y$ ) de  $|\varepsilon(y/x)|$  %.

La duration apparaît donc comme l'élasticité, au signe près, du prix de l'obligation par rapport à  $(1+r)$  et une augmentation (relative) de  $(1+r)$  de 1 % provoquera une diminution (relative) de  $P$  de  $D$  %.

### *Illustration*

Reprenons les données utilisées ci-dessus (une obligation émise et remboursée au pair, de valeur égale à 1000, de durée 8 ans et de coupon constant égal à 100 ;  $r = 10$  % et  $D = 5,868$ ).

Si l'augmentation relative de  $(1+r)$  est de 1 % (c'est-à-dire que  $(1+r)$  passe de 1,1 à  $1,1 \cdot 1,01 = 1,111$ , ou encore que  $r$  passe de 10 % à 11,1 %), il est facile de voir par le calcul direct que le prix passe de  $P(10 \%) = 1000$  à

$$P(11,1 \%) = \sum_{t=1}^8 100 \cdot 1,111^{-t} + 1000 \cdot 1,111^{-8} = 943,59.$$

Celui-ci diminue donc de 5,64 %, la différence par rapport à la duration provenant simplement du fait que nous n'avons pas utilisé un accroissement infinitésimal.♣

### ***Duration d'obligations particulières***

La duration d'une obligation zéro-coupon est égale à sa maturité. En effet, elle est donnée par



$$D = \frac{T \cdot R(1+r)^{-T}}{R(1+r)^{-T}} = T.$$

Dans tous les autres cas, la duration est strictement inférieure à la maturité ; il s'agit en effet d'une moyenne pondérée des nombres 1, 2, ...,  $T$ .

Une *obligation perpétuelle* est un titre donnant des coupons jusqu'à l'infini (et aucun remboursement) ; nous supposons ici que les coupons sont de même montant  $C$ .

La relation définissant le taux de rendement actuariel prend ici la forme<sup>89</sup>

$$P = C \sum_{t=1}^{\infty} (1+r)^{-t} = \frac{C}{r}.$$

Ceci permet le calcul de la duration d'une obligation perpétuelle :

$$D = -(1+r) \frac{P'(r)}{P(r)} = -(1+r) \frac{-\frac{C}{r^2}}{\frac{C}{r}} = \frac{1+r}{r}.$$

### ***Duration et caractéristiques d'une obligation***

La duration est une fonction des paramètres  $r_n$  (taux nominal),  $r$  (taux actuariel) et  $T$  (maturité). Signalons, sans justification rigoureuse<sup>90</sup>, la manière dont elle dépend de ces trois paramètres.

La duration est une fonction décroissante de  $r_n$ , les deux paramètres  $r$  et  $T$  étant fixés. De la même manière, elle est une fonction décroissante de  $r$ , les deux paramètres  $r_n$  et  $T$  étant fixés.

La dépendance de  $D$  vis-à-vis de  $T$  ( $r$  et  $r_n$  étant fixés) est plus complexe. On a déjà signalé que la duration d'une obligation zéro-coupon est égale à son échéance et que celle d'une obligation perpétuelle est égale à  $(1+r)/r$ . Pour une obligation remboursable au pair et de coupons constants, on peut montrer que

- si elle est émise au pair ou au-dessus du pair,  $D$  croît avec  $T$ , de manière concave avec

$$\lim_{T \rightarrow \infty} D(T) \begin{cases} = \\ < \end{cases} \frac{1+r}{r}$$

selon que l'obligation est émise au pair ou au-dessus du pair ;

- si elle est émise au-dessous du pair,  $D$  présente une croissance en fonction de  $T$  jusqu'à un certain moment, pour décroître ensuite.

<sup>89</sup> En utilisant la notion de *série géométrique*.

<sup>90</sup> Les raisonnements ne sont pas très compliqués mais assez fastidieux.

### ***Immunisation d'une obligation***

La propriété la plus importante de la duration est que le risque de réinvestissement et le risque de réalisation s'équilibrent si l'obligation est revendue à l'instant  $D$  : une variation de taux d'intérêt n'a plus d'influence sur le taux de rendement actuariel ex post. On dit que l'obligation est *immunisée* à l'horizon  $D$  contre le risque de taux d'intérêt.

En fait, ce résultat n'est exact que si les fluctuations de taux d'intérêt sont de faible amplitude et si elles se produisent juste après l'achat du titre.

Supposons donc qu'une petite variation  $\Delta r$  du taux du marché (c'est-à-dire du taux actuariel à l'émission) se produit juste après l'émission du titre : ce taux passe de  $r$  à  $r + \Delta r$ .

Plaçons-nous à une époque quelconque  $t + \theta$  ( $t$  entier et  $0 \leq \theta < 1$ ) de la vie de l'obligation. Nous devons donc prouver que si le titre est immunisé à l'horizon  $t + \theta$  (c'est-à-dire si le taux de rendement actuariel  $r(t + \theta)$  est égal au taux  $r$  à l'émission), alors  $t + \theta$  est égal à la *duration*  $D$  de l'obligation.

Les valeurs de réinvestissement et de réalisation à l'instant  $t + \theta$  sont respectivement données par

$$A_{t+\theta} = \sum_{h=1}^t C_h (1 + r + \Delta r)^{t+\theta-h}$$
$$B_{t+\theta} = \sum_{k=t+1}^T C_k (1 + r + \Delta r)^{-k+t+\theta} + R(1 + r + \Delta r)^{-T+t+\theta}$$

et leur somme est donc égale à

$$A_{t+\theta} + B_{t+\theta} = (1 + r + \Delta r)^{t+\theta} \left[ \sum_{k=1}^T C_k (1 + r + \Delta r)^{-k} + R(1 + r + \Delta r)^{-T} \right].$$

Exprimons les puissances apparaissant dans cette expression grâce à un développement de Taylor jusqu'au premier degré :

$$(1 + r + \Delta r)^m = (1 + r)^m + m(1 + r)^{m-1} \Delta r + O((\Delta r)^2).$$

Ceci permet d'écrire

$$\begin{aligned}
& A_{t+\theta} + B_{t+\theta} \\
&= \left[ (1+r)^{t+\theta} + (t+\theta)(1+r)^{t+\theta-1} \Delta r + O((\Delta r)^2) \right] \\
&\quad \cdot \left\{ \sum_{k=1}^T C_k \left[ (1+r)^{-k} - k(1+r)^{-k-1} \Delta r + O((\Delta r)^2) \right] + R \left[ (1+r)^{-T} - T(1+r)^{-T-1} \Delta r + O((\Delta r)^2) \right] \right\} \\
&= (1+r)^{t+\theta} \left[ 1 + \frac{t+\theta}{1+r} \Delta r + O((\Delta r)^2) \right] \\
&\quad \cdot \left\{ \left[ \sum_{k=1}^T C_k (1+r)^{-k} + R(1+r)^{-T} \right] - \left[ \sum_{k=1}^T k \cdot C_k (1+r)^{-k} + T \cdot R(1+r)^{-T} \right] \Delta r + O((\Delta r)^2) \right\} \\
&= (1+r)^{t+\theta} \left[ 1 + \frac{t+\theta}{1+r} \Delta r + O((\Delta r)^2) \right] \cdot \left\{ P - \frac{D \cdot P}{1+r} \Delta r + O((\Delta r)^2) \right\} \\
&= (1+r)^{t+\theta} P \cdot \left[ 1 + \frac{t+\theta-D}{1+r} \Delta r + O((\Delta r)^2) \right]
\end{aligned}$$

Par définition du taux actuariel à l'instant  $t$ , on a également

$$A_{t+\theta} + B_{t+\theta} = P \cdot (1+r(t+\theta))^{t+\theta}.$$

Ceci n'est possible que si le coefficient

$$\frac{t+\theta-D}{1+r},$$

de  $\Delta r$  est nul, c'est-à-dire si  $t+\theta=D$ , ce que nous voulions prouver.

### Illustration

Reprenons les données utilisées ci-dessus (une obligation émise et remboursée au pair, de valeur égale à 1000, de durée 8 ans et de coupon constant égal à 100 ;  $r = 10\%$  et  $D = 5,868$ ) et supposons que, juste après son émission, le taux d'intérêt passe de  $10\%$  à  $9\%$ , pour se maintenir ensuite à ce niveau.

Les valeurs de réinvestissement et de réalisation à l'instant  $D$  sont respectivement données par

$$A_D = 100 \cdot (1,09^{D-1} + 1,09^{D-2} + 1,09^{D-3} + 1,09^{D-4} + 1,09^{D-5}) = 644,98$$

$$B_D = \frac{100}{1,09^{6-D}} + \frac{100}{1,09^{7-D}} + \frac{1100}{1,09^{8-D}} = 1104,99$$

et le taux de rendement actuariel à l'instant  $D$  est donné par

$$1000 \cdot (1 + r(D))^D = 1104,99 + 644,98$$

On trouve donc  $r(D) = 10,005 \%$ , ce qui montre clairement que les deux sources de risque s'anéantissent à un horizon égal à la durée<sup>91</sup>. ♣

### ***Approximation du prix d'une obligation***

L'application de la formule de Taylor à l'expression  $P(r + \Delta r)$  donne

$$\begin{aligned} P(r + \Delta r) &= P(r) + P'(r) \cdot \Delta r + O((\Delta r)^2) \\ &= P(r) - \frac{D \cdot P(r)}{1 + r} \Delta r + O((\Delta r)^2) \end{aligned}$$

ce qui permet d'écrire une formule approximative évaluant  $P(r + \Delta r)$  :

$$\begin{aligned} P(r + \Delta r) &\approx P(r) \left[ 1 - \frac{D}{1 + r} \Delta r \right] \\ &\approx P(r) [1 - D_m \cdot \Delta r] \end{aligned}$$

### ***Illustration***

Poursuivant l'exemple précédent (une obligation émise et remboursée au pair, de valeur égale à 1000, de durée 8 ans et de coupon constant égal à 100 ;  $r = 10 \%$  et  $D = 5,868$ ), on a, pour la durée modifiée,

$$D_m = \frac{5,868}{1,1} = 5,335.$$

Vérifions l'interprétation de ce paramètre pour une augmentation de  $r$  de 0,8 %. Le nouveau prix (évaluation exacte) est alors donné par

$$P(0,108) = 100 \cdot \sum_{t=1}^8 1,108^{-t} + 1000 \cdot 1,108^{-8} = 958,536,$$

alors que l'approximation ci-dessus donne

$$P(0,108) \approx 1000 \cdot (1 - 5,335 \cdot 0,008) = 957,321. \clubsuit$$

### **La relation**

<sup>91</sup> La très légère différence entre  $r(D)$  et le taux de rendement actuariel à l'émission ( $r = 10 \%$ ) provient du fait que le résultat que nous venons de donner (démontré à l'aide de la formule de Taylor) n'est valable que pour une variation infinitésimale du taux d'intérêt, alors que dans notre exemple, cette variation est de 1 %.

$$P(r + \Delta r) - P(r) \approx -D_m P(r) \cdot \Delta r ,$$

utilisée avec  $\Delta r = 0,0001$  (le centième de pour cent est appelé *point de base*<sup>92</sup>), donne la variation de prix d'une obligation provoquée par une augmentation du taux actuariel. Cette variation porte le nom de *valeur d'un point de base*<sup>93</sup> et se note *VBP*.

### 4.2.3. Convexité

#### Définition

La duration (ou la duration modifiée) permet d'appréhender les variations du prix en fonction du taux d'intérêt, mais seulement de manière linéaire. Cette variation linéaire est bien mise en évidence par la relation du premier degré

$$P(r + \Delta r) \approx P(r) \cdot \left[ 1 - D \frac{\Delta r}{1 + r} \right] .$$

Il est cependant facile de montrer que la courbe du prix n'est pas linéaire (voir figure 4.3). C'est pour prendre cette courbure en considération que l'on introduit un nouveau paramètre.

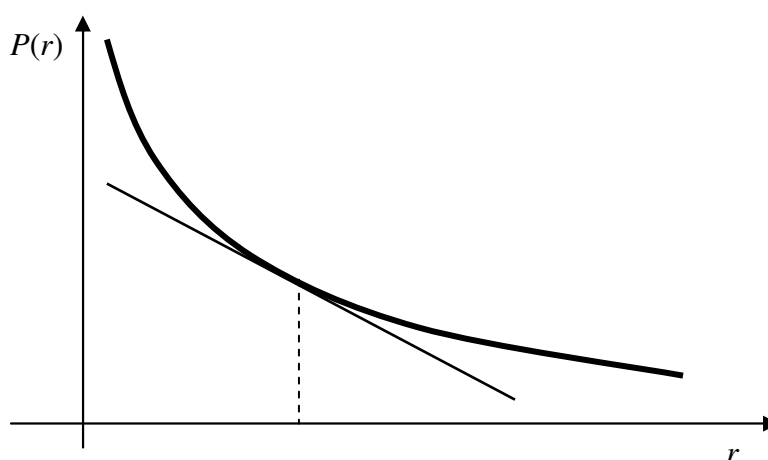


Figure 4.3 – Prix d'une obligation

On définit ainsi la *convexité* d'une obligation, notée *C*, par la relation

$$C = \frac{\sum_{t=1}^T t(t+1)C_t(1+r)^{-t} + T(T+1)R(1+r)^{-T}}{(1+r)^2 \cdot \left[ \sum_{t=1}^T C_t(1+r)^{-t} + R(1+r)^{-T} \right]} .$$

<sup>92</sup> On utilise couramment le terme anglais « *basis point* ».

<sup>93</sup> En anglais « *value of one basis point* ».

### Illustration

Reprenons les données utilisées ci-dessus (une obligation émise et remboursée au pair, de valeur égale à 1000, de durée 8 ans et de coupon constant égal à 100 ;  $r = 10\%$ ).

$$C = \frac{100 \cdot [1 \cdot 2 \cdot 1,1^{-1} + 2 \cdot 3 \cdot 1,1^{-2} + \dots + 8 \cdot 9 \cdot 1,1^{-8}] + 1000 \cdot 8 \cdot 9 \cdot 1,1^{-8}}{1,1^2 \cdot \{100 \cdot [1,1^{-1} + 1,1^{-2} + \dots + 1,1^{-8}] + 1000 \cdot 1,1^{-8}\}} = 38,843 \text{ .♣}$$

### Interprétation

La notion mathématique permettant de prendre en compte cette courbure est la dérivée seconde de la fonction. On a donc

$$\begin{aligned} P''(r) &= \sum_{t=1}^T t(t+1) \cdot C_t (1+r)^{-t-2} + T(T+1) \cdot R(1+r)^{-T-2} \\ &= (1+r)^{-2} \left\{ \sum_{t=1}^T t(t+1) \cdot C_t (1+r)^{-t} + T(T+1) \cdot R(1+r)^{-T} \right\} \\ &= C \cdot P(r) \end{aligned}$$

ce qui permet d'écrire

$$C = \frac{P''(r)}{P(r)}.$$

### Approximation du prix d'une obligation

On peut maintenant appliquer la formule de Taylor à l'expression  $P(r + \Delta r)$  jusqu'à l'ordre 2 :

$$\begin{aligned} P(r + \Delta r) &= P(r) + P'(r) \cdot \Delta r + \frac{P''(r)}{2} (\Delta r)^2 + O((\Delta r)^3) \\ &= P(r) - \frac{D \cdot P(r)}{1+r} \Delta r + \frac{C \cdot P(r)}{2} (\Delta r)^2 + O((\Delta r)^3) \end{aligned}$$

ce qui donne la formule d'approximation

$$P(r + \Delta r) \approx P(r) \left[ 1 - D_m \cdot \Delta r + \frac{1}{2} C \cdot (\Delta r)^2 \right].$$

### Illustration

Poursuivons l'exemple précédent (une obligation émise et remboursée au pair, de valeur égale à 1000, de durée 8 ans et de coupon constant égal à 100 ;  $r = 10\%$ ,  $D = 5,868$  et  $C = 38,843$ ).

En supposant à nouveau une augmentation de  $r$  de  $0,8\%$ , on a l'évaluation du prix :

$$P(0,108) \approx 1000 \cdot \left[ 1 - 5,335 \cdot 0,008 + \frac{1}{2} 38,843 \cdot 0,008^2 \right] = 958,564,$$

approximation beaucoup plus précise que celle n'utilisant que la duration, cette dernière donnant une valeur de 957,321 alors que la valeur exacte est égale à 958,536.♣

Le terme du second degré  $C(\Delta r)^2/2$  de la formule d'approximation étant toujours positif, il apparaît donc que lorsque l'on est amené à choisir entre deux obligations de même rendement (taux actuariel) et de même duration, on aura intérêt à choisir celle présentant la plus haute convexité, indépendamment du sens de la variation potentielle du taux de rendement.

## 4.3. STRUCTURE DETERMINISTE DES TAUX D'INTERET<sup>94</sup>

### 4.3.1. Courbes de taux

Le taux de rendement actuariel à l'émission d'une obligation tel qu'il a été défini au § 4.1.2 est bien évidemment une caractéristique propre du titre en question. Celui-ci variera d'une obligation à l'autre, en fonction essentiellement de la qualité de l'émetteur (évaluée par les ratings émis par les sociétés de notations) et de la maturité du titre.

Le premier facteur est bien sûr très difficile à modéliser ; nous n'en tiendrons donc pas compte, en supposant dans toute cette section 4.3 que l'on a affaire à un émetteur public, ne présentant en principe aucun risque de défaut. Quant au second facteur, on pourrait croire que pour des maturités égales, le taux soit – en vertu de la loi de l'offre et de la demande – le même pour tous les titres. En réalité, la politique de coupons des différents émetteurs introduit ici aussi des différences ; nous n'envisagerons donc dans ce qui suit que des obligations zéro-coupon, le taux ne dépendant donc plus maintenant que de la maturité. Cette simplification se justifie par le fait qu'une obligation classique est une simple « superposition » de titres zéro-coupon, qu'on évaluera par actualisation des différents flux financiers (coupons et remboursement) au taux correspondant<sup>95</sup>.

<sup>94</sup> On trouvera une présentation détaillée de ces notions par exemple dans BISIERE C., *La structure par terme des taux d'intérêt*, Presses Universitaires de France, 1997.

<sup>95</sup> Ceci justifie que le titre de la présente section 4.3 parle de *taux d'intérêt* et plus d'obligations.

Ne seront envisagées ici que des structures déterministes des taux d'intérêt, le cas aléatoire étant abordé au § 4.5.

Si on note  $P(s)$  le prix à l'émission d'une obligation zéro-coupon de maturité  $s$  et  $R(s)$  le taux constaté sur le marché à l'instant 0 pour ce type de titre – appelé *taux au comptant* – ces deux grandeurs sont évidemment liées par la relation

$$P(s) = (1 + R(s))^{-s}.$$

La grandeur  $R(s)$ , pour toutes les valeurs de  $s > 0$ , constitue la *structure par terme des taux d'intérêts* à l'instant 0 et le graphe de cette fonction est la *courbe des taux d'intérêt*<sup>96</sup>.

L'allure la plus « naturelle » d'une courbe des taux d'intérêt est bien évidemment croissante : l'investisseur doit être mieux rémunéré s'il investit pour une durée plus longue. Il n'en est toutefois pas toujours ainsi : on observe fréquemment en pratique, outre les courbes croissantes, des courbes plates ( $R(s)$  constant), des courbes inversées ( $R(s)$  décroissant) ainsi que des courbes bosselées (voir figure 4.4).

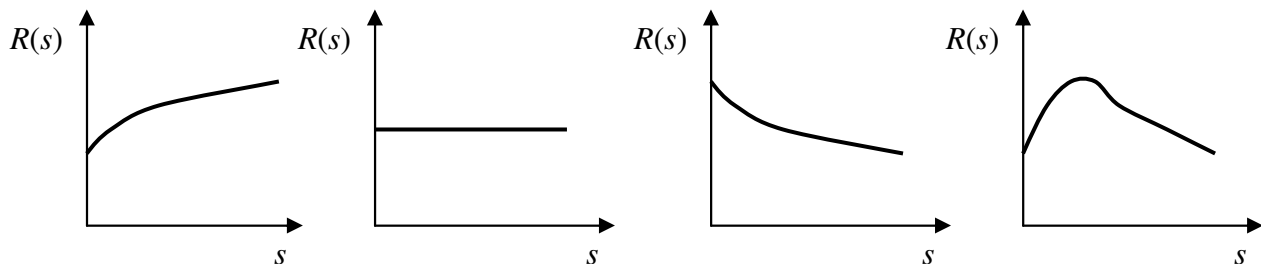


Figure 4.4 – Courbes de taux d'intérêt

### 4.3.2. Structure statique des taux d'intérêt

Les modèles statiques examinent la structure des taux d'intérêt à un instant fixé, que nous noterons 0 et ils traitent d'une obligation zéro-coupon donnant lieu à un remboursement de 1, ce qui n'est pas une restriction.

Dans ce paragraphe (comme dans le suivant), nous détaillerons la modélisation dans le cas discret, puis nous la généraliserons au cas continu ensuite. Ce sont ces aspects continus qui seront utilisés au § 4.5 pour les modèles d'évolution stochastiques.

#### **Modélisation discrète**

Le prix en 0 d'une obligation de maturité  $s$  se notant<sup>97</sup>  $P_0(s)$  et le taux au comptant associé étant représenté par  $R_0(s)$ , on a bien entendu

$$P_0(s) = (1 + R_0(s))^{-s}.$$

<sup>96</sup> En anglais *yield curve*.

<sup>97</sup> Bien entendu,  $P_0(0) = 1$ .



Le taux d'intérêt au comptant  $R_0(s)$  agrège en fait toute l'information sur les taux d'intérêt relatifs aux périodes  $]0 ; 1]$ ,  $]1 ; 2]$ , ...,  $]s - 1 ; s]$ . Nous noterons  $r(t)$  et nous appellerons *taux d'intérêt à terme* ou *taux d'intérêt court terme*, celui qui est relatif à la période  $]t - 1 ; t]$ . On a par conséquent

$$(1 + R_0(s))^s = (1 + r(1)) \cdot (1 + r(2)) \cdot \dots \cdot (1 + r(s)).$$

Réciproquement, il est aisé d'exprimer les taux à terme en fonction des taux au comptant :

$$\begin{cases} r(1) = R_0(1) \\ 1 + r(s) = \frac{(1 + R_0(s))^s}{(1 + R_0(s-1))^{s-1}} \quad s = 2, 3, \dots \end{cases}$$

De la même manière, on a

$$r(s) = \frac{P_0(s-1)}{P_0(s)} - 1 \quad (s > 0).$$

En résumé, on peut facilement passer de chacune des trois structures suivantes à une autre : celle des prix  $\{P_0(s) : s = 1, 2, \dots\}$ , celle des taux d'intérêt au comptant  $\{R_0(s) : s = 1, 2, \dots\}$  et celle des taux d'intérêt à terme  $\{r(s) : s = 1, 2, \dots\}$ .

### Illustration

Considérons une structure de taux au comptant définie pour les échéances 1 à 6 ans donnée dans le tableau 4.1.

$s$	1	2	3	4	5	6
$R_0(s)$	6,0%	6,6%	7,0%	7,3%	7,5%	7,6%

Tableau 4.1 – Structure de taux au comptant

Cette structure (croissante) est représentée dans la figure 4.5.

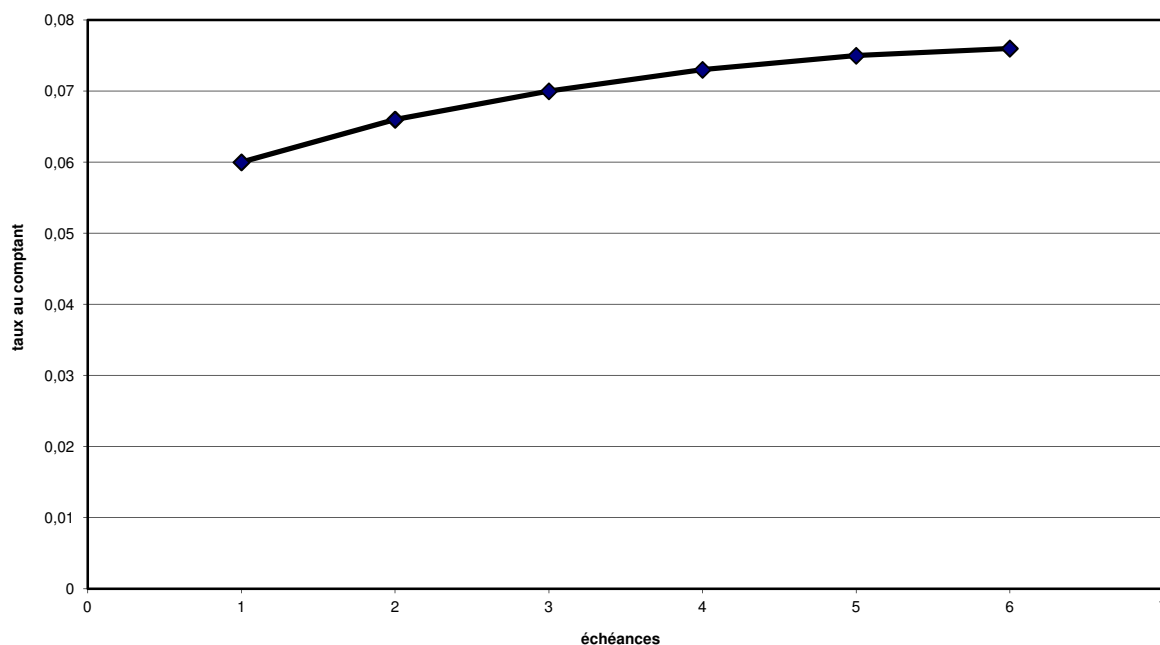


Figure 4.5 – Structure de taux au comptant

On déduit aisément de là les prix et les taux à terme ; par exemple

$$P_0(5) = 1,075^{-5} = 0,6966$$

$$r(5) = \frac{1,075^5}{1,073^4} - 1 = 0,0830$$

ce qui donne de manière générale,

$s$	$R_0(s)$	$P_0(s)$	$r(s)$
0		1,0000	
1	0,060	0,9434	0,0600
2	0,066	0,8800	0,0720
3	0,070	0,8163	0,0780
4	0,073	0,7544	0,0821
5	0,075	0,6966	0,0830
6	0,076	0,6444	0,0810

Tableau 4.2 – Structures de prix et de taux en 0♣

### Modélisation continue

Dans le cas où l'ensemble des époques est  $]0 ; +\infty[$ , on conserve les mêmes définitions et les mêmes notations pour les structures de prix et de taux au comptant :  $\{P_0(s) : s > 0\}$  et  $\{R_0(s) : s > 0\}$ . Ce dernier sera un taux instantané : après une durée  $s$ , un montant 1 deviendra, à ce taux,

$$e^{s \cdot R_0(s)}.$$

Nous noterons encore, le temps de réaliser les passages à la limite,  $R_0^d(s)$  le taux au comptant du modèle discret (même appliqué à une durée non entière). Il est donc lié au taux au comptant du modèle continu par la relation  $R_0(s) = \ln(1 + R_0^d(s))$ .

En ce qui concerne le taux à terme, nous introduisons provisoirement la notation  $r(t_1, t_2)$  pour représenter le taux d'intérêt relatif à la période  $]t_1 ; t_2]$  et nous définissons le *taux d'intérêt instantané à terme* ou encore *taux spot* par

$$r(t) = \lim_{s \rightarrow t+} \frac{1}{s-t} \int_t^s r(t, u) du.$$

On obtient aisément comme ci-dessus

$$(1 + r(s, s + \Delta s))^{\Delta s} = \frac{(1 + R_0^d(s + \Delta s))^{s + \Delta s}}{(1 + R_0^d(s))^s},$$

qui s'écrit, grâce à la formule de Taylor,

$$\left[1 + \Delta s \cdot r(s, s + \Delta s) + O((\Delta s)^2)\right] \cdot (1 + R_0^d(s))^s = (1 + R_0^d(s + \Delta s))^{s + \Delta s}.$$

Cette relation peut se réécrire

$$r(s, s + \Delta s) \cdot (1 + R_0^d(s))^s + O(\Delta s) = \frac{(1 + R_0^d(s + \Delta s))^{s + \Delta s} - (1 + R_0^d(s))^s}{\Delta s},$$

qui devient, après passage à la limite,

$$r(s) = \frac{\left[(1 + R_0^d(s))^s\right]'}{(1 + R_0^d(s))^s} = \left[\ln(1 + R_0^d(s))^s\right]' = \left[s \cdot \ln(1 + R_0^d(s))\right]' = [s \cdot R_0(s)]'.$$

Cette relation exprimant le taux instantané à terme en fonction du taux au comptant s'inverse aisément en intégrant :

$$R_0(s) = \frac{1}{s} \int_0^s r(u) du$$

et peut se traduire aisément en disant que le taux au comptant pour la période  $]0 ; s]$  est la moyenne du taux instantané à terme pour la même période.

Le prix est bien sûr lié aux deux taux par les relations

$$P_0(s) = e^{-s \cdot R_0(s)} = e^{-\int_0^s r(u) du}.$$

### Remarque

En cas de structure de taux plate (c'est-à-dire  $R_0(s)$  indépendant de  $s$ ), il est facile de voir, en développant la relation  $[s \cdot R_0(s)]' = r_0(s)$ , que

$$R_0(s) = r(s) = r \quad \text{pour tout } s$$

et que la structure de prix est donnée par

$$P_0(s) = e^{-rs}.$$

### 4.3.3. Structure dynamique des taux d'intérêt

Les modèles dynamiques examinent la structure des taux d'intérêt à un instant quelconque  $t$ . Ils traitent toujours d'une obligation zéro-coupon, émise en 0 et donnant lieu à un remboursement de 1.

Ils permettent éventuellement de prendre en compte les déformations de la courbe des taux ; en fait, nous allons étudier le lien existant entre les structures de prix et de taux pour différentes époques d'observation.

### Modélisation discrète

On note<sup>98</sup>  $P_t(s)$  le prix à l'instant  $t$  de l'obligation émise en 0 et échéant en  $s$ . On note également  $R_t(s)$  le taux au comptant relatif à l'intervalle  $]t ; s]$ . Enfin, on représente par  $r(t)$  le taux à terme relatif à la période  $]t - 1 ; t]$ .

Par des raisonnements en tous points similaires à ceux des modèles statiques, on obtient aisément les relations

$$P_t(s) = (1 + R_t(s))^{-(s-t)}$$

$$(1 + R_t(s))^{s-t} = (1 + r(t+1)) \cdot (1 + r(t+2)) \cdot \dots \cdot (1 + r(s))$$

Cette dernière s'inverse aisément en

<sup>98</sup> On suppose bien sûr  $0 \leq t < s$ .

$$\begin{cases} r(t+1) = R_t(t+1) \\ 1+r(s) = \frac{(1+R_t(s))^s}{(1+R_t(s-1))^{s-1}} \end{cases} \quad s = t+2, t+3, \dots$$

On a également, entre la structure des prix et celle des taux à terme,

$$r(s) = \frac{P_t(s-1)}{P_t(s)} - 1 \quad (s > t).$$

Le lien entre les structures des prix à différentes époques d'observation s'exprime par la relation suivante :

$$\begin{aligned} P_t(s) &= [(1+r(t+1)) \cdot (1+r(t+2)) \cdot \dots \cdot (1+r(s))]^{-1} \\ &= \left[ \frac{(1+r(t)) \cdot (1+r(t+1)) \cdot (1+r(t+2)) \cdot \dots \cdot (1+r(s))}{(1+r(t))} \right]^{-1} \\ &= \frac{(1+R_{t-1}(s))^{-(s-t+1)}}{(1+R_{t-1}(t))^{-1}} \\ &= \frac{P_{t-1}(s)}{P_{t-1}(t)} \end{aligned}$$

Ce résultat se généralise aisément : quel que soit  $u$  compris entre  $t$  et  $s$  ( $t \leq u \leq s$ ), on a

$$P_t(s) = \frac{P_u(s)}{P_u(t)}.$$

De cette relation peut se déduire un lien – dont l'expression n'est pas très élégante – entre les structures de taux au comptant à différentes époques.

### *Illustration*

Reprenons la structure de taux d'intérêt au comptant utilisée au paragraphe précédent : 6%, 6,6%, 7%, 7,3%, 7,5% et 7,6% pour les échéances respectives 1, 2, 3, 4, 5 et 6 ans. Voyons ce que deviennent les structures après deux ans. On trouve aisément

$$P_2(5) = \frac{P_0(5)}{P_0(2)} = \frac{0,69656}{0,88001} = 0,7915$$

$$R_2(5) = P_2(5)^{\frac{-1}{5-2}} - 1 = 0,7915^{-1/3} - 1 = 0,0810$$

et, de manière plus générale,

$s$	$P_2(s)$	$R_2(s)$
2	1,0000	
3	0,9276	0,0780
4	0,8573	0,0800
5	0,7915	0,0810
6	0,7322	0,0810

Tableau 4.3 – Structures de prix et de taux en 2

Remarquons qu'on a bien

$$r(5) = \frac{P_0(4)}{P_0(5)} - 1 = \frac{P_2(4)}{P_2(5)} - 1 = 0,0830 \text{ . } \clubsuit$$

### Modélisation continue

Les prix  $P_t(s)$  et les taux au comptant  $R_t(s)$  sont définis comme pour les modèles dynamiques, mais avec une observation à l'instant  $t$  au lieu de 0. Les taux instantané à terme  $r(t)$  sont définis de la même manière.

On voit aisément que les relations liant les deux taux sont

$$r(s) = \left[ (s-t) \cdot R_t(s) \right]'_s \quad \forall t$$

$$R_t(s) = \frac{1}{s-t} \int_t^s r(u) du$$

tandis que celles liant les taux aux prix sont données par

$$P_t(s) = e^{-(s-t) \cdot R_t(s)} = e^{-\int_t^s r(u) du}.$$

#### 4.3.4. Modélisation déterministe et modélisation aléatoire

Les relations ci-dessus ont été établies dans un contexte déterministe. Entre autres, on assimilé le taux instantané court et le taux à terme. Dans une approche plus générale (modélisation stochastique), il conviendra de distinguer

a) le *taux spot*, défini par

$$r(t) = \lim_{s \rightarrow t+} R_t(s)$$

- b) le *taux instantané à terme* ou *taux forward*, défini de la manière suivante : si  $f_t(s_1, s_2)$  représente le taux d'intérêt, vu depuis l'époque  $t$ , d'une obligation émise en  $s_1$  et d'échéance  $s_2$ , le taux forward (en  $s$  vu depuis  $t$ , avec  $t < s$ ) est

$$f_t(s) = \lim_{u \rightarrow s^+} f_t(s, u).$$

Dans une modélisation générale, c'est ce taux forward qu'il faut utiliser pour retrouver les structures de prix et de taux au comptant :

$$P_t(s) = e^{-\int_t^s f_t(u) du}$$

$$R_t(s) = \frac{1}{s-t} \int_t^s f_t(u) du$$

Il est facile de voir que ces deux taux – spot et forward – sont liés par la relation

$$r(t) = f_t(t).$$

On peut démontrer que, dans le cas déterministe,  $f_t(s)$  est indépendant de  $t$  et qu'on peut donc identifier les deux taux :  $f_t(s) = r(s)$ . C'est donc uniquement dans ce contexte que l'on a

$$P_t(s) = e^{-\int_t^s r(u) du}$$

$$R_t(s) = \frac{1}{s-t} \int_t^s r(u) du$$

$$P_t(s) = \frac{P_u(s)}{P_u(t)}$$

## 4.4. STRATEGIES DE GESTION D'UN PORTEFEUILLE OBLIGATAIRE

### 4.4.1. Stratégie passive : immunisation

L'objectif d'une gestion passive est de neutraliser le risque du portefeuille, dû aux fluctuations des taux d'intérêt.

### *Duration et convexité d'un portefeuille*

Considérons un portefeuille obligataire constitué à l'instant 0 de  $N$  titres ( $j = 1, \dots, N$ ), caractérisés chacun par

- une maturité (vie résiduelles)  $T_j$ ;

- des coupons à venir  $C_{j,t}$  ( $t = 1, \dots, T_j$ ) ;
- une valeur de remboursement  $R_j$  ;
- un taux actuariel à l'émission  $r_j$  ;
- un prix  $P_j$ .

On notera encore  $T$  la plus grande des maturités  $T_j$  et  $F_{j,t}$  le flux financier généré par le titre ( $j$ ) à l'instant  $t$  :

$$F_{j,t} = \begin{cases} C_j & \text{si } t < T_j \\ C_{T_j} + R_j & \text{si } t = T_j \\ 0 & \text{si } t > T_j \end{cases}$$

La duration du  $j$ -ème titre est donnée par

$$D_j = \frac{\sum_{t=1}^{T_j} t \cdot C_{j,t} (1+r_j)^{-t} + T_j \cdot R_j (1+r_j)^{-T_j}}{\sum_{t=1}^{T_j} C_{j,t} (1+r_j)^{-t} + R_j (1+r_j)^{-T_j}} = \frac{\sum_{t=1}^T t \cdot F_{j,t} (1+r_j)^{-t}}{P_j}$$

Supposons enfin que le  $j$ -ème titre est présent au sein du portefeuille en nombre  $n_j$ .

Le flux financier actualisé généré par le portefeuille à l'époque  $t$  est égal à

$$\sum_{j=1}^N n_j F_{j,t} (1+r_j)^{-t}$$

et son prix à

$$\sum_{j=1}^N n_j P_j .$$

La duration du portefeuille peut donc s'écrire



$$\begin{aligned}
D_p &= \frac{\sum_{t=1}^T t \cdot \sum_{j=1}^N n_j F_{j,t} (1+r_j)^{-t}}{\sum_{k=1}^N n_k P_k} \\
&= \sum_{j=1}^N \frac{n_j}{\sum_{k=1}^N n_k P_k} \sum_{t=1}^{T_j} t \cdot F_{j,t} (1+r_j)^{-t} \\
&= \sum_{j=1}^N \frac{n_j P_j}{\sum_{k=1}^N n_k P_k} \cdot \frac{\sum_{t=1}^{T_j} t \cdot F_{j,t} (1+r_j)^{-t}}{P_j} \\
&= \sum_{j=1}^N X_j D_j
\end{aligned}$$

où

$$X_j = \frac{n_j P_j}{\sum_{k=1}^N n_k P_k}$$

représente la proportion du  $j$ -ème titre au sein du portefeuille, exprimée en terme de capitalisation.

Le même raisonnement conduit à la convexité du portefeuille :

$$C_p = \sum_{j=1}^N X_j C_j .$$

### ***Immunisation d'un portefeuille***

Un portefeuille est dit immunisé à l'horizon  $H$  si sa valeur à cette date est au moins celle qu'il aurait eue si les taux d'intérêt étaient restés constants durant la période  $[0 ; H]$ . En appliquant le résultat établi au § 4.2.2 pour une obligation au portefeuille, on obtient le même résultat : un portefeuille d'obligations est immunisé à un horizon correspondant à sa duration.

Bien évidemment, lors de chaque modification des taux d'intérêt, la duration résiduelle varie brusquement. Le gestionnaire prudent d'un portefeuille obligataire qui désire immuniser celui-ci pour un horizon  $H$  qu'il se fixe devra donc

- constituer un portefeuille de duration  $H$  ;
- après chaque variation (significative) des taux d'intérêt, modifier par des achats et des ventes la composition du portefeuille (c'est-à-dire adapter les proportions  $X_j$ ) de manière à « poursuivre » la duration résiduelle.

Bien entendu, ces adaptations de la composition du portefeuille entraînent des frais de transaction qu'il convient de prendre en considération et de mettre en balance avec les avantages que cette stratégie immunisante procure.

#### *Remarque*

On a noté au § 4.2.3 qu'entre deux obligations de même rendement (taux actuariel) et de même duration, celle présentant la plus haute convexité est plus intéressante.

Ce résultat reste valable pour un portefeuille et peut donc être pris en considération par le gestionnaire lors des différentes révisions de son portefeuille.♣

### **4.4.2. Stratégie active**

L'objectif d'une gestion active est d'obtenir un rendement supérieur à celui obtenu grâce à l'immunisation, c'est-à-dire au taux de rendement actuariel à l'émission.

Dans le cas (le plus courant) d'une courbe des taux croissante, et si la courbe des taux reste inchangée au cours du temps, la technique consiste en l'achat de titres de maturité plus élevée que l'horizon d'investissement et en la revente de ceux-ci avant leur échéance<sup>99</sup>.

#### *Illustration*

Reprenons la structure de taux introduite à la section précédente (tableau 4.4)

$s$	$R_0(s)$	$P_0(s)$
0		1,0000
1	0,060	0,9434
2	0,066	0,8800
3	0,070	0,8163
4	0,073	0,7544
5	0,075	0,6966
6	0,076	0,6444

*Tableau 4.4 – Structure de prix et de taux en 0*

et supposons que l'investisseur se fixe un horizon de deux ans.

S'il achète simplement un titre de maturité 2 ans, il obtiendra simplement un rendement annuel de 6,6% ; d'ailleurs, le rendement sur 2 ans peut se calculer par

<sup>99</sup> Si la courbe des taux est plate (et le reste), la stratégie présentée donne le même rendement que l'achat du titre dont la maturité est l'horizon d'investissement.

$$\frac{1 - 0,8800}{0,8800} = 0,1364$$

et  $\sqrt{1,1364} = 1,066$ .

S'il achète un titre de maturité 5 ans (au prix de 0,6966) et le revend après 2 ans (au prix du titre 3 ans si les courbes de taux restent inchangées, c'est-à-dire 0,8163), il réalise un rendement total de

$$\frac{0,8163 - 0,6966}{0,6966} = 1,1719,$$

ce qui donne  $\sqrt{1,1719} = 1,0825$ , c'est-à-dire un rendement annuel de 8,25%, qui est significativement plus intéressant que celui (6,6%) obtenu avec le titre 2 ans.

Notons que l'on retrouve ici une interprétation du taux à terme puisque le rendement total sur la période [3 ; 5] qui a été effectivement utilisée est donné par

$$(1 + r(4)) \cdot (1 + r(5)) = 1,0821 \cdot 1,0830 = 1,1719. \clubsuit$$

Le taux intéressant obtenu grâce à cette technique suppose que la courbe des taux reste inchangée au cours du temps. Si cette courbe fluctue – et plus particulièrement le taux au comptant utilisé pour le calcul du prix de revente – l'investisseur est exposé au risque de fluctuation des taux d'intérêt. Cette fluctuation lui est favorable (resp. défavorable) si en cas de baisse (resp. hausse) du taux en question.

Dans ce cas, c'est à l'investisseur de choisir entre un rendement sûr et un rendement potentiellement plus élevé mais risqué.

### *Illustration*

Avec les mêmes données, si, après l'acquisition du titre de maturité 5 ans, le taux au comptant du titre 3 ans passe de 7% à 8%, le prix de celui-ci tombe de 0,8163 à 0,7938 et le rendement sur les deux années vaut

$$\frac{0,7938 - 0,6966}{0,6966} = 1,1396.$$

On obtient alors  $\sqrt{1,1396} = 1,0675$  qui correspond à un rendement annuel de 6,75%.♣

## 4.5. MODELES STOCHASTIQUES D'EVOLUTION DES OBLIGATIONS

Les modèles qui vont être présentés ici constituent en réalité des généralisations des structures déterministes des taux d'intérêt. L'objectif est d'obtenir des relations gouvernant l'évolution des prix  $P_t(s)$  et des taux au comptant  $R_t(s)$ .

Il existe deux grandes catégories de tels modèles : les modèles de déformation et les modèles d'arbitrage.

Les *modèles de déformation* examinent l'évolution du prix  $P_t(s)$  lorsque la structure des taux d'intérêt subit des déformations. Un modèle simple est celui de Ho et Lee<sup>100</sup>, pour lequel la déformation de la courbe de taux se matérialise par deux mouvements possibles à chaque époque ; il s'agit donc d'un modèle discret de type binomial. Un modèle plus évolué est celui de Heath, Jarrow et Morton<sup>101</sup>, dont il existe une version discrète et une version continue et pour lesquels les déformations de la courbe des taux sont plus complexes.

Les *modèles d'arbitrage* consistent en l'élaboration – et si possible la résolution – d'une équation aux dérivées partielles pour le prix  $P_t(s, v_1, v_2, \dots)$ , considéré comme fonction de  $t, v_1, v_2, \dots$  ( $s$  fixé) à l'aide

- d'un raisonnement d'arbitrage ;
- d'hypothèses quant aux processus stochastiques gouvernant l'évolution des *variables d'état*  $v_1, v_2, \dots$

Les modèles à une seule variable d'état sont, pour citer les plus courants, le modèle de Merton<sup>102</sup>, celui de Vasicek<sup>103</sup> et celui de Cox, Ingersoll et Ross<sup>104</sup> ; ils utilisent tous comme variable d'état le taux spot  $r(t)$ . Parmi les modèles à deux variables d'état, citons

- le modèle de Brennan et Schwartz<sup>105</sup>, qui utilise comme variables d'état le taux spot  $r$  et le taux long  $l$  ;
- le modèle de Nelson et Schaefer<sup>106</sup> et celui de Schaefer et Schwartz<sup>107</sup>, pour lesquels les variables d'état sont le taux long  $l$  et le spread  $s = l - r$  ;
- le modèle de Richard<sup>108</sup>, qui utilise le taux spot et le taux d'inflation ;

<sup>100</sup> HO T. et LEE S., Term structure movement and pricing interest rate contingent claims, *Journal of Finance*, Vol. 41, n° 5, 1986, pp. 1011-1029.

<sup>101</sup> HEATH D., JARROW R. et MORTON A., Bond pricing and the term structure of interest rates : a new methodology, *Cornell University*, 1987.

HEATH D., JARROW R. et MORTON A., Bond pricing and the term structure of interest rates : discrete time approximation, *Journal of Financial and Quantitative Analysis*, Vol. 25, 1990, pp. 419-440.

<sup>102</sup> MERTON R., Theory of rational option pricing, *Bell Journal of Economics and Management Science*, Vol. 4, n° 1, 1973, pp. 141-183.

<sup>103</sup> VASICEK O., An equilibrium characterization of the term structure, *Journal of Financial Economics*, Vol. 5, n° 2, 1977, pp. 177-188.

<sup>104</sup> COX J., INGERSOLL J. et ROSS J., A theory of the term structure of interest rates, *Econometrica*, Vol. 53, n° 2, 1985, pp. 385-406.

<sup>105</sup> BRENNAN M. et SCHWARTZ E., A continuous time approach to the pricing of bonds, *Journal of Banking and Finance*, Vol. 3, n° 2, 1979, pp. 133-155.

<sup>106</sup> NELSON J. et SCHAEFER S., The dynamics of the term structure and alternative portfolio immunization strategies, in BIERWAG G., KAUFMAN G. et TOEVS A. (éds.), *Innovations in bond portfolio management : duration analysis and immunization*, JAI Press, 1983.

<sup>107</sup> SCHAEFER S. et SCHWARTZ E., A two-factor model of the term structure : an approximate analytical solution, *Journal of Financial and Quantitative Analysis*, Vol. 19, n° 4, 1984, pp. 413-424.

<sup>108</sup> RICHARD S., An arbitrage model of the term structure of interest rates, *Journal of Financial Economics*, Vol. 6, n° 1, 1978, pp. 33-57.

- le modèle de Ramaswamy et Sundaresan<sup>109</sup>, qui, à côté du taux spot, prend en considération la prime de risque instantanée liée au risque de défaut.

Nous n'aborderons ici que les plus simples des modèles d'arbitrage : après une présentation générale du principe de ces modèles (§ 4.5.1), nous examinerons successivement les modèles de Vasicek (§ 4.5.2) et de Cox, Ingersoll et Ross (§ 4.5.3)<sup>110</sup> ; nous présenterons ensuite (§ 4.5.4) la notion de duration stochastique.

#### 4.5.1. Modèles d'arbitrage à une variable d'état

##### *Principe général*

Rappelons (voir § 4.3) que les processus stochastiques qui nous intéressent ici sont

- le prix  $P_t(s)$  en  $t$  d'une obligation zéro-coupon (de valeur de remboursement unité) échéant à l'instant  $s$  (avec  $t < s$ ) ;
- le taux au comptant  $R_t(s)$ , lié au prix par la relation

$$P_t(s) = e^{-(s-t)R_t(s)} ;$$

- le taux spot, que nous noterons désormais<sup>111</sup>  $r_t$  ou  $r$  si aucune confusion n'est possible, et qui est le taux instantané à l'instant  $t$  et qui peut s'écrire

$$r_t = \lim_{s \rightarrow t+} R_t(s) = \lim_{s \rightarrow t+} \frac{1}{s-t} \int_t^s f_t(u) du .$$

C'est ce taux spot qui sera la variable d'état. Le prix et le taux au comptant s'écriront désormais  $P_t(s, r)$  et  $R_t(s, r)$  et seront considérés comme fonctions des seules variables  $t$  et  $r$ , l'échéance  $s$  étant fixe ; on suppose de plus que ces expressions sont par l'intermédiaire de  $r_t$  seulement.

On supposera ici que l'évolution de la variable d'état  $r_t$  est gouvernée par l'équation différentielle stochastique<sup>112</sup> générale

$$dr_t = a(t, r_t)dt + b(t, r_t)dw_t ,$$

où les coefficients  $a$  et  $b$  représentent respectivement le rendement moyen instantané du taux spot et la volatilité de ce dernier, et où  $w_t$  est un mouvement brownien standard.

L'application de la formule d'Itô à la fonction  $P_t(s, r_t)$  conduit à (avec des notations simplifiées)

<sup>109</sup> RAMASWAMY K. et SUNDARESAN M., The valuation of floating-rates instruments : theory and evidence, Journal of Financial Economics, Vol. 17, n° 2, 1986, pp. 251-272.

<sup>110</sup> Dans le CD-Rom joint, on trouvera des fichiers Excel<sup>TM</sup> présentant des simulations de ces processus stochastiques et ses courbes de taux relatives aux divers modèles, rassemblés dans le dossier nommé « Oblig ».

<sup>111</sup> Au lieu de  $r(t)$  comme au § 4.3, pour la légèreté des notations.

<sup>112</sup> Voir l'annexe 2.

$$\begin{aligned}
dP_t(s, r_t) &= \left( P'_t + P'_r a + \frac{1}{2} P''_{rr} b^2 \right) \cdot dt + P'_r b \cdot dw_t \\
&= P_t(s, r_t) \cdot \mu_t(s, r_t) \cdot dt - P_t(s, r_t) \cdot \sigma_t(s, r_t) \cdot dw_t
\end{aligned}$$

où on a posé

$$\begin{cases} \mu_t = \frac{P'_t + P'_r a + \frac{1}{2} P''_{rr} b^2}{P} \\ \sigma_t = -\frac{P'_r b}{P} \end{cases}$$

(notons que  $\sigma_t > 0$  puisque  $P'_r < 0$ ). L'expression  $\mu_t(s, r_t)$  est généralement appelée *taux de rendement moyen instantané* du titre.

Considérons maintenant deux échéances fixées quelconques  $s_1$  et  $s_2$  ( $> t$ ) et appliquons un raisonnement d'arbitrage en construisant à l'instant  $t$  un portefeuille constitué de

- l'émission d'une obligation d'échéance  $s_1$  ;
- l'achat de  $X$  obligations d'échéance  $s_2$  ;

avec un  $X$  choisi de manière que ce portefeuille ne contienne plus de composante aléatoire : le terme faisant intervenir  $dw_t$  doit disparaître.

La valeur de ce portefeuille à l'instant  $t$  est donnée par

$$V_t = -P_t(s_1) + XP_t(s_2)$$

et l'hypothèse d'absence d'opportunité d'arbitrage permet d'exprimer que le rendement moyen de ce portefeuille sur l'intervalle  $[t ; t + dt]$  est donné par le taux spot  $r_t$  :

$$\frac{dV_t}{V_t} = r_t \cdot dt + 0 \cdot dw_t.$$

En différenciant la valeur du portefeuille, on obtient

$$\begin{aligned}
dV_t &= -P_t(s_1)(\mu_t(s_1)dt - \sigma_t(s_1)dw_t) + X \cdot P_t(s_2)(\mu_t(s_2)dt - \sigma_t(s_2)dw_t) \\
&= [-P_t(s_1)\mu_t(s_1) + XP_t(s_2)\mu_t(s_2)] \cdot dt + [P_t(s_1)\sigma_t(s_1) - XP_t(s_2)\sigma_t(s_2)] \cdot dw_t
\end{aligned}$$

Le raisonnement d'arbitrage nous conduit donc à :

$$\begin{cases} \frac{-P_t(s_1)\mu_t(s_1) + XP_t(s_2)\mu_t(s_2)}{-P_t(s_1) + XP_t(s_2)} = r_t \\ \frac{P_t(s_1)\sigma_t(s_1) - XP_t(s_2)\sigma_t(s_2)}{-P_t(s_1) + XP_t(s_2)} = 0 \end{cases}$$

c'est-à-dire,

$$\begin{cases} XP_t(s_2) \cdot (\mu_t(s_2) - r_t) = P_t(s_1) \cdot (\mu_t(s_1) - r_t) \\ XP_t(s_2) \cdot \sigma_t(s_2) = P_t(s_1) \cdot \sigma_t(s_1) \end{cases}$$

On élimine  $X$ , par exemple en divisant les deux équations membre à membre, ce qui donne

$$\frac{\mu_t(s_1) - r_t}{\sigma_t(s_1)} = \frac{\mu_t(s_2) - r_t}{\sigma_t(s_2)},$$

ce qui montre que l'expression

$$\lambda_t(r_t) = \frac{\mu_t(s) - r_t}{\sigma_t(s)},$$

est indépendante de  $s$  ; cette expression porte le nom de *prime unitaire de risque de taux*.

En remplaçant  $\mu_t$  et  $\sigma_t$  par leur valeur dans la relation précédente, on obtient

$$P'_t + (a + \lambda b)P'_r + \frac{b^2}{2}P''_{rr} - rP = 0.$$

Il s'agit là de l'équation aux dérivées partielles du second ordre qui, avec la condition initiale  $P_s(s, r_t) = 1$ , définit le processus du prix. Cette équation devra être résolue pour chaque spécification de  $a(t, r_t)$ ,  $b(t, r_t)$  et  $\lambda_t(r_t)$ .

### **Modèle de Merton<sup>113</sup>**

Présentons, pour son intérêt historique<sup>114</sup>, le modèle le plus simple : celui de Merton. Celui-ci suppose que le taux spot suit un modèle de marche aléatoire :

$$dr_t = \alpha \cdot dt + \sigma \cdot dw_t,$$

avec  $\alpha$  et  $\sigma$  constants, et que la prime de risque est nulle :  $\lambda = 0$ .

L'équation aux dérivées partielles pour le prix prend la forme

$$P'_t + \alpha P'_r + \frac{\sigma^2}{2}P''_{rr} - rP = 0.$$

<sup>113</sup> MERTON R., Theory of rational option pricing, *Bell Journal of Economics and Management Science*, Vol. 4, n° 1, 1973, pp. 141-183.

<sup>114</sup> Il s'agit en effet là du premier modèle basé sur la représentation de l'évolution du taux spot par une équation différentielle stochastique.

Il est facile de vérifier que la solution de cette équation (avec la condition initiale) est donnée par

$$P_t(s, r_t) = \exp \left[ -(s-t)r_t - \frac{\alpha}{2}(s-t)^2 + \frac{\sigma^2}{6}(s-t)^3 \right].$$

Le taux de rendement moyen instantané est donné par

$$\mu_t(s, r_t) = \frac{P'_t + \alpha P'_r + \frac{\sigma^2}{2} P''_{rr}}{P} = \frac{r_t \cdot P}{P} = r_t,$$

ce qui montre que le rendement moyen est, dans ce cas, indépendant de l'échéance.

Le taux au comptant est quant à lui égal à

$$\begin{aligned} R_t(s, r_t) &= -\frac{1}{s-t} \ln P_t(s, r_t) \\ &= r_t + \frac{\alpha}{2}(s-t) - \frac{\sigma^2}{6}(s-t)^2 \end{aligned}$$

Cette expression montre que le taux au comptant est proche du taux spot à court terme, ce qui est logique, mais également (à cause du troisième terme) qu'il finit à coup sûr par devenir négatif pour des échéances lointaines, ce qui est beaucoup moins logique.

### Remarque

Si on généralise le modèle de Merton au cas où la prime de risque  $\lambda$  est une constante strictement positive, on obtient un rendement moyen  $\mu_t$  croissant avec l'échéance, mais la « propriété » désagréable du taux au comptant  $R_t$  demeure.♣

Ce modèle de Merton, par trop irréaliste, a été remplacé par les modèles plus proches de la réalité qui font l'objet des deux paragraphes suivants.

### 4.5.2. Modèle de Vasicek<sup>115</sup>

Dans ce modèle, la variable d'état  $r_t$  évolue suivant un processus d'Ornstein-Uhlenbeck :

$$dr_t = \delta(\theta - r_t) \cdot dt + \sigma \cdot dw_t,$$

<sup>115</sup> VASICEK O., An equilibrium characterization of the term structure, *Journal of Financial Economics*, Vol. 5, n° 2, 1977, pp. 177-188.



où les paramètres  $\delta$ ,  $\theta$  et  $\sigma$  sont des constantes strictement positives, et la prime unitaire de risque de taux est également une constante strictement positive :  $\lambda_t(r_t) = \lambda > 0$ .

La propriété essentielle du *processus d'Ornstein-Uhlenbeck* est de « rappeler » la variable  $r_t$  vers  $\theta$  si elle s'en éloigne trop et que  $\delta$  représente la « force de rappel ».

### Illustration

La figure 4.6 représente une trajectoire (évolution au cours du temps) simulée d'un tel processus sur 1000 très courtes périodes avec les valeurs  $\delta = 100$ ,  $\theta = 0,1$  et  $\sigma = 0,8$  avec une valeur de départ de  $r_0 = 10\%$ .

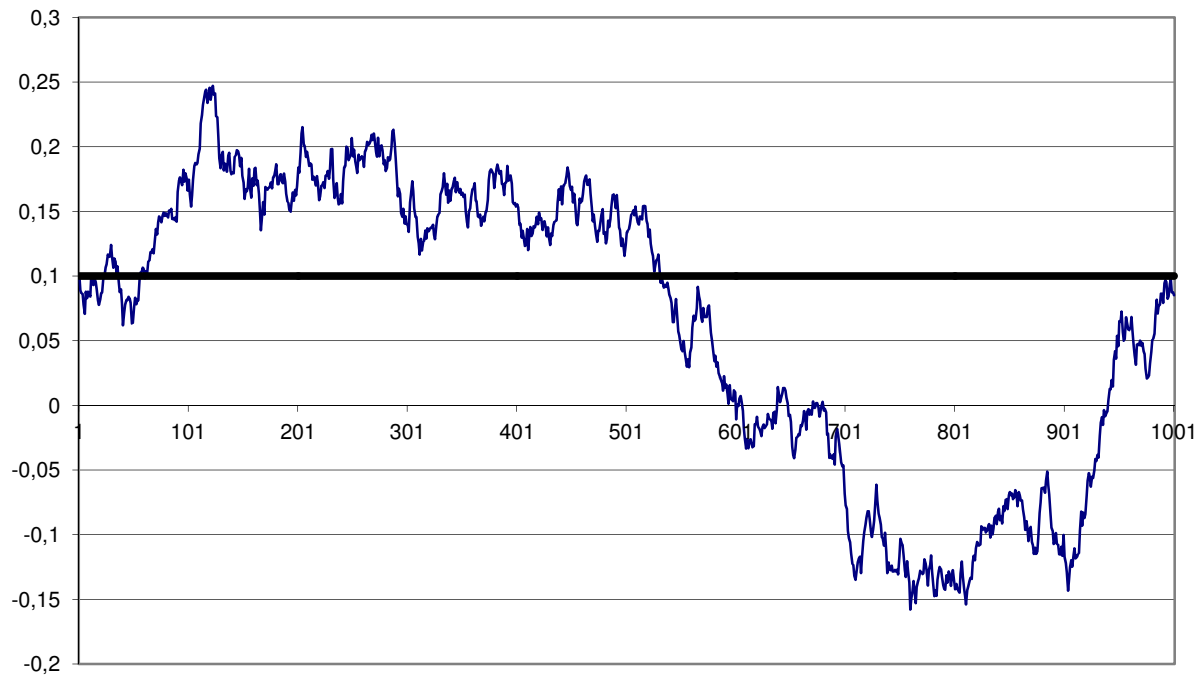


Figure 4.6 – Processus d'Ornstein-Uhlenbeck♣

L'équation aux dérivées partielles pour le prix s'écrit ici

$$P'_t + (\delta(\theta - r) + \lambda\sigma)P'_r + \frac{\sigma^2}{2}P''_{rr} - rP = 0.$$

La solution de cette équation et de sa condition initiale est donnée par

$$P_t(s, r_t) = \exp \left[ -k(s-t) + \frac{k-r_t}{\delta} (1 - e^{-\delta(s-t)}) - \frac{\sigma^2}{4\delta^3} (1 - e^{-\delta(s-t)})^2 \right],$$

où on a posé

$$k = \theta + \frac{\lambda\sigma}{\delta} - \frac{\sigma^2}{2\delta^2}.$$

Le taux de rendement moyen instantané est donné par

$$\begin{aligned}\mu_t(s, r_t) &= \frac{P'_t + \delta(\theta - r_t)P'_r + \frac{\sigma^2}{2}P''_{rr}}{P} \\ &= \frac{r_t \cdot P - \lambda\sigma \cdot P'_r}{P} \\ &= r_t + \frac{\lambda\sigma}{\delta}(1 - e^{-\delta(s-t)})\end{aligned}$$

Ce rendement moyen est croissant en fonction de l'échéance et présente, à long terme, une asymptote horizontale (voir figure 4.7).

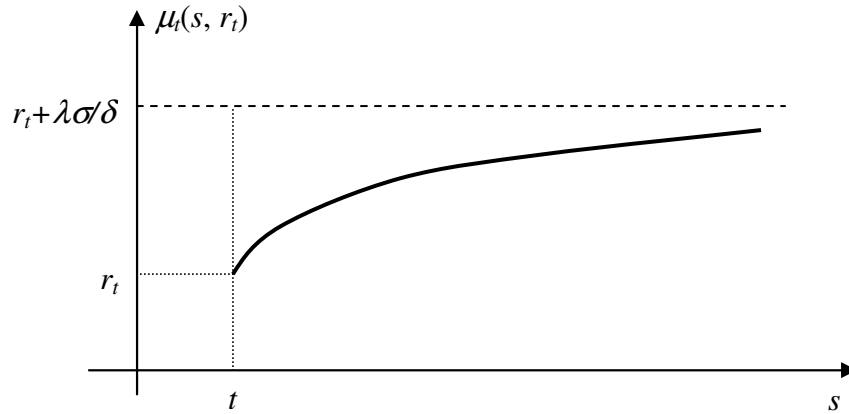


Figure 4.7 – Modèle de Vasicek : rendement moyen instantané

Le taux au comptant est, lui, donné par

$$\begin{aligned}R_t(s, r_t) &= -\frac{1}{s-t} \ln P_t(s, r_t) \\ &= k - \frac{k - r_t}{\delta(s-t)}(1 - e^{-\delta(s-t)}) + \frac{\sigma^2}{4\delta^3(s-t)}(1 - e^{-\delta(s-t)})^2\end{aligned}$$

Cette expression montre que, d'une part, le taux au comptant se stabilise pour les échéances lointaines, et ce indépendamment de la valeur initiale du taux spot :

$$\lim_{(s-t) \rightarrow +\infty} R_t(s, r_t) = k$$

et que, d'autre part, suivant la valeur actuelle du taux spot vis-à-vis des paramètres, on peut avec ce modèle représenter différentes allures de la courbe de taux : suivant que  $r_t$  appartient aux intervalles

$$\left[0; k - \frac{\sigma^2}{4\delta^2}\right], \quad \left[k - \frac{\sigma^2}{4\delta^2}; k + \frac{\sigma^2}{2\delta^2}\right], \quad \left[k + \frac{\sigma^2}{2\delta^2}; +\infty\right],$$

on obtient une courbe de taux respectivement croissante, bosselée ou décroissante.

### Illustration

La figure 4.8 représente des courbes de taux au comptant résultant du modèle de Vasicek pour les valeurs suivantes des paramètres :  $\delta = 0,2$ ,  $\theta = 0,08$ ,  $\sigma = 0,05$  et  $\lambda = 0,02$ . Les trois courbes correspondent, de bas en haut, à  $r_0 = 2\%$ ,  $r_0 = 6\%$  et  $r_0 = 10\%$ .

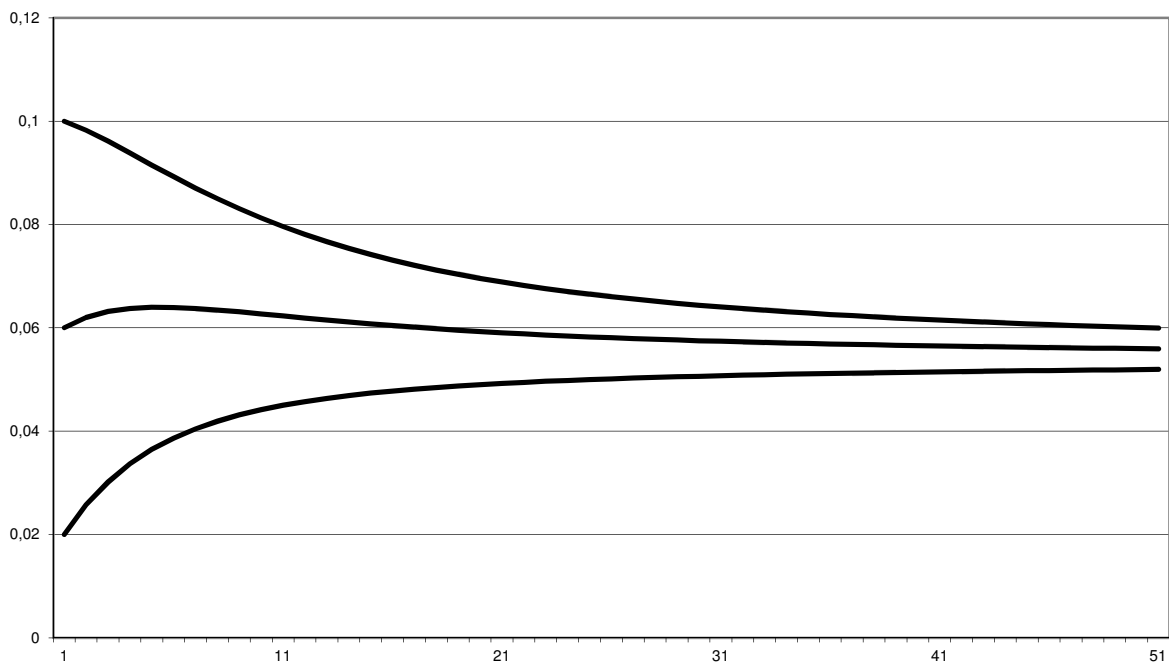


Figure 4.8 – Modèle de Vasicek : courbes de taux ♣

Le modèle de Vasicek présente deux inconvénients majeurs. D'une part, le processus d'Ornstein-Uhlenbeck à partir duquel il est construit

$$dr_t = \delta(\theta - r_t) \cdot dt + \sigma \cdot dw_t,$$

permet parfois, à cause du second terme, au taux spot de prendre des valeurs négatives. D'autre part, la fonction de taux au comptant qu'il génère  $R_t(s)$  peut, dans certains cas, prendre également des valeurs négatives.

#### 4.5.3. Modèle de Cox, Ingersoll et Ross<sup>116</sup>

Ce modèle fait partie d'une famille parfois appelée modèles d'équilibre car le point de départ est un raisonnement de type macroéconomique et le raisonnement s'appuie sur l'hypothèse que le consommateur a un comportement consistant à maximiser son utilité espérée.

Ces considérations – que nous ne détaillerons pas – conduisent, comme pour les autres modèles d'arbitrage, à préciser le processus stochastique gouvernant l'évolution du taux spot  $r_t$  ainsi que la prime de risque  $\lambda_t(r_t)$ .

Si, dans le processus d'Ornstein-Uhlenbeck, on modifie le deuxième terme de la manière suivante

$$dr_t = \delta(\theta - r_t) \cdot dt + \sigma r_t^\alpha \cdot dw_t,$$

avec  $\alpha > 0$ , on évite l'inconvénient signalé plus tôt : le taux spot ne peut plus devenir négatif. En effet, dès qu'il devient nul, seul le premier terme subsiste et la variation de taux s'effectue alors nécessairement vers le haut : l'axe horizontal agit alors comme une « barrière » repoussante.

Dans le raisonnement macroéconomique à la base du modèle de Cox, Ingersoll et Ross, on a cette situation avec  $\gamma = 1/2$  et le processus stochastique porte le nom de *processus de racine carrée* :

$$dr_t = \delta(\theta - r_t) \cdot dt + \sigma \sqrt{r_t} \cdot dw_t.$$

Le même raisonnement conduit à une prime unitaire de risque de taux donnée par

$$\lambda_t(r_t) = \frac{\gamma}{\sigma} \sqrt{r_t},$$

où  $\gamma$  est une constante strictement positive ; la prime de risque augmente donc, dans ce cas, avec le taux spot.

#### *Illustration*

La figure 4.9 représente un tel processus de racine carrée avec les paramètres  $\delta = 100$ ,  $\theta = 0,1$  et  $\sigma = 0.8$ .

---

<sup>116</sup> COX J., INGERSOLL J. et ROSS J., A theory of the term structure of interest rates, *Econometrica*, Vol. 53, n° 2, 1985, pp. 385-406.

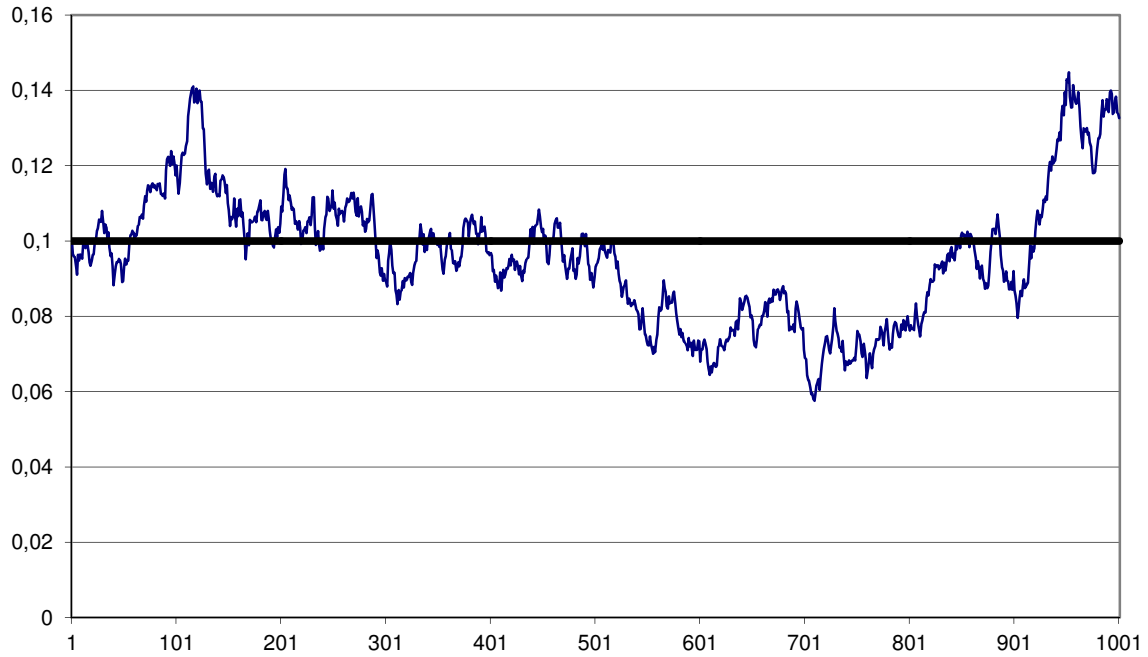


Figure 4.9 – Processus de racine carrée ♣

L'équation aux dérivées partielles pour le prix est ici donnée par

$$P'_t + (\delta(\theta - r) + \gamma r)P'_r + \frac{\sigma^2}{2} r P''_{rr} - rP = 0.$$

La solution de cette équation et de sa condition initiale est donnée par

$$P_t(s, r_t) = x_t(s) \cdot e^{-y_t(s)r},$$

où on a posé

$$\left\{ \begin{array}{l} x_t(s) = \left( \frac{2ke^{\frac{1}{2}(\delta - \gamma + k)(s-t)}}{z_t(s)} \right)^{\frac{2\delta\theta}{\sigma^2}} \\ y_t(s) = \frac{2(e^{k(s-t)} - 1)}{z_t(s)} \\ z_t(s) = 2k + (\delta - \gamma + k)(e^{k(s-t)} - 1) \\ k = \sqrt{(\delta - \gamma)^2 + 2\sigma^2} \end{array} \right.$$

Le taux de rendement moyen instantané est donné par

$$\begin{aligned}
\mu_t(s, r_t) &= \frac{P'_t + \delta(\theta - r_t)P'_r + \frac{\sigma^2}{2} r P''_{rr}}{P} \\
&= \frac{r_t \cdot P - \gamma r_t \cdot P'_r}{P} \\
&= r_t (1 + \gamma y_{t(s)})
\end{aligned}$$

Dans ce cas, le taux de rendement moyen est proportionnel au taux spot.

Finalement, le taux au comptant est donné par

$$\begin{aligned}
R_t(s, r_t) &= -\frac{1}{s-t} \ln P_t(s, r_t) \\
&= -\frac{1}{s-t} (\ln x_t(s) - r_t y_t(s))
\end{aligned}$$

### Illustration

La figure 4.10 représente des courbes de taux au comptant résultant du modèle de Cox, Ingersoll et Ross pour les valeurs suivantes des paramètres :  $\delta = 0,2$ ,  $\theta = 0,08$ ,  $\sigma = 0,05$  et  $\gamma = 0,02$ . Les trois courbes correspondent, de bas en haut, à  $r_0 = 2\%$ ,  $r_0 = 6\%$  et  $r_0 = 10\%$ .

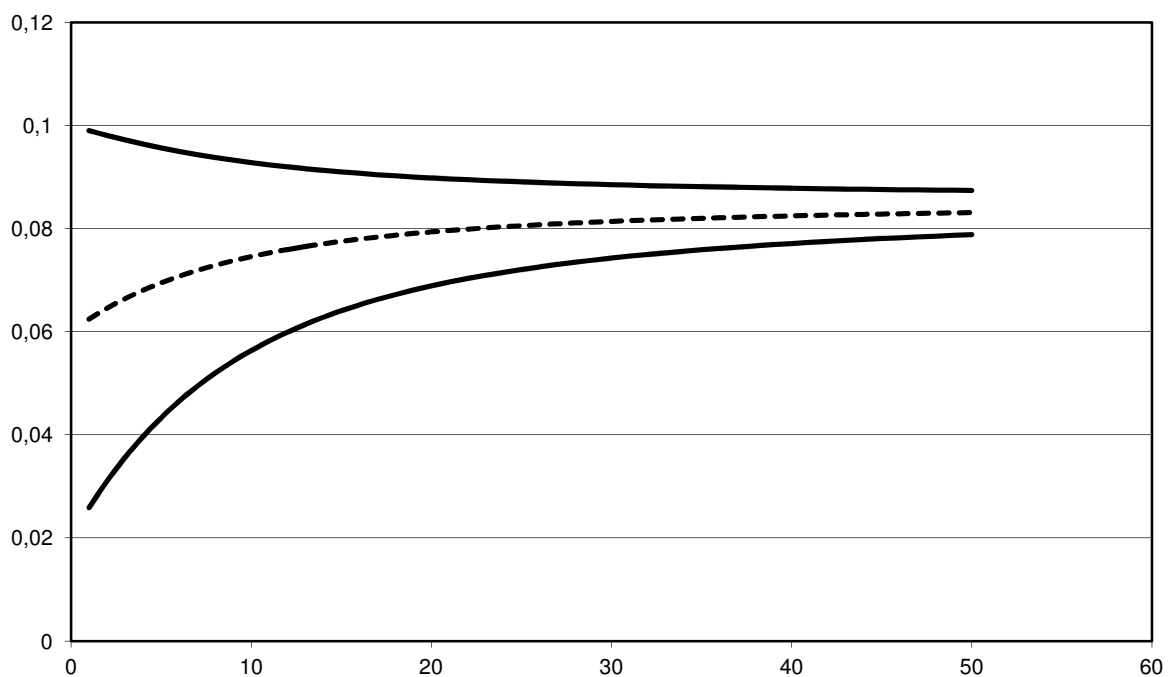


Figure 4.10 – Modèle de Cox, Ingersoll et Ross : courbes de taux♣

Signalons finalement que, contrairement au cas du modèle de Vasicek, le modèle de Cox, Ingersoll et Ross ne peut jamais conduire à des taux au comptant négatifs. Par ailleurs, comme pour le modèle de Vasicek, le taux au comptant se stabilise pour les échéances lointaines, et ce indépendamment de la valeur initiale du taux spot :

$$\lim_{(s-t) \rightarrow +\infty} R_t(s, r_t) = \frac{2\delta\theta}{\delta - \gamma + k}.$$

## CHAPITRE 5

### OPTIONS

#### 5.1. DEFINITIONS

##### 5.1.1. Caractéristiques

L'*option*<sup>117</sup> est un contrat qui confère à son acquéreur, contre paiement d'une *prime*, le droit d'acheter ou de vendre un actif (l'*actif sous-jacent*) à une date future à un prix déterminé d'avance (le *prix d'exercice* de l'option). L'option d'achat et l'option de vente portent respectivement le nom de *call* et de *put*. L'ensemble des actifs sur lesquels peut porter un contrat d'option est très large : action, obligation, taux de change, matière première ou encore certains produits dérivés tels que FRA, future, swap ou option.

L'option représente toujours un droit pour le détenteur, une obligation de vendre ou d'acheter pour l'émetteur. L'exercice de ce droit d'option peut se faire soit seulement à l'échéance du contrat (on parle alors d'*option européenne*) soit à n'importe quelle date jusqu'à cette échéance (*option américaine*). Le détenteur d'une option d'achat exercera donc son droit d'option si le prix de l'actif sous-jacent est supérieur au prix d'exercice de l'option (*strike* en anglais) ; une option de vente sera par contre utilement exercée dans le cas contraire.

Les actifs étudiés dans les deux chapitres précédents ont certes un comportement aléatoire (théorie moyenne-variance pour les actions, modèles de taux d'intérêt pour les obligations). Ils admettent néanmoins des approches déterministes également (formule de Gordon-Shapiro, duration et convexité). Dans le cas des options, le caractère aléatoire est plus intrinsèque encore puisque tout dépend d'une décision liée à un événement futur.

Ce genre de contrat peut constituer tant une source de profit (avec les risques liés à la spéculation) qu'une technique de couverture. Nous étairons notre propos en raisonnant sur une option d'achat européenne.

L'acquisition d'une telle option est susceptible de procurer des rendements intéressants puisque, dans le cas où le prix du sous-jacent à l'échéance est inférieur au prix d'exercice, l'option n'est pas exercée et la perte se limite au prix de l'option (la prime), alors qu'en cas de hausse du prix du sous-jacent au-delà du prix d'exercice, on reçoit le sous-jacent pour un prix inférieur à sa valeur.

---

<sup>117</sup> COLMANT B. et KLEYNEN G., *Gestion du risque de taux d'intérêt et instruments financiers dérivés*, Kluwer, 1995.

HULL J.C., *Options, futures, and others derivatives*, Prentice Hall, 1997.

HICKS A., *Foreign exchange options*, Woodhead, 1993.



La vente (l'émission) d'une option d'achat est par contre une opération beaucoup plus spéculative : le bénéfice se limitera à la prime si le prix du sous-jacent reste inférieur au prix d'exercice, alors que des pertes considérables pourront résulter d'une hausse plus importante du prix du sous-jacent. Cette opération n'est donc envisageable que si l'émetteur a une confiance absolue en la baisse (au pire la hausse modérée) du prix de l'actif sous-jacent.

### Illustration

Considérons une option d'achat sur un actif dont le prix actuel est de 100, la prime étant de 3 et le prix d'exercice étant fixé à 105. Calculons le gain réalisé (ou la perte subie) par l'acquéreur et par l'émetteur de ce contrat en fonction du prix atteint par le sous-jacent à l'échéance.

Prix du sous-jacent	Gain	
	acquéreur	émetteur
90	-3	3
95	-3	3
100	-3	3
105	-3	3
106	-2	2
107	-1	1
108	0	0
109	1	-1
110	2	-2
115	7	-7
120	12	-12

Tableau 5.1 – Gain sur option en fonction du prix du sous-jacent♣

Bien entendu, l'émetteur qui constate, en cours de contrat, la montée du prix du sous-jacent peut partiellement se prémunir contre cette perte en achetant la même option : il clôt sa position. Néanmoins, compte tenu du prix plus élevé du sous-jacent, la prime de l'option achetée pourrait être sensiblement plus élevée que celle de l'option qu'il avait émise.

Le prix (la prime) d'une option dépend de nombreux facteurs :

- le prix du sous-jacent  $S_t$  à l'instant  $t$  (*spot* en anglais) ;
- le prix d'exercice  $K$  de l'option ;
- la durée  $T - t$  restant à courir jusqu'à l'échéance de l'option<sup>118</sup> ;
- la volatilité  $\sigma_R$  du return du sous-jacent ;
- le taux  $R_F$  de l'actif sans risque<sup>119</sup>.

<sup>118</sup> On désignera souvent cette durée résiduelle  $T - t$  par la lettre  $\tau$  dans la suite.

<sup>119</sup> Cette description correspond par exemple à une option sur action (qui est celle qui nous intéressera principalement dans ce chapitre). Pour une option de change, le taux d'intérêt sera dédoublé : dans la devise domestique et dans la devise étrangère. Par ailleurs, ce taux – qui sera utilisé pour réaliser des actualisations –

La prime  $p$  peut donc s'écrire

$$p_t = f(S_t, K, T - t, \sigma_R, R_F).$$

Les différentes manières de spécifier la fonction  $f$  (qu'on notera  $C$  ou  $P$  selon que l'on a affaire à un call ou à un put) donnent lieu à ce que l'on appelle les modèles d'évaluation, dont le paragraphe 5.3 fait l'objet.

### 5.1.2. Utilisation

L'illustration qui a été présentée au paragraphe précédent correspond à la situation où l'acquéreur peut espérer des gains intéressants. Le profit réalisé est représenté graphiquement sur la figure 5.1.

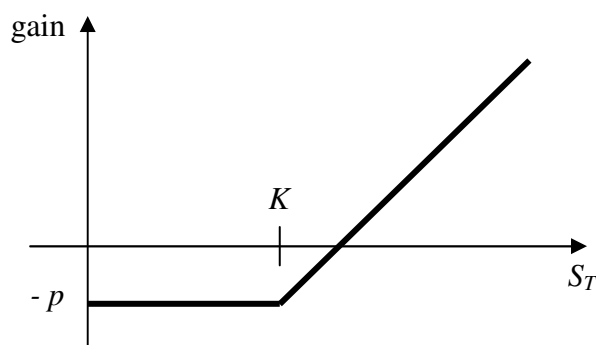


Figure 5.1 – Acquisition d'une option d'achat (call)

A côté de cet aspect spéculatif, l'émission d'une option d'achat peut devenir intéressante si elle est détenue simultanément avec l'actif sous-jacent. En effet, en cas de baisse (ou de hausse modérée) du prix du sous-jacent, la perte subie sur celui-ci est partiellement compensée par l'encaissement de la prime, alors qu'en cas de forte hausse, le bénéfice qui aurait été réalisé sur le sous-jacent se limite au prix de l'option augmenté du différentiel entre le prix d'exercice et le prix du sous-jacent à l'origine du contrat.

### Illustration 1

Poursuivant l'exemple introduit ci-dessus, calculons le gain réalisé (ou la perte subie) en cas de détention du sous-jacent seul et en cas de couverture de celui-ci par l'option d'achat.

---

peut être considéré de manière discrète (facteur d'actualisation  $(1 + R_F)^{-1}$ ) ou de manière continue (et on utilisera alors la notation  $r : e^{-rt}$ ).

<i>Prix du sous-jacent</i>	<i>Gain</i>	
	<i>seul</i>	<i>couverture</i>
90	-10	-7
95	-5	-2
100	0	3
105	5	8
106	6	8
107	7	8
108	8	8
109	9	8
110	10	8
115	15	8
120	20	8

Tableau 5.2 – Gain sur sous-jacent seul ou couvert ♣

### Illustration 2

Donnons maintenant un exemple plus réaliste. L'entreprise européenne X a souvent des factures libellées en USD, payables à la livraison. Bien sûr, les prix ont été fixés au moment de l'achat (bien avant la livraison). En cas de hausse du cours du dollar entre l'instant de l'achat et celui de la livraison, l'entreprise X subira une perte si elle achète ses dollars au moment du paiement.

Supposons de manière plus précise que le cours du dollar à l'instant  $t$  est de  $S_t$  (1 USD =  $S_t$  EUR) et que X achète aujourd'hui ( $t = 0$ ) des marchandises pour 1000 USD, le cours étant  $S_0 = x$  (1 USD =  $x$  EUR), livrables en  $t = T$ . L'entreprise X acquiert (en  $t = 0$ ) 1000 calls européens USD/EUR échéant en  $T$ , le prix d'exercice étant  $K = x$  EUR pour 1 USD.

- Si  $S_T > x$ , l'option sera exercée et X achètera ses dollars au cours  $x$  (celui du libellé de la facture) et l'entreprise ne perdra que le montant de la prime ;
- Si  $S_T \leq x$ , l'option ne sera pas exercée et X achètera ses dollars au cours  $S_T$  et l'entreprise réalisera un gain de  $1000 \cdot (x - S_T)$ , moins la prime.

L'acquisition de l'option d'achat agit ici comme une assurance contre les variations des cours. Celle-ci ne peut bien entendu pas être gratuite (considérez le point de vue de l'émetteur de l'option ...) ; son prix est la prime de l'option. ♣

Le cas envisagé ci-dessus correspond à l'acquisition d'une option d'achat. On peut effectuer le même genre de raisonnement pour quatre situations, correspondant d'une part à l'acquisition ou à l'émission d'une option d'achat ou de vente d'autre part. On a ainsi les figures 5.2 et 5.3.

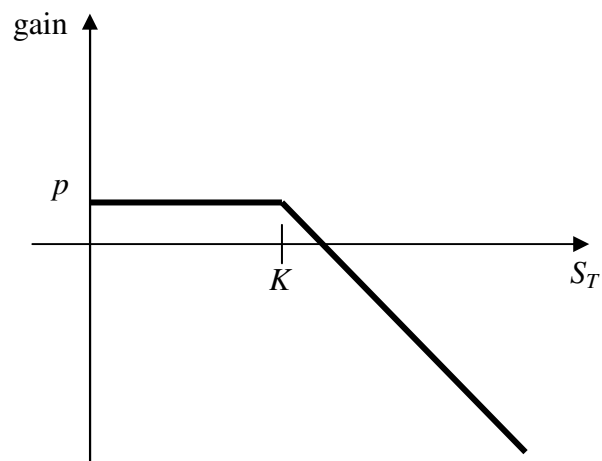


Figure 5.2 – Emission d’une option d’achat

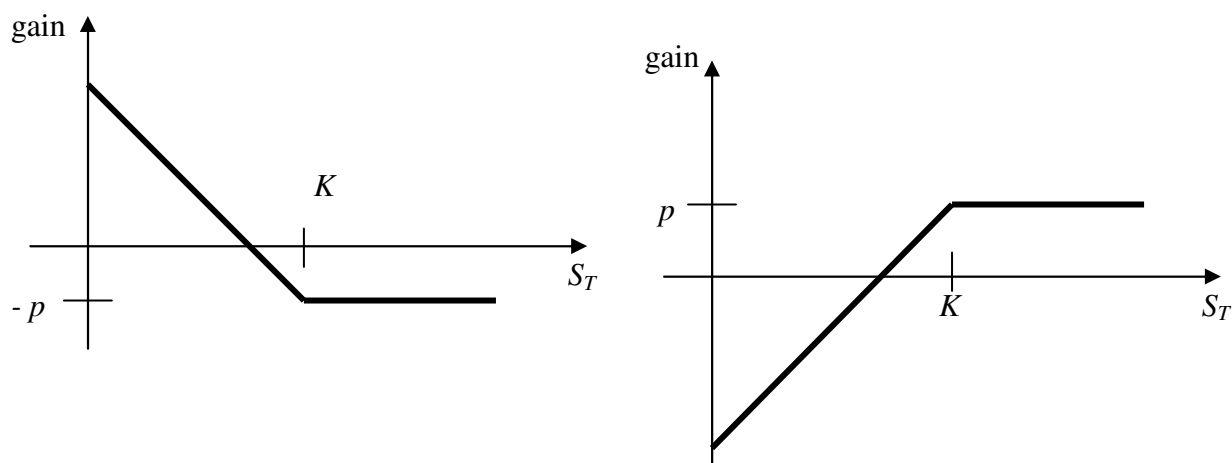


Figure 5.3 – Acquisition et émission d’une option de vente

A côté de la stratégie simple de couverture présentée ci-dessus, on peut réaliser des combinaisons plus complexes du sous-jacent, d’options d’achat et d’options de vente. Ces stratégies plus élaborées font l’objet du paragraphe 5.4.

## 5.2. VALEUR D’UNE OPTION

### 5.2.1. Valeur intrinsèque et valeur-temps

La prime d’une option peut se décomposer en deux termes : la valeur intrinsèque d’une option et sa valeur-temps.

La *valeur intrinsèque* d'une option à un instant  $t$  est simplement le gain réalisé par l'acquéreur (compte non tenu de la prime) si l'option était exercée en  $t$ . De manière plus précise, pour une option d'achat, il s'agit de la différence, si celle-ci est positive<sup>120</sup>, entre le prix du sous-jacent  $S_t$  à cet instant et le prix d'exercice<sup>121</sup>  $K$  de l'option ; si la différence en question est négative, la valeur intrinsèque est par définition 0. Pour une option de vente, la valeur intrinsèque est, cette fois, la différence entre le prix d'exercice et le prix du sous-jacent<sup>122</sup>. Si donc on note  $VI$  la valeur intrinsèque de l'option, on a

$$VI_t = \max(0, S_t - K) = (S_t - K)^+,$$

pour une option d'achat et

$$VI_t = \max(0, K - S_t) = (K - S_t)^+,$$

pour une option de vente, avec les représentations graphiques de la figure 5.4

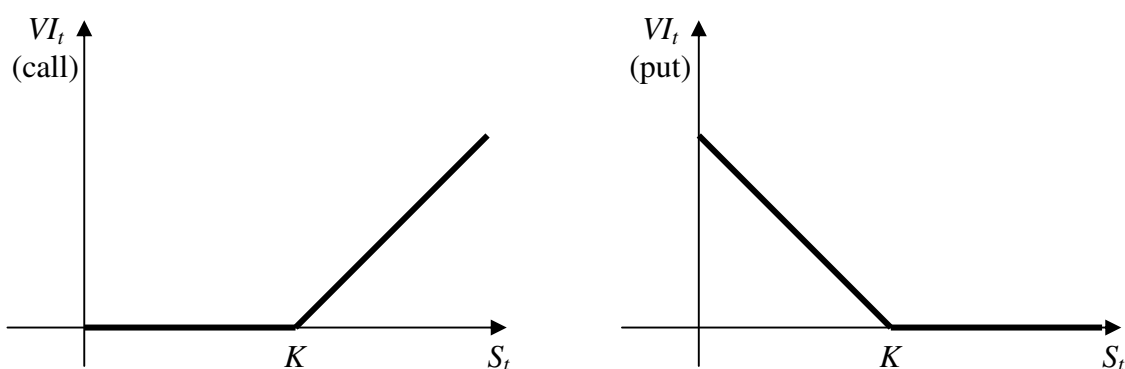


Figure 5.4 – Valeur intrinsèque d'une option d'achat et d'une option de vente

Le prix de l'option est bien évidemment au moins égal à sa valeur intrinsèque. La partie de la prime qui dépasse la valeur intrinsèque est appelée *valeur-temps* et est notée  $VT$  :

$$VT_t = p_t - VI_t.$$

Cette valeur-temps, qui vient s'ajouter à la valeur intrinsèque pour donner la prime, représente la rémunération de l'espoir pour l'acquéreur d'un gain supplémentaire. Du point de vue de l'émetteur, elle représente donc en quelque sorte une prime de risque.

La valeur-temps décroît bien entendu lorsque le temps qui reste à courir diminue, pour finir par s'annuler à l'échéance (voir figure 5.5).

<sup>120</sup> On dit alors que l'option est dans la monnaie (en anglais *in-the-money*). Si la différence est négative, l'option est dite hors de la monnaie (*out-of-the-money*) ; si le prix du sous-jacent est égal ou voisin du prix d'exercice, elle est dite à la monnaie (*at-the-money*). Ces définitions s'inversent pour une option de vente.

<sup>121</sup> L'option ne peut en fait être exercée immédiatement que si elle est de type américain. Pour une option européenne, le prix d'exercice doit normalement être actualisé sur la durée restant à courir jusqu'à l'échéance.

<sup>122</sup> Cette définition est donnée pour une option américaine ; pour une option européenne, il suffit de remplacer, dans l'interprétation de  $S_t$ , le prix à l'instant  $t$  par le prix à l'échéance.

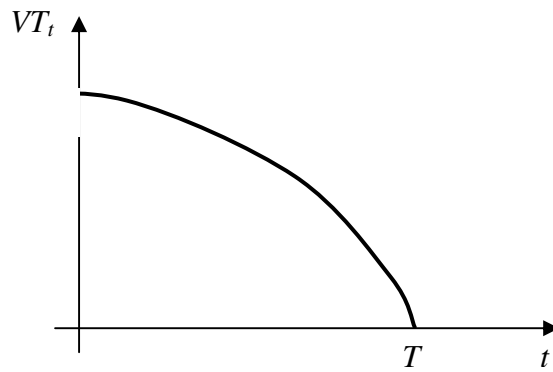


Figure 5.5 – Valeur-temps en fonction du temps

Il est facile de voir que – les autres paramètres étant constants – la valeur-temps est d’autant plus grande que le prix du sous-jacent est proche du prix d’exercice, comme représenté sur la figure 5.6.

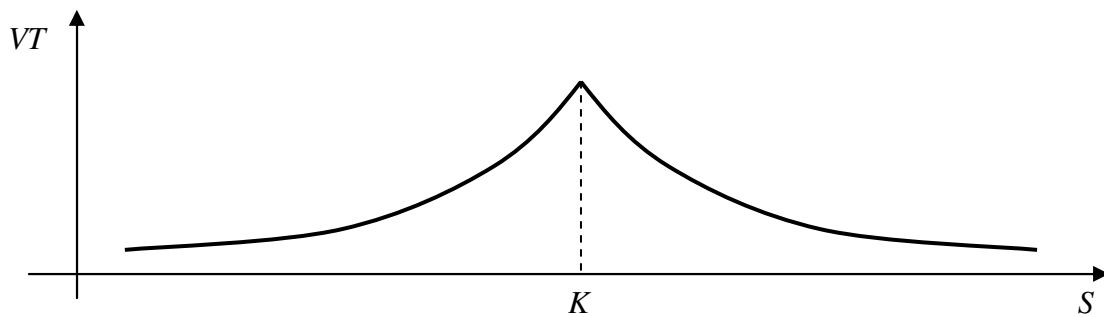


Figure 5.6 – Valeur-temps en fonction du prix de sous-jacent

Pour justifier cette propriété, plaçons-nous pour fixer les idées du point de vue de l’émetteur d’un call. Si l’option est out-of-the-money, elle ne sera vraisemblablement pas exercée et l’émetteur peut se dispenser d’acquérir le sous-jacent ; son risque (forte hausse du prix de ce sous-jacent) est donc faible et il ne devra guère être rémunéré. De la même manière, une option in-the-money sera vraisemblablement exercée, l’émetteur aura donc intérêt à acquérir le sous-jacent ; une forte baisse du prix de ce dernier représente un risque fort improbable et la valeur-temps sera faible également. A l’opposé, pour une option at-the-money, l’émetteur n’a aucune certitude quant à l’exercice ou non de l’option, ni quant à l’évolution du prix du sous-jacent ; le risque d’une chute du prix de ce dernier s’il a acheté le sous-jacent (ou d’une flambée du prix sans avoir acquis le sous-jacent) est donc important et il exigera une prime de risque en conséquence. Ce phénomène est schématisé sur la figure 5.7.

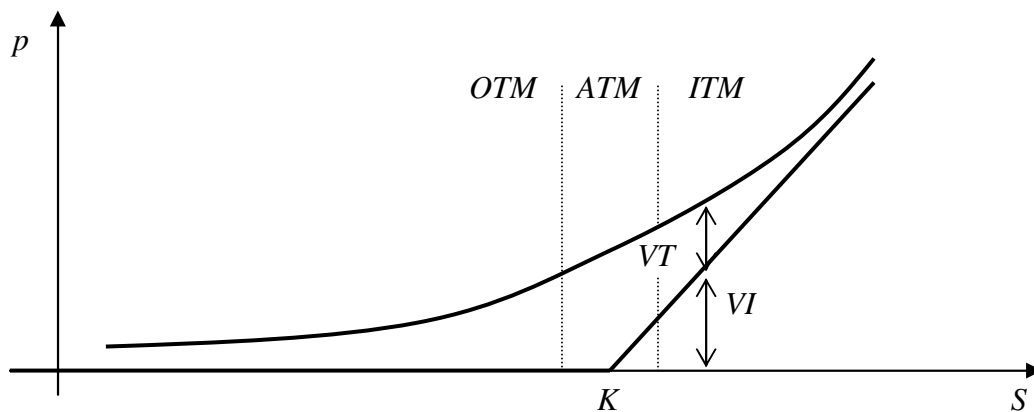


Figure 5.7 – Décomposition de la prime d'une option d'achat

Par ailleurs, il est clair que plus le délai à courir jusqu'à l'échéance du contrat d'option est long, plus les risques sont importants et plus la valeur-temps est grande (voir figure 5.8).

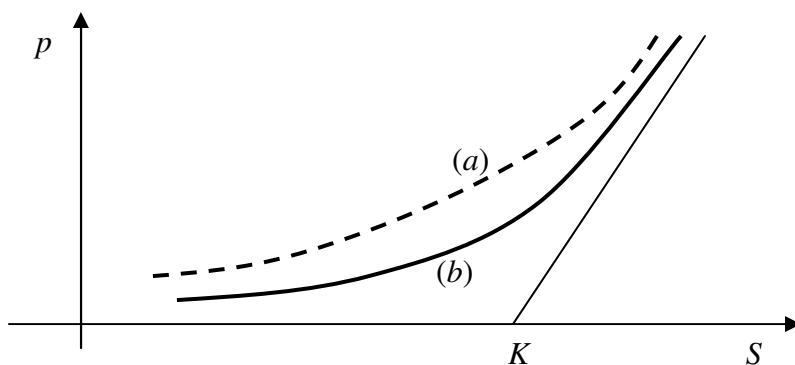


Figure 5.8 – Prime d'un call et maturité longue (a) et courte (b)

Bien entendu, la valeur d'une option à l'échéance est identique à sa valeur intrinsèque :

$$\begin{cases} C_T = (S_T - K)^+ \\ P_T = (K - S_T)^+ \end{cases}$$

### 5.2.2. Volatilité

Parmi les paramètres définissant le prix d'une option, examinons plus spécifiquement la volatilité  $\sigma_R$  du return du sous-jacent. On définit la *volatilité* d'une option comme étant une mesure de la dispersion du return de l'actif sous-jacent. En pratique, elle est généralement ramenée à une durée de référence d'une année et est exprimée en pourcentage. Cette notion de volatilité peut être envisagée de deux points de vue : la volatilité historique et la volatilité implicite.

La *volatilité historique* est simplement l'écart type annualisé du return du sous-jacent, obtenu à partir d'observations journalières de ce return dans le passé :

$$\sigma_R = \sqrt{J \cdot \frac{1}{n} \sum_{t=1}^n (R_t - \bar{R})^2},$$

où le facteur  $J$  représente le nombre de jours ouvrables de l'année,  $n$  est le nombre d'observations et  $R_t$  est le return du sous-jacent. Elle est simple à calculer, mais elle présente l'inconvénient majeur d'être « tournée vers le passé » alors qu'elle doit servir à l'analyse de l'évolution future du prix de l'option.

C'est pourquoi la notion de *volatilité implicite* a été introduite. Il s'agit là de l'estimation, avec l'aide d'un modèle d'évaluation, de la dispersion du return du sous-jacent pour la durée restant à courir jusqu'à l'échéance du contrat. La valeur de la prime de l'option est déterminée en pratique par la loi de l'offre et de la demande. Par ailleurs, cette dernière est reliée à différents facteurs par un modèle d'évaluation

$$p_t = f(S_t, K, T - t, \sigma_R, R_F),$$

du type binomial ou encore Black et Scholes (voir § 5.3). La résolution de cette relation par rapport à  $\sigma_R$  définit la volatilité implicite. Bien que d'un accès plus compliqué, cette notion est préférable et c'est celle qui sera souvent utilisée en pratique.

### 5.2.3. Paramètres de sensibilité

#### *Les « grecques »*

Des variations de prime sont susceptibles de se produire suite à des fluctuations de chacun des paramètres déterminant le prix de l'option (prix spot, prix d'exercice, maturité, ...). C'est l'objectif de ce paragraphe d'étudier des indices<sup>123</sup> – appelés en anglais « the greeks » – mesurant la sensibilité de la prime aux variations d'un certain nombre de ces caractéristiques par l'intermédiaire de la relation

$$p_t = f(S_t, K, \tau, \sigma_R, R_F).$$

Nous nous contenterons d'examiner ici les coefficients de sensibilité les plus utilisés : ceux qui mettent en relation le prix de l'option et le prix du sous-jacent, le temps, la volatilité et le taux sans risque. De plus, les indications de signe fournies sont valables pour une option sur action ne versant pas de dividende.

<sup>123</sup> A l'instar de la duration et de la convexité qui mesurent la sensibilité de la valeur d'une obligation suite à des fluctuations de taux d'intérêt (voir chapitre 4).



Le coefficient  $\Delta$  (*delta*) représente la sensibilité du prix de l'option par rapport à celui du sous-jacent. Il se mesure par le rapport des variations de ces deux prix pour un petit accroissement  $\delta S_t$  du prix du sous-jacent :

$$\Delta = \frac{f(S_t + \delta S_t, K, \tau, \sigma_R, R_F) - f(S_t, K, \tau, \sigma_R, R_F)}{\delta S_t}$$

ou, de manière plus précise,

$$\begin{aligned}\Delta &= \lim_{\delta S_t \rightarrow 0} \frac{f(S_t + \delta S_t, K, \tau, \sigma_R, R_F) - f(S_t, K, \tau, \sigma_R, R_F)}{\delta S_t} \\ &= f'_S(S_t, K, \tau, \sigma_R, R_F)\end{aligned}$$

Ainsi, pour un call, si le prix du sous-jacent augmente de 1 EUR, le prix de l'option augmentera de  $\Delta$  EUR. Il est compris entre 0 et 1 pour un call et entre  $-1$  et 0 pour un put.

Un autre coefficient exprime la sensibilité du prix de l'option par rapport à celui du sous-jacent, mais au second ordre cette fois. Il s'agit du coefficient  $\Gamma$  (*gamma*) qui s'exprime par le rapport des variations de  $\Delta$  d'une part et du prix  $S_t$  d'autre part :

$$\Gamma = f''_{SS}(S_t, K, \tau, \sigma_R, R_F).$$

Si on veut comparer la dépendance de la prime d'option vis-à-vis du prix du sous-jacent et celle du prix d'une obligation en fonction du taux actuariel, on peut dire que  $\Delta$  est à la duration ce que  $\Gamma$  est à la convexité. Ce coefficient  $\Gamma$ , toujours positif, est le même pour une option d'achat et pour une option de vente.

Le coefficient suivant, noté  $\Theta$  (*thêta*), mesure la dépendance du prix de l'option en fonction de la durée courante du contrat. Si on note celle-ci  $t$ , on a donc

$$\Theta = f'_t(S_t, K, T-t, \sigma_R, R_F)$$

ou, en introduisant la durée  $\tau$  à courir jusqu'à l'échéance,

$$\Theta = -f'_\tau(S_t, K, \tau, \sigma_R, R_F).$$

Lorsque l'on s'approche de l'échéance du contrat d'option, la valeur de celui-ci diminue, ce qui implique que  $\Theta$  est généralement négatif.

Le coefficient  $V$  (*véga*)<sup>124</sup> mesure quant à lui la sensibilité de la prime d'option par rapport à la volatilité :

$$V = f'_\sigma(S_t, K, \tau, \sigma_R, R_F).$$

<sup>124</sup> Parfois aussi noté  $\kappa$  (*kappa*). Parce que véga n'est pas une lettre grecque ?

Il est toujours positif et a la même valeur pour un call et pour un put. Il s'interprète bien sûr de la manière suivante : si la volatilité augmente de 1 %, le prix de l'option s'accroît de  $V$ .

Enfin, le coefficient  $\rho$  (*rhô*) exprime la manière dont le prix de l'option dépend du taux sans risque  $R_F$  :

$$\rho = f'_{R_F}(S_t, K, T - t, \sigma_R, R_F).$$

Ce coefficient est positif ou négatif selon que l'on a affaire à un call ou à un put.

### ***Le « delta hedging »***

Ces coefficients étant définis, donnons une interprétation pratique intéressante du delta. Il s'agit là d'un élément intervenant dans la couverture d'une position courte (émission) d'option d'achat (on parle de « *delta hedging* »). La question est : combien l'émetteur d'un call doit-il acquérir d'unités de l'actif sous-jacent de manière à couvrir sa position ? Appelons  $X$  cette quantité. Si la valeur actuelle de l'actif sous-jacent est  $S$ , la valeur de son portefeuille constitué de

- l'achat de  $X$  unités de l'actif sous-jacent ;
- l'émission d'un call sur ce sous-jacent ;

est de

$$V(S) = X \cdot S - C(S).$$

Si le prix du sous-jacent passe à  $S + \delta S$ , la valeur du portefeuille passe à

$$V(S + \delta S) = X \cdot (S + \delta S) - C(S + \delta S).$$

Comme

$$\Delta \approx \frac{C(S + \delta S) - C(S)}{\delta S},$$

la nouvelle valeur du portefeuille est

$$\begin{aligned} V(S + \delta S) &= X \cdot (S + \delta S) - [C(S) + \Delta \cdot \delta S] \\ &= X \cdot S - C(S) + (X - \Delta) \cdot \delta S \\ &= V(S) + (X - \Delta) \cdot \delta S \end{aligned}$$

La position sera donc couverte suite à un mouvement (à la hausse ou à la baisse) du prix du sous-jacent si le deuxième terme est nul ( $X = \Delta$ ) c'est-à-dire si l'émetteur du call détient  $\Delta$  unités de l'actif sous-jacent.

## 5.2.4. Propriétés générales

### *Relation de parité call-put pour une option européenne*

Nous allons ici établir une relation liant la prime d'un call et la prime d'un put européens, portant tous deux sur le même sous-jacent et ayant le même prix d'exercice et la même échéance : la *relation de parité « call-put »*.

Nous allons établir cette relation pour une option européenne sur action ne distribuant pas de dividende pendant la durée du contrat d'option.

Considérons un portefeuille constitué à l'instant  $t$  par

- l'achat du sous-jacent, dont la valeur est  $S_t$  ;
- l'achat d'un put sur ce sous-jacent, de prix d'exercice  $K$  et d'échéance  $T$  ; sa valeur est donc  $P_t(S_t, K, \tau, \sigma_R, R_F)$  ;
- la vente d'un call sur ce même sous-jacent, de prix d'exercice  $K$  et d'échéance  $T$  ; sa valeur est  $C_t(S_t, K, \tau, \sigma_R, R_F)$  ;
- l'emprunt (au taux sans risque  $R_F$ ) d'un montant qui vaudra  $K$  à l'époque  $T$  ; ce montant est donc  $K \cdot (1 + R_F)^{-\tau}$ .

La valeur de ce portefeuille à l'échéance  $T$  sera  $S_T + P_T - C_T - K$ . Comme on a montré précédemment que  $C_T = (S_T - K)^+$  et que  $P_T = (K - S_T)^+$ , cette valeur à l'échéance est égale à

$$\begin{cases} \text{si } S_T > K, & S_T + 0 - (S_T - K) - K = 0 \\ \text{si } S_T \leq K, & S_T + (K - S_T) - 0 - K = 0 \end{cases}$$

Ce portefeuille a, quelle que soit l'évolution de la valeur du sous-jacent entre  $t$  et  $T$  et pour  $K$  et  $R_F$  constants, une valeur nulle à l'instant  $T$ . En vertu de l'hypothèse d'absence d'opportunité d'arbitrage<sup>125</sup>, le portefeuille ne peut donc avoir qu'une valeur nulle à l'instant  $t$ .

La nullité de la valeur de ce portefeuille à l'instant  $t$  s'exprime par

$$S_t + P_t - C_t - K \cdot (1 + R_F)^{-\tau} = 0,$$

ou encore, de manière plus classique,

$$C_t + K \cdot (1 + R_F)^{-\tau} = P_t + S_t,$$

qui est la relation de parité annoncée.

### *Remarque*

<sup>125</sup> Rappelons qu'aucun flux financier n'est intervenu entre  $t$  et  $T$  puisque nous avons exclu le versement de dividendes.

La relation de parité « call-put » n'est pas valable pour une option de change en raison du différentiel d'intérêt entre les deux devises. Si on désigne respectivement par  $R_F^{(D)}$  et par  $R_F^{(F)}$  le taux d'intérêt sans risque de la devise domestique et de la devise étrangère (on les suppose encore constants et valables pour toute échéance), il est facile de voir que la relation de parité prend la forme

$$C_t + K \cdot (1 + R_F^{(D)})^{-\tau} = P_t + S_t \cdot (1 + R_F^{(F)})^{-\tau} \cdot \clubsuit$$

### ***Relation entre call européen et call américain***

Etablissons maintenant une relation liant la valeur d'un call européen et celle d'un call américain, portant tous deux sur le même sous-jacent et ayant le même prix d'exercice et la même échéance. Comme pour la relation de parité, nous allons traiter uniquement le cas d'options sur action ne distribuant pas de dividende pendant la durée du contrat d'option.

L'option américaine pouvant s'exercer à n'importe quel moment avant l'échéance, sa valeur est toujours au moins égale à la valeur de l'option européenne de mêmes caractéristiques :

$$C_t^{(a)}(S_t, K, T-t, \sigma_R, R_F) \geq C_t^{(e)}(S_t, K, T-t, \sigma_R, R_F).$$

La relation de parité permet d'écrire successivement

$$\begin{aligned} C_t^{(e)} + K \cdot (1 + R_F)^{-\tau} &= P_t^{(e)} + S_t \\ C_t^{(e)} &\geq S_t - K \cdot (1 + R_F)^{-\tau} > S_t - K \\ C_t^{(a)} &\geq C_t^{(e)} > (S_t - K)^+ \end{aligned}$$

Comme  $(S_t - K)^+$  représente ce que rapporterait le call américain s'il était exercé à l'instant  $t$ , son détenteur a tout intérêt à le conserver jusqu'à l'instant  $T$ . Cette option aura donc à tout moment la même valeur que l'option européenne correspondante :

$$C_t^{(a)} = C_t^{(e)} \quad \forall t \in [0; T].$$

Signalons que cette identité entre call américain et call européen n'est pas vraie pour les puts, ni pour les autres types d'option (de change par exemple).

### ***Inégalités sur les prix***

Les valeurs des calls et des puts obéissent aux inégalités suivantes :

$$\begin{aligned} [S_t - K(1 + R_F)^{-\tau}]^+ &\leq C_t \leq S_t \\ [K(1 + R_F)^{-\tau} - S_t]^+ &\leq P_t^{(e)} \leq K(1 + R_F)^{-\tau} \\ [K - S_t]^+ &\leq P_t^{(a)} \leq K \end{aligned}$$

Ces inégalités limitent les zones où se trouve le graphe du prix de l'option en fonction du prix du sous-jacent. On a ainsi la figure 5.9 pour un call (européen ou américain) et la figure 5.10 pour les puts.

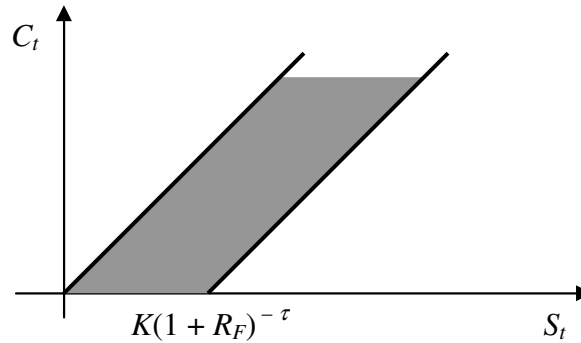


Figure 5.9 – Inégalités pour la valeur d'un call

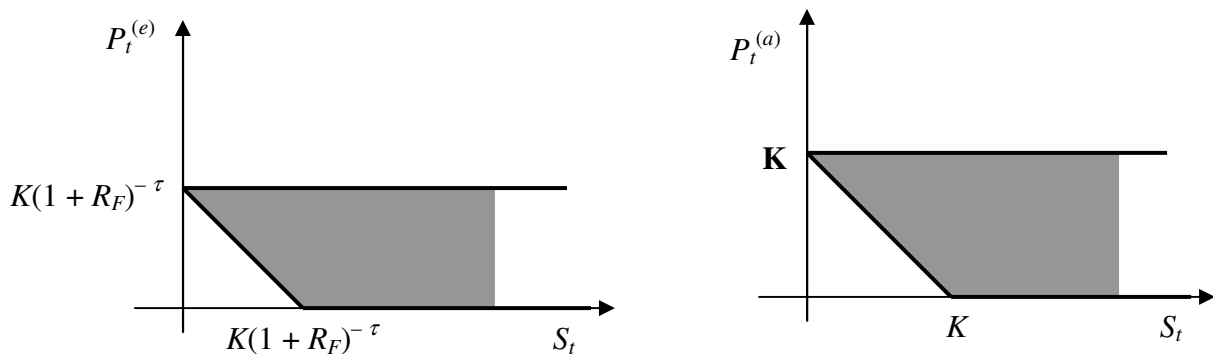


Figure 5.10 – Inégalités pour la valeur d'un put européen et d'un put américain

Les inégalités de droite sont évidentes : elles expriment simplement qu'une option ne peut valoir plus que ce qu'elle permet d'obtenir. Ainsi, un call ne peut valoir plus que le sous-jacent qu'il permet d'acquérir. De même, un put ne peut valoir plus que le prix d'exercice  $K$  auquel il permet de vendre le sous-jacent ; il ne peut même, pour un put européen, dépasser la valeur actualisée du prix d'exercice en question (l'exercice ne pouvant avoir lieu qu'à l'échéance).

Justifions maintenant l'inégalité de gauche pour un call. Constituons pour ce faire à l'instant  $t$  un portefeuille composé de

- l'achat d'un call ;
- un placement financier sans risqué qui vaudra  $K$  à l'échéance :  $K(1 + R_F)^{-\tau}$  ;
- la vente d'une unité du sous-jacent.

Sa valeur à l'instant  $t$  est bien sûr égale à

$$V_t = C_t + (1 + R_F)^{-\tau} - S_t.$$

Sa valeur à l'échéance dépendra de l'évolution du sous-jacent :

$$V_T = \begin{cases} \text{si } S_T > K, & (S_T - K) + K - S_T = 0 \\ \text{si } S_T \leq K, & 0 + K - S_T \end{cases}$$

c'est-à-dire  $V_T = (K - S_T)^+$ , qui n'est négatif pour aucun des scénarios d'évolution possibles. En l'absence d'opportunité d'arbitrage, on a aussi  $V_t \geq 0$ , c'est-à-dire

$$C_t \geq S_t - (1 + R_F)^{-\tau}.$$

Comme le prix de l'option ne peut être négatif, on a l'inégalité annoncée.

L'inégalité de gauche pour un put européen s'obtient de la même manière, par un raisonnement d'arbitrage à partir du portefeuille composé de

- l'achat d'un put ;
- l'achat d'une unité du sous-jacent ;
- l'emprunt d'un montant qui vaudrait  $K$  à l'échéance :  $K(1 + R_F)^{-\tau}$ .

L'inégalité de gauche pour un put américain résulte de celle pour le put européen en notant qu'il n'y a pas lieu d'actualiser le prix d'exercice, l'instant où le droit d'option étant exercé étant inconnu.

### 5.3. MODELES D'EVALUATION

Avant d'aborder les méthodes évoluées de détermination de la valeur d'une option, dégageons, à l'aide d'un exemple volontairement simplifié à l'extrême, le principe fondamental de la tarification des options.

#### *Illustration*

Considérons une option européenne d'achat sur le taux de change USD/EUR dont le prix d'exercice est  $K = 1$  et supposons qu'à l'instant présent ( $t = 0$ ), le cours soit de  $S_0 = 0,95$  (1 USD = 0,95 EUR). Nous travaillerons avec un taux sans risque nul ( $R_F = 0$ ) afin de simplifier les développements.

Supposons par ailleurs que l'évolution, aléatoire, du sous-jacent entre les instants  $t = 0$  et  $t = T$  puisse correspondre à deux scénarios  $s_1$  et  $s_2$  pour lesquels  $S_T$  vaut respectivement 1,1 EUR et 0,9 EUR et que ces scénarios surviennent avec les probabilités respectives 0,6 et 0,4 :

$$S_T = \begin{cases} 1,1 & \Pr(s_2) = 0,6 \\ 0,9 & \Pr(s_1) = 0,4 \end{cases}$$

L'évolution, aléatoire également, de l'option de change sera donc décrite par

$$C_T = \begin{cases} 0,1 & \Pr(s_2) = 0,6 \\ 0 & \Pr(s_1) = 0,4 \end{cases}$$

Considérons que l'on constitue, à l'instant  $t = 0$ , un portefeuille composé de

- l'émission d'un call USD/EUR (au prix actuel de  $C_0$ ) ;
- l'emprunt de  $X$  EUR ;
- l'achat de  $Y$  USD ;

de manière telle que

- la valeur initiale  $V_0$  du portefeuille est nulle : l'achat des  $Y$  USD se fait exactement avec ce qu'ont donné l'émission du call et l'emprunt ;
- le portefeuille soit sans risque : il subira la même évolution quel que soit le scénario (en fait sa valeur n'évoluera pas puisque nous avons supposé  $R_F$  nul).

La valeur initiale en EUR de ce portefeuille est donc de

$$V_0 = -C_0 - X + 0,95 Y = 0.$$

Suivant le scénario, la valeur finale sera donnée par

$$\begin{cases} V_T(s_1) = -0,1 - X + 1,1 \cdot Y \\ V_T(s_2) = -X + 0,9 \cdot Y \end{cases}$$

L'hypothèse d'absence d'opportunité d'arbitrage permet d'affirmer que  $V_T(s_1) = V_T(s_2) = 0$  et donc de déduire les valeurs suivantes :  $X = 0,45$  et  $Y = 0,5$ . A partir de la valeur initiale du portefeuille, on déduit donc la valeur initiale de l'option :

$$C_0 = -X + 0,95 Y = 0,025.$$

Il est important de remarquer que cette valeur est totalement indépendante des probabilités 0,6 et 0,4 associées aux deux scénarios d'évolution du prix du sous-jacent, sinon on aurait  $C_0 = 0,1 \cdot 0,6 + 0 \cdot 0,4 = 0,06$ .

Si maintenant on détermine une autre loi de probabilité

$$\Pr(s_1) = q \quad \Pr(s_2) = 1 - q$$

pour laquelle  $C_0 = E_q(C_T)$ , on a

$$0,025 = 0,1 \cdot q + 0 \cdot (1 - q),$$

c'est-à-dire  $q = 0,25$ .

Il s'agit précisément là de la loi de probabilité pour laquelle  $S_0 = E_q(S_T)$  :

$$E_q(S_T) = 1,1 \cdot 0,25 + 0,9 \cdot 0,75 = 0,95 = S_0 \cdot \clubsuit$$

Nous venons donc de voir – dans un cas très particulier et cela demande à être généralisé – que la valeur actuelle de l'option est égale à l'espérance mathématique de sa valeur future, par rapport à la loi de probabilité pour laquelle la valeur actuelle du sous-jacent est égale à l'espérance de sa valeur future<sup>126</sup>. Cette loi de probabilité porte le nom de *probabilité risque-neutre*.

### 5.3.1. Modèle binomial pour les options sur action

Ce modèle est dû à Cox, Ross et Rubinstein<sup>127</sup>. Dans ce modèle discret, on considère uniquement une liste d'époques  $0, 1, 2, \dots, T$ , séparées par une certaine unité de durée (la *période*), généralement assez courte.

Nous nous plaçons dans un marché parfait et nous envisageons une option européenne sur action, ne distribuant pas de dividende pendant la durée du contrat et dont la volatilité est constante pendant cette durée.

Nous supposons encore que le taux d'intérêt sans risque ne varie pas pendant cette période, qu'il est valable pour n'importe quelle échéance (courbe des taux plate et constante) et qu'il est le même pour un emprunt et pour un placement. Ce taux d'intérêt, noté  $R_F$ , sera exprimé en fonction d'une durée égale à une période ; il en sera de même des autres paramètres (return, volatilité, ...).

Rappelons (voir § 3.4.2) que l'évolution, d'époque en époque, de la valeur de l'action sous-jacente se réalise de manière dichotomique : l'action  $a$ , à l'instant  $t$ , la valeur  $S_t$  et aura à l'instant suivant  $t + 1$  une des deux valeurs  $S_t \cdot u$  (supérieure à  $S_t$ ) ou  $S_t \cdot d$  (inférieure à  $S_t$ ) avec les probabilités respectives  $\alpha$  et  $(1 - \alpha)$ . On a  $d \leq 1 < 1 + R_F \leq u$  et les paramètres  $u$ ,  $d$  et  $\alpha$ , qu'on suppose constants au cours du temps, sont à estimer à partir des observations.

On a la représentation graphique suivante de l'évolution du prix de l'action pour une période :

$$\begin{array}{lcl} & \nearrow & S_{t+1} = S_t \cdot u \quad (\alpha) \\ S_t & & \\ & \searrow & S_{t+1} = S_t \cdot d \quad (1 - \alpha) \end{array}$$

et donc, de manière plus générale :

<sup>126</sup> Lorsque le taux sans risque est nul, rappelons-le.

<sup>127</sup> COX J., ROSS S. et RUBINSTEIN M., Option pricing : a simplified approach, *Journal of Financial Economics*, n°7, 1979, pp. 229-263.



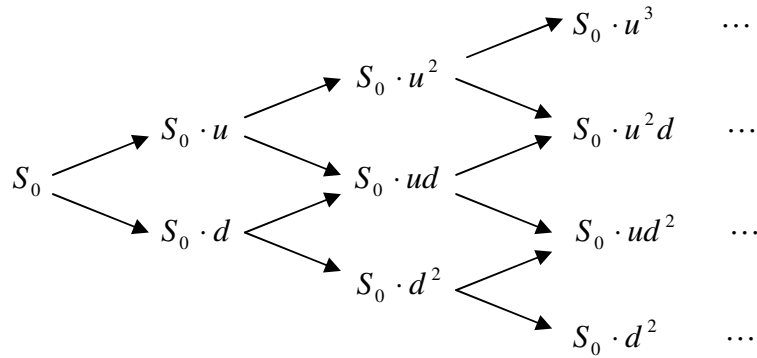


Figure 5.11 – Treillis binomial pour l'action sous-jacente

Abordons à présent la question de l'évaluation de l'option à l'instant initial. Nous raisonnerons pour une option d'achat.

On sait que la valeur de l'option au terme du contrat s'exprime en fonction de celle de l'action par

$$C_T = (S_T - K)^+.$$

Après avoir construit l'arbre pour l'action de l'instant 0 à l'instant  $T$ , nous allons, pour l'option, parcourir celui-ci de  $T$  à 0, à partir de chacune des extrémités de l'arbre de l'action afin de reconstituer la valeur  $C_0$  de l'option en 0.

Nous réaliserons ce raisonnement par étapes.

### Une période

Supposons que  $T = 1$ . A partir de l'arbre de l'action, on voit aisément que le call  $C_0$  (inconnu !) aura pu évoluer vers deux valeurs, avec les probabilités respectives  $\alpha$  et  $(1 - \alpha)$  :

$$\begin{array}{l} \nearrow C_1 = C(u) = (S_0 \cdot u - K)^+ \\ C_0 \searrow C_1 = C(d) = (S_0 \cdot d - K)^+ \end{array}$$

La valeur de  $C_1$  (c'est-à-dire celle de  $C(u)$  et de  $C(d)$ ) étant connue, nous allons déterminer celle de  $C_0$ . Nous allons pour ce faire construire un portefeuille constitué en  $t = 0$  par

- l'achat de  $X$  actions sous-jacentes, dont la valeur est  $S_0$  ;
- la vente d'un call sur ce sous-jacent, de valeur  $C_0$ .

La valeur  $V_0$  de ce portefeuille et son évolution  $V_1$  dans le contexte décrit sont données par

$$\begin{array}{l} \nearrow V_1 = X \cdot S_0 \cdot u - C(u) \\ V_0 = X \cdot S_0 - C_0 \searrow V_1 = X \cdot S_0 \cdot d - C(d) \end{array}$$

Choisissons ensuite  $X$  de manière à ce que ce portefeuille soit non risqué (les deux valeurs de  $V_1$  sont alors identiques). L'hypothèse d'absence d'opportunité d'arbitrage montre que dans ce cas, le rendement de ce portefeuille doit être donné par le taux sans risque  $R_F$ .

On obtient donc

$$\begin{cases} V_1 = X \cdot S_0 \cdot u - C(u) = X \cdot S_0 \cdot d - C(d) \\ V_1 = (X \cdot S_0 - C_0)(1 + R_F) \end{cases}$$

La première équation fournit aisément

$$X \cdot S_0 = \frac{C(u) - C(d)}{u - d},$$

et par conséquent

$$V_1 = \frac{d \cdot C(u) - u \cdot C(d)}{u - d}.$$

La seconde équation devient alors

$$\frac{d \cdot C(u) - u \cdot C(d)}{u - d} = \left( \frac{C(u) - C(d)}{u - d} - C_0 \right) (1 + R_F),$$

qui se résout aisément par rapport à  $C_0$  :

$$C_0 = (1 + R_F)^{-1} \left[ \frac{(1 + R_F) - d}{u - d} C(u) + \frac{u - (1 + R_F)}{u - d} C(d) \right].$$

Les coefficients de  $C(u)$  et  $C(d)$  sont manifestement compris entre 0 et 1 et ont pour somme 1. On posera donc

$$q = \frac{(1 + R_F) - d}{u - d} \quad 1 - q = \frac{u - (1 + R_F)}{u - d},$$

qui constituent la loi de probabilité risque-neutre ou encore *de couverture*.

Nous avons donc la valeur du call à l'origine :

$$C_0 = (1 + R_F)^{-1} [q \cdot C(u) + (1 - q) \cdot C(d)].$$

### Remarque 1

Comme on l'a fait remarquer lors de l'illustration introductive, la probabilité de croissance  $\alpha$  n'apparaît pas dans la relation ci-dessus. La seule loi de probabilité intervenant est celle

relative à la probabilité risque-neutre  $q$ , par rapport à laquelle  $C_0$  apparaît comme la valeur actualisée de la valeur moyenne du call à l'échéance ( $t = 1$ ).

L'appellation « probabilité risque-neutre » ou « probabilité de couverture » se justifie de la manière suivante : la valeur moyenne de l'action sous-jacente à l'échéance ( $t = 1$ ) par rapport à cette loi de probabilité est donnée par

$$\begin{aligned} E_q(S_1) &= q \cdot S_0 \cdot u + (1 - q) \cdot S_0 \cdot d \\ &= S_0 \left[ \frac{(1 + R_F) - d}{u - d} u + \frac{u - (1 + R_F)}{u - d} d \right] \\ &= S_0 (1 + R_F) \end{aligned}$$

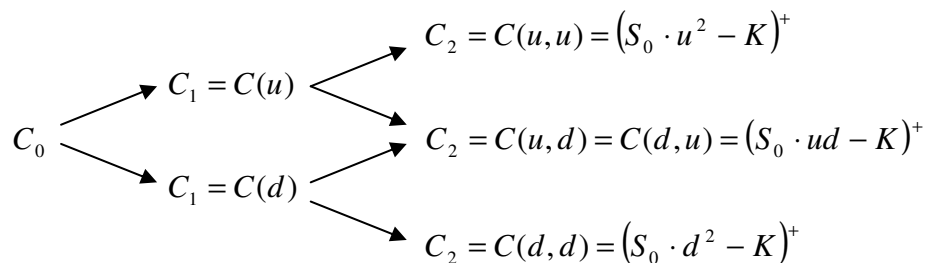
L'évolution du titre sans risque est la même que l'évolution moyenne (pour cette loi de probabilité) du titre risqué.♣

### Remarque 2

Dans l'utilisation pratique du modèle binomial, il est plus simple d'appliquer le raisonnement relatif à une seule période pour chacun des nœuds de l'arbre, en allant de  $T$  à 0. Nous poussons cependant l'analyse plus avant afin d'obtenir un résultat général.

### Deux périodes

Supposons maintenant que  $T = 2$ . L'arbre binomial pour l'option s'écrit alors



Le raisonnement précédent permet de passer de l'époque 2 à l'époque 1 :

$$\begin{cases} C(u) = (1 + R_F)^{-1} [q \cdot C(u, u) + (1 - q) \cdot C(u, d)] \\ C(d) = (1 + R_F)^{-1} [q \cdot C(u, d) + (1 - q) \cdot C(d, d)] \end{cases}$$

ainsi que de l'époque 1 à l'époque 0 :

$$\begin{aligned} C_0 &= (1 + R_F)^{-1} [q \cdot C(u) + (1 - q) \cdot C(d)] \\ &= (1 + R_F)^{-2} [q^2 \cdot C(u, u) + 2q(1 - q) \cdot C(u, d) + (1 - q)^2 \cdot C(d, d)] \end{aligned}$$

La considération des coefficients de  $C(u, u)$ ,  $C(u, d)$  et  $C(d, d)$  permet de préciser la remarque ci-dessus :  $C_0$  est la valeur actualisée de la valeur moyenne du call à l'échéance ( $t = 2$ ) par rapport à une loi de probabilité binomiale<sup>128</sup> de paramètres  $(2 ; q)$ .

### *T périodes*

En généralisant ce qui vient d'être fait, on voit que  $C_0$  est la valeur actualisée de la valeur moyenne du call à l'échéance ( $t = T$ ) par rapport à une loi de probabilité binomiale de paramètres  $(T ; q)$ . On peut donc écrire

$$\begin{aligned} C_0 &= (1 + R_F)^{-T} \sum_{j=0}^T \binom{T}{j} q^j (1-q)^{T-j} C(\underbrace{u, \dots, u}_j, \underbrace{d, \dots, d}_{T-j}) \\ &= (1 + R_F)^{-T} \sum_{j=0}^T \binom{T}{j} q^j (1-q)^{T-j} (S_0 u^j d^{T-j} - K)^+ \end{aligned}$$

Comme  $u^j d^{T-j}$  est une fonction croissante de  $j$ , si on introduit

$$J = \min\{j : S_0 u^j d^{T-j} - K > 0\},$$

c'est-à-dire la plus petite valeur de  $j$  strictement supérieure à

$$\frac{\ln K - \ln(S_0 d^T)}{\ln u - \ln d},$$

l'évaluation du call prend la forme

$$\begin{aligned} C_0 &= (1 + R_F)^{-T} \sum_{j=J}^T \binom{T}{j} q^j (1-q)^{T-j} (S_0 u^j d^{T-j} - K) \\ &= S_0 \sum_{j=J}^T \binom{T}{j} \left( \frac{uq}{1 + R_F} \right)^j \left( \frac{d(1-q)}{1 + R_F} \right)^{T-j} - K(1 + R_F)^{-T} \sum_{j=J}^T \binom{T}{j} q^j (1-q)^{T-j} \end{aligned}$$

Puisque

$$\frac{uq}{1 + R_F} + \frac{d(1-q)}{1 + R_F} = \frac{u[(1 + R_F) - d] + d[u - (1 + R_F)]}{(1 + R_F)(u - d)} = 1,$$

on pose

$$q' = \frac{uq}{1 + R_F} \quad 1 - q' = \frac{d(1-q)}{1 + R_F}.$$

<sup>128</sup> Voir l'annexe 2.

En introduisant la notation  $B(n ; p)$  pour une variable aléatoire binomiale de paramètres  $(n ; p)$ , on peut donc écrire

$$C_0 = S_0 \cdot \Pr[B(T ; q') \geq J] - K(1 + R_F)^{-T} \cdot \Pr[B(T ; q) \geq J].$$

La relation de parité « call-put »

$$C_0 + K(1 + R_F)^{-T} = P_0 + S_0,$$

permet d'obtenir immédiatement la formule d'évaluation du put ayant les mêmes caractéristiques :

$$P_0 = -S_0 \cdot \Pr[B(T ; q') < J] + K(1 + R_F)^{-T} \cdot \Pr[B(T ; q) < J].$$

### Remarque

Les paramètres  $u$  et  $d$  se déterminent par exemple à partir de la volatilité  $\sigma_R$  du return du sous-jacent. En effet, le return relatif à une période prenant les valeurs  $(u - 1)$  ou  $(d - 1)$  avec les probabilités respectives  $\alpha$  et  $(1 - \alpha)$ , on a

$$\begin{aligned} E_R &= \alpha(u - 1) + (1 - \alpha)(d - 1) \\ \sigma_R^2 &= \alpha(u - 1)^2 + (1 - \alpha)(d - 1)^2 - [\alpha(u - 1) + (1 - \alpha)(d - 1)]^2 \\ &= \alpha(1 - \alpha)(u - d)^2 \end{aligned}$$

En choisissant  $\alpha = 1/2$ , on obtient  $u - d = 2\sigma_R$ . Cox, Ross et Rubinstein suggèrent de prendre  $d = 1/u$ , ce qui conduit à une équation du second degré aisée à résoudre. ♣

### Illustration

Considérons une option d'achat de durée 7 mois, qui porte sur une action dont la valeur actuelle est 100 EUR et dont le prix d'exercice est de 110 EUR. On suppose que sa volatilité est de  $\sigma_R = 0,25$  calculée sur base annuelle et que le taux sans risque est de 4 % par an.

Nous allons évaluer la valeur de ce call en  $t = 0$  en construisant un arbre binomial avec le mois pour période de base. La volatilité et le taux sans risque équivalents sont donnés par

$$\sigma_R = \sqrt{\frac{1}{12}} \cdot 0,25 = 0,07219$$

$$R_F = \sqrt[12]{1,04} - 1 = 0,003274$$

On a par conséquent

$$u - \frac{1}{u} = 0,1443,$$

dont la seule racine positive est<sup>129</sup>  $u = 1,07477$  (et donc  $d = 0,93043$ ). La probabilité neutre vaut quant à elle

$$q = \frac{1,003274 - 0,93043}{1,07477 - 0,93043} = 0,5047$$

Présentons d'abord la manière pratique de travailler : construction des deux arbres binomiaux (progressive pour l'action et rétrogressive pour l'option).

On a par exemple, pour les deux valeurs de  $S_1$  :

$$\begin{cases} S_0 \cdot u = 100 \cdot 1,07477 = 107,477 \\ S_0 \cdot d = 100 \cdot 0,93043 = 93,043 \end{cases}$$

Le treillis binomial pour l'action sous-jacente est donné par le tableau 5.3.

0	1	2	3	4	5	6	7
100	107,477	115,513	124,150	133,432	143,409	154,132	165,656
	93,043	100,000	107,477	115,513	124,150	133,432	143,409
		86,570	93,043	100,000	107,477	115,513	124,150
			80,548	86,570	93,043	100,000	107,477
				74,944	80,548	86,570	93,043
					69,731	74,944	80,548
						64,880	69,731
							60,366

Tableau 5.3 – Treillis binomial pour l'action sous-jacente

L'arbre binomial pour l'option se construit à rebours. Ainsi, la dernière colonne se construit grâce à la relation

$$C_T = (S_T - K)^+,$$

dont le premier élément vaut  $\max(165,656 - 110 ; 0) = 55,656$  et les éléments des colonnes précédentes s'en déduisent, par exemple

$$\frac{1}{1,003274} [0,5047 \cdot 55,656 + 0,4953 \cdot 33,409] = 44,491.$$

On obtient ainsi le tableau 5.4.

<sup>129</sup> Si on avait choisi  $\alpha = 1/3$  au lieu de  $1/2$ , on aurait trouvé  $u = 1,0795$ , c'est-à-dire une différence relativement faible ; l'estimation de  $u$  ne dépend donc que peu de  $\alpha$ .

0	1	2	3	4	5	6	7
4,657	7,401	11,462	17,197	24,809	34,126	44,491	55,656
	1,891	3,312	5,696	9,555	15,482	23,791	33,409
		0,456	0,906	1,801	3,580	7,118	14,150
			0	0	0	0	0
				0	0	0	0
					0	0	0
						0	0
							0

Tableau 5.4 – Treillis binomial pour l'option

La valeur initiale du call est donc  $C_0 = 4,657$  EUR.

Présentons à présent le calcul de la valeur de l'option à partir de la formule finale. La probabilité auxiliaire est donnée par

$$q' = \frac{1,07477 \cdot 0,5047}{1,003274} = 0,5406.$$

Par ailleurs, comme

$$\frac{\ln 110 - \ln(100 \cdot d^7)}{\ln u - \ln d} = 4,1609,$$

on trouve  $J = 5$ . Ceci permet de calculer

$$\begin{aligned} \Pr[B(7; p) \geq 5] &= \binom{7}{5} p^5 (1-p)^2 + \binom{7}{6} p^6 (1-p) + \binom{7}{7} p^7 \\ &= p^5 (21 - 35p + 15p^2) \end{aligned}$$

et donc

$$\Pr[B(7; q) \geq 5] = 0,2343 \quad \Pr[B(7; q') \geq 5] = 0,2984.$$

Le prix du call est donc égal à

$$C_0 = 100 \cdot 0,2984 - 110 \cdot (1 + R_F)^{-7} \cdot 0,2343 = 4,657.$$

La prime du put ayant les mêmes caractéristiques vaut quant à elle

$$P_0 = -100 \cdot (1 - 0,2984) + 110 \cdot (1 + R_F)^{-7} \cdot (1 - 0,2343) = 12,168.$$

Notons qu'il est logique que le prix du put soit plus élevé que celui du call puisque l'option est actuellement out-of-the-money. ♣

### **Prise en compte des dividendes**

Nous avons supposé jusqu'ici que l'action sous-jacente ne versait pas de dividende. Examinons maintenant le cas où des dividendes sont versés.

Si un seul dividende est versé, durant la  $i$ -ème période (intervalle  $[i - 1 ; i]$ ), et qu'on note  $\delta$  le taux de dividende (rapport du montant du dividende à la valeur du titre), la valeur de l'action va diminuer à ce taux  $\delta$  lors du versement du dividende et l'arbre binomial pour le sous-jacent doit donc être modifié de la manière suivante :

- jusqu'à l'époque  $(i - 1)$ , aucune modification : les valeurs portées par les nœuds de l'arbre pour l'époque  $j \leq i - 1$  sont

$$S_0 u^k d^{j-k} \quad (k = 0, \dots, j) ;$$

- à partir de l'époque  $i$  (disons en  $j \geq i$ ), les valeurs deviennent<sup>130</sup>

$$S_0 (1 - \delta) u^k d^{j-k} \quad (k = 0, \dots, j) .$$

La construction rétrogressive de l'arbre pour l'option se fait classiquement à partir de là.

Si plusieurs dividendes sont versés, à des époques différentes, durant la durée du contrat d'option, la procédure décrite ci-dessus doit être appliquée lors de chaque versement.

### **5.3.2. Modèle de Black et Scholes pour les options sur actions**

Développons maintenant le célèbre modèle en temps continu élaboré par Black et Scholes<sup>131</sup>. L'option, conclue à l'instant  $0$  et échéant à l'instant  $T$ , sera ici évaluée à un instant quelconque  $t \in [0 ; T]$  et on note comme d'habitude  $\tau = T - t$ .

#### **Modèle de base**

Comme pour le modèle binomial, nous nous plaçons dans un marché parfait et nous envisageons une option européenne sur action ne distribuant aucun dividende pendant la durée du contrat d'option et dont la volatilité est constante pendant cette durée.

Nous supposons encore que le taux d'intérêt sans risque ne varie pas pendant cette période, qu'il est valable pour n'importe quelle échéance (courbe des taux plate et constante) et qu'il est le même pour un emprunt et pour un placement. Le taux d'intérêt annuel, noté  $R_F$  jusqu'à

<sup>130</sup> Ceci signifie que, lors d'une construction purement numérique de l'arbre, la prise en compte du facteur  $(1 - \delta)$  n'est effective que pour le passage de l'époque  $i - 1$  à l'époque  $i$ .

<sup>131</sup> BLACK F. et SCHOLES M., The pricing of options and corporate liabilities, *Journal of Political Economy*, Vol. 81, 1973, pp. 637-659.



présent, sera remplacé dans cette modélisation continue par le taux instantané correspondant  $r = \ln(1 + R_F)$ , de sorte qu'un montant unitaire placé pendant une durée de  $t$  années devient

$$(1 + R_F)^t = e^{rt}.$$

Rappelons (voir § 3.4.2) que l'évolution de la valeur de l'action sous-jacente est gouvernée par l'équation différentielle stochastique

$$\frac{dS_t}{S_t} = E_R \cdot dt + \sigma_R \cdot dw_t.$$

Nous allons dans un premier temps établir<sup>132</sup> la formule de Black & Scholes pour une option d'achat, dont la valeur est considérée comme fonction de la valeur  $S_t$  du sous-jacent et du temps  $t$ , les autres paramètres étant considérés comme constants :  $C_t = C(S_t, t)$ .

En appliquant la formule d'Itô à la fonction  $C(S_t, t)$ , on obtient

$$dC(S_t, t) = \left( C'_t + E_R S_t C'_S + \frac{\sigma_R^2}{2} S_t^2 C''_{SS} \right) \cdot dt + \sigma_R S_t C'_S \cdot dw_t.$$

Construisons maintenant un portefeuille constitué à l'instant  $t$  par

- l'achat de  $X$  actions sous-jacentes, dont la valeur est  $S_t$  ;
- la vente d'un call sur ce sous-jacent, de valeur  $C(S_t, t)$ .

La valeur  $V_t$  de ce portefeuille est donnée par

$$V_t = X \cdot S_t - C(S_t, t),$$

ce qui donne, par différentiation,

$$\begin{aligned} dV_t &= X \cdot [E_R S_t \cdot dt + \sigma_R S_t \cdot dw_t] - \left[ \left( C'_t + E_R S_t C'_S + \frac{\sigma_R^2}{2} S_t^2 C''_{SS} \right) \cdot dt + \sigma_R S_t C'_S \cdot dw_t \right] \\ &= \left[ X \cdot E_R S_t - \left( C'_t + E_R S_t C'_S + \frac{\sigma_R^2}{2} S_t^2 C''_{SS} \right) \right] \cdot dt + [X \cdot \sigma_R S_t - \sigma_R S_t C'_S] \cdot dw_t \end{aligned}$$

Nous choisissons ensuite  $X$  de manière à ce que ce portefeuille n'ait plus de composante aléatoire (le coefficient de  $dw_t$  dans la relation précédente doit être nul). L'hypothèse d'absence d'opportunité d'arbitrage montre que si cela est possible, le rendement de ce portefeuille doit être donné par le taux sans risque  $r$  :

$$\frac{dV_t}{V_t} = r \cdot dt + 0 \cdot dw_t.$$

<sup>132</sup> Nous ne développerons que la partie du raisonnement ayant un caractère financier, la fin de la démonstration étant exclusivement analytique. Le lecteur intéressé par le détail des calculs peut consulter la littérature originale ou DEVOLDER P., *Finance stochastique*, Editions de l'Université Libre de Bruxelles, 1993.

On obtient donc

$$\begin{cases} \frac{X \cdot E_R S_t - \left( C'_t + E_R S_t C'_S + \frac{\sigma_R^2}{2} S_t^2 C''_{SS} \right)}{X \cdot S_t - C(S_t, t)} = r \\ \frac{X \cdot \sigma_R S_t - \sigma_R S_t C'_S}{X \cdot S_t - C(S_t, t)} = 0 \end{cases}$$

ou, de manière équivalente,

$$\begin{cases} X \cdot (E_R - r) S_t - \left( C'_t + E_R S_t C'_S + \frac{\sigma_R^2}{2} S_t^2 C''_{SS} - rC(S_t, t) \right) = 0 \\ X - C'_S = 0 \end{cases}$$

La deuxième équation fournit la valeur de  $X$  qui annule la composante aléatoire du portefeuille :  $X = C'_S$ . En substituant dans la première équation, on trouve

$$(E_R - r) S_t \cdot C'_S - \left( C'_t + E_R S_t C'_S + \frac{\sigma_R^2}{2} S_t^2 C''_{SS} - rC(S_t, t) \right) = 0,$$

c'est-à-dire

$$C'_t + r S_t C'_S + \frac{\sigma_R^2}{2} S_t^2 C''_{SS} - rC(S_t, t) = 0,$$

équation dont le rendement moyen instantané  $E_R$  a disparu<sup>133</sup>.

Il s'agit là d'une équation (dont plus aucun élément n'est aléatoire) aux dérivées partielles du second ordre pour la fonction inconnue  $C(S_t, t)$ . Elle admet une solution unique si on impose deux conditions aux limites :

$$\begin{cases} C(0, t) = 0 \\ C(S_T, T) = (S_T - K)^+ \end{cases}$$

Cette équation se ramène, grâce à un changement de variables, à une équation bien connue des physiciens : l'équation de la chaleur<sup>134</sup>. Il est en effet facile (mais fastidieux) de voir que si on introduit la nouvelle fonction inconnue  $u(x, s) = C(S_t, t) \cdot e^{r\tau}$  où on effectue le changement de variables

<sup>133</sup> Tout comme on avait remarqué l'indépendance du résultat obtenu par le modèle binomial vis-à-vis de la probabilité  $\alpha$  gouvernant l'évolution du prix de l'action sous-jacente.

<sup>134</sup> Voir par exemple

KRASNOV M., KISSELEV A., MAKARENKO G. et CHIKINE E., *Mathématiques supérieures pour ingénieurs et polytechniciens*, De Boeck, 1993.

SOKOLNIKOFF I.S. et REDHEFFER R.M., *Mathematics of physics and modern engineering*, Mac Graw Hill, 1966.

$$\begin{cases} S_t = K \cdot \exp \left( \frac{\sigma_R^2 (x-s)}{2 \left( r - \frac{\sigma_R^2}{2} \right)} \right) \\ t = T - \frac{s \sigma_R^2}{2 \left( r - \frac{\sigma_R^2}{2} \right)^2} \end{cases}$$

qui s'inverse en

$$\begin{cases} x = \frac{2}{\sigma_R^2} \left( r - \frac{\sigma_R^2}{2} \right) \cdot \left[ \ln \frac{S_t}{K} + \left( r - \frac{\sigma_R^2}{2} \right) \tau \right] \\ s = \frac{2}{\sigma_R^2} \left( r - \frac{\sigma_R^2}{2} \right)^2 \tau \end{cases}$$

l'équation obtenue se transforme en

$$u''_{xx} = u'_s,$$

avec les conditions aux limites

$$\begin{cases} \lim_{x \rightarrow -\infty} u(x, s) = 0 \\ u(x, 0) = v(x) = \begin{cases} K \cdot \left[ \exp \left( \frac{x \frac{\sigma_R^2}{2}}{r - \frac{\sigma_R^2}{2}} \right) - 1 \right] & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \end{cases}$$

Cette équation de la chaleur a comme solution

$$u(x, s) = \frac{1}{2\sqrt{\pi s}} \int_{-\infty}^{+\infty} v(y) e^{-(x-y)^2 / 4s} dy.$$

En effectuant les calculs avec l'expression particulière de  $v(y)$  puis en réalisant le changement de variables inverse, on obtient alors la formule de Black & Scholes pour l'option d'achat :

$$C(S_t, t) = S_t \Phi(d_1) - K e^{-r\tau} \Phi(d_2),$$

où on a posé

$$\left. \begin{matrix} d_1 \\ d_2 \end{matrix} \right\} = \frac{\ln \frac{S_t}{K} + \left( r \pm \frac{\sigma_R^2}{2} \right) \tau}{\sigma_R \sqrt{\tau}}$$

et où la fonction  $\Phi$  représente la fonction de répartition de la loi normale réduite :

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx.$$

L'évaluation du prix  $P_t$  d'une option de vente s'obtient à partir de celle de l'option d'achat grâce à la relation de parité « call-put »

$$C_t + K \cdot e^{-r\tau} = P_t + S_t.$$

En effet,

$$\begin{aligned} P(S_t, t) &= C(S_t, t) + Ke^{-r\tau} - S_t \\ &= S_t \Phi(d_1) - Ke^{-r\tau} \Phi(d_2) + Ke^{-r\tau} - S_t \\ &= -S_t [1 - \Phi(d_1)] + Ke^{-r\tau} [1 - \Phi(d_2)] \end{aligned}$$

et par conséquent

$$P(S_t, t) = -S_t \Phi(-d_1) + Ke^{-r\tau} \Phi(-d_2),$$

puisque

$$1 - \Phi(t) = \Phi(-t)$$

en raison de la parité de la densité normale réduite.

### ***Illustration***

Considérons une option ayant les mêmes caractéristiques qu'au § 5.3.1 :  $S_0 = 100$ ,  $K = 110$ ,  $t = 0$ ,  $T = 7$  mois,  $\sigma_R = 0,25$  sur base annuelle et  $R_F = 4$  % par an.

Nous travaillerons ici avec l'année comme base du temps, de sorte que

$$\tau = \frac{7}{12} \quad r = \ln 1,04 = 0,03922.$$

$$\left. \begin{matrix} d_1 \\ d_2 \end{matrix} \right\} = \frac{\ln \frac{100}{110} + \left( 0,03922 \pm \frac{0,25^2}{2} \right) \cdot \frac{7}{12}}{0,25 \cdot \sqrt{\frac{7}{12}}} = \begin{cases} -0,2839 \\ -0,4748 \end{cases}$$

d'où  $\Phi(d_1) = 0,3823$  et  $\Phi(d_2) = 0,3175$ . Ceci permet de calculer le prix du call :

$$C = C(S_0, 0) = 100 \cdot \Phi(d_1) - 110 \cdot e^{-0,03922 \cdot \frac{7}{12}} \cdot \Phi(d_2) = 4,695.$$

La prime du put ayant les mêmes caractéristiques vaut quant à elle

$$P = P(S_0, 0) = -100 \cdot [1 - \Phi(d_1)] + 110 \cdot e^{-0,03922 \cdot \frac{7}{12}} \cdot [1 - \Phi(d_2)] = 12,207.$$

On peut remarquer la similitude de ces chiffres avec les valeurs obtenues à l'aide du modèle binomial (4,657 et 12,168 respectivement).♣

### *Paramètres de sensibilité*

Lorsque l'évaluation du prix d'une option s'effectue à l'aide de la formule de Black et Sholes, les paramètres de sensibilité (les « grecques ») prennent une forme pratique.

Examinons tout d'abord le cas du delta d'une option d'achat. Si on désigne par  $\phi$  la densité normale réduite

$$\phi(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

on obtient par dérivation,

$$\Delta(C) = C'_S = \Phi(d_1) + \frac{1}{S_t \sigma_R \sqrt{\tau}} [S_t \phi(d_1) - K e^{-r\tau} \phi(d_2)].$$

Il est facile de voir que la quantité entre crochets est nulle et que donc

$$\Delta(C) = \Phi(d_1)$$

et que, par un raisonnement en tout point similaire, on a pour un put :

$$\Delta(P) = \Phi(d_1) - 1.$$

La formule ci-dessus fournit un moyen très simple de déterminer le nombre d'actions à détenir pour un émetteur de call pour couvrir son risque (delta hedging). C'est d'ailleurs une utilisation courante de la relation de Black et Scholes : le prix d'une option est déterminé par la loi de l'offre et de la demande et son « inversion » fournit la volatilité implicite ; on utilise alors cette dernière dans la relation  $\Delta(C) = \Phi(d_1)$ , qui porte alors le nom de *formule de couverture*.

Les autres paramètres de sensibilité (gamma, thêta, véga, rho) s'obtiennent de manière similaire :

$$\Gamma(C) = \Gamma(P) = \frac{\phi(d_1)}{S_t \sigma_R \sqrt{\tau}}$$

$$\begin{cases} \Theta(C) = -\frac{S_t \sigma_R \phi(d_1)}{2\sqrt{\tau}} - rKe^{-r\tau} \Phi(d_2) \\ \Theta(P) = -\frac{S_t \sigma_R \phi(d_1)}{2\sqrt{\tau}} + rKe^{-r\tau} \Phi(-d_2) \end{cases}$$

$$V(C) = V(P) = \tau S_t \phi(d_1)$$

$$\begin{cases} \rho(C) = \tau Ke^{-r\tau} \Phi(d_2) \\ \rho(P) = -\tau Ke^{-r\tau} \Phi(-d_2) \end{cases}$$

Citons pour terminer une relation liant les paramètres delta, gamma et thêta. L'équation aux dérivées partielles obtenue lors de la démonstration de la formule de Black et Scholes pour un call est en réalité valable également pour un put (nous noterons donc  $p$  le prix, sans préciser) :

$$p'_t + rS_t p'_s + \frac{\sigma_R^2}{2} S_t^2 p''_{ss} - rp(S_t, t) = 0,$$

ce qui donne, en utilisant les paramètres de sensibilité :

$$\Theta + rS_t \Delta + \frac{\sigma_R^2}{2} S_t^2 \Gamma = r \cdot p.$$

### ***Prise en compte des dividendes***

Si un dividende de taux continu<sup>135</sup>  $\delta$  est versé entre  $t$  et  $T$  et que l'action sous-jacente vaut  $S_t$  (resp.  $S_T$ ) à l'instant  $t$  (resp.  $T$ ), on peut dire que si elle n'avait pas versé de dividende, elle serait passée de la valeur  $S_t$  à la valeur  $e^{\delta\tau} S_T$  ou encore qu'une même action sans dividende passerait de la valeur  $e^{-\delta\tau} S_t$  à l'instant  $t$  à la valeur  $S_T$  à l'instant  $T$ . Pour tenir compte du dividende, il suffit donc, dans la formule de Black et Scholes, de remplacer  $S_t$  par  $e^{-\delta\tau} S_t$ , ce qui donne

$$C(S_t, t) = S_t e^{-\delta\tau} \Phi(d_1) - Ke^{-r\tau} \Phi(d_2)$$

$$P(S_t, t) = -S_t e^{-\delta\tau} \Phi(-d_1) + Ke^{-r\tau} \Phi(-d_2)$$

<sup>135</sup> On utilisera, comme dans tout ce paragraphe, un facteur d'actualisation/capitalisation de type exponentiel.

où on a posé

$$\left. \begin{matrix} d_1 \\ d_2 \end{matrix} \right\} = \frac{\ln \frac{S_t}{K} + \left( r - \delta \pm \frac{\sigma_R^2}{2} \right) \tau}{\sigma_R \sqrt{\tau}}.$$

### 5.3.3. Autres modèles d'évaluation

#### *Options sur obligations*

L'application des méthodes vues ci-dessus (arbre binomial ou formule de Black et Scholes) à des options sur obligations n'est pas adéquate. En effet,

- il faut tenir compte des coupons régulièrement versés ;
- la constance de la volatilité du sous-jacent – hypothèse valable pour les actions – n'est pas vraie dans le cas d'une obligation puisque sa valeur converge, à maturité, vers la valeur de remboursement  $R$ .

L'adaptation du modèle binomial à cette situation est possible, mais n'est pas une généralisation évidente de la méthode exposée ci-dessus<sup>136</sup>.

L'adaptation du modèle de Black et Scholes consistera à remplacer le mouvement brownien géométrique qui représentait l'évolution de la valeur de l'action par un processus stochastique qui gouverne l'évolution des taux d'intérêt, comme ceux qui sont à la base des modèles de Vasicek ou de Cox, Ingersoll et Ross (voir § 4.5). Malheureusement, l'équation aux dérivées partielles qu'on en déduit n'admet généralement pas de solution analytique et des techniques numériques doivent alors être utilisées<sup>137</sup>.

#### *Options de change*

Pour une option de change, deux taux sans risque doivent être pris en considération : un relatif à la devise domestique et un relatif à la devise étrangère.

Pour le modèle discret, notons respectivement  $R_F^{(D)}$  et  $R_F^{(F)}$  ces deux taux. En refaisant le raisonnement du § 5.3.1 avec cette généralisation, il est possible de déterminer le prix d'une option de change grâce à la technique de l'arbre binomial. On voit que le principe exposé précédemment reste valable à une légère adaptation des formules près :  $C_0$  est la valeur actualisée de la valeur moyenne du call à l'échéance (pour une période)

$$C_0 = (1 + R_F^{(D)})^{-1} [q \cdot C(u) + (1 - q) \cdot C(d)],$$

<sup>136</sup> On consultera par exemple COPELAND T.E. et WESTON J.F., *Financial theory and corporate finance*, Addison-Wesley, 1988.

<sup>137</sup> Voir par exemple COURTADON G., The pricing of options on default-free bonds, *Journal of Financial and Quantitative Analysis*, Vol. 17, 1982, pp. 75-100.

par rapport à la probabilité risque neutre

$$q = \frac{(1 + (R_F^{(D)} - R_F^{(F)})) - d}{u - d} \quad 1 - q = \frac{u - (1 + (R_F^{(D)} - R_F^{(F)}))}{u - d}.$$

En ce qui concerne le modèle continu, notons les taux d'intérêt dans les devises domestique et étrangère respectivement  $r^{(D)}$  et  $r^{(F)}$ . Par un raisonnement similaire à celui tenu pour les options sur action versant des dividendes, on voit que la formule de Black et Scholes est encore valable à condition de remplacer le prix du sous jacent  $S_t$  par  $S_t e^{-r^{(F)}\tau}$ , ce qui donne les formules

$$C(S_t, t) = S_t e^{-r^{(F)}\tau} \Phi(d_1) - K e^{-r^{(D)}\tau} \Phi(d_2)$$

$$P(S_t, t) = -S_t e^{-r^{(F)}\tau} \Phi(-d_1) + K e^{-r^{(D)}\tau} \Phi(-d_2)$$

où on a posé

$$\left. \begin{matrix} d_1 \\ d_2 \end{matrix} \right\} = \frac{\ln \frac{S_t}{K} + \left( (r^{(D)} - r^{(F)}) \pm \frac{\sigma_R^2}{2} \right) \tau}{\sigma_R \sqrt{\tau}}.$$

Il s'agit là de la formule de Garman-Kohlhagen<sup>138</sup>.

<sup>138</sup> GARMAN M. et KOHLHAGEN S., Foreign currency option values, *Journal of International Money and Finance*, n° 2, 1983, pp. 231-237.