

Chapitre 2

Marches aléatoires

De la marche au hasard de l'ivrogne “modèle” qui oublie à chaque instant d'où il vient et où il veut aller, au mouvement brownien d'une particule colloïdale en suspension dans un fluide moléculaire, en passant par l'évolution d'actifs financiers, une marche aléatoire est décrite par un processus stochastique en termes de probabilité. La modélisation et la compréhension d'un tel processus s'inscrivent pleinement dans le cadre de la physique mais sont également à la base de la modélisation financière depuis les travaux précurseurs de Louis Bachelier en 1900, cinq ans avant l'étude du mouvement brownien par Albert Einstein.

En effet, quel que soit l'actif financier (action, obligations, matière première...) on observe la même forme d'évolution erratique dans le temps. Cette régularité, qui résulte pourtant des actions innombrables d'individus dont les motivations propres ne sont pas forcément rationnelles, suggère une possible modélisation. La question est alors : comment tirer de l'information de la “trajectoire” passée.¹ L'hypothèse d'efficience des marchés, qui implique que toute l'information sur un actif est reflétée par sa valeur à l'instant t , est assez bien vérifiée dans la réalité (corrélations temporelles faibles sur des temps courts). Mais comme en physique, la modélisation financière repose sur une idéalisation de la réalité...

En partant d'un modèle simple sur réseau, comme le jeu de pile ou face, nous présenterons des résultats non triviaux et contraires à l'intuition qui nous permettront, à la limite continue, d'aborder les phénomènes de diffusion en termes d'équations stochastiques. Enfin, nous introduirons le modèle de Black-Scholes, pierre angulaire de la modélisation financière. Mais pour commencer, intéressons-nous à un modèle phénoménologique et historique : l'équation de Langevin.

¹On suppose ici que les mécanismes sous-jacents ne changent pas. En fait, en finance, les paramètres comme la volatilité évoluent. La stabilité est donc faible et les prédictions ne sont valables que pour un avenir proche.

2.1 Mouvement Brownien et équation de Langevin

Le mouvement brownien décrit la dynamique macroscopique et aléatoire d'une "grosse"² particule en suspension dans un fluide. Les collisions incessantes des molécules du fluide avec la particule mettent cette dernière en mouvement. L'origine microscopique du mouvement faisant intervenir un nombre gigantesque de molécules on ne peut traiter la dynamique de la particule qu'en termes de probabilité.

C'est le botaniste Robert Brown qui en 1827 a étudié systématiquement le mouvement qui porte son nom en observant au microscope des particules de pollen.³ Ce mouvement imprévisible et erratique est représenté sur la figure 2.1 dans le cas d'une bille microscopique en suspension dans de l'eau.

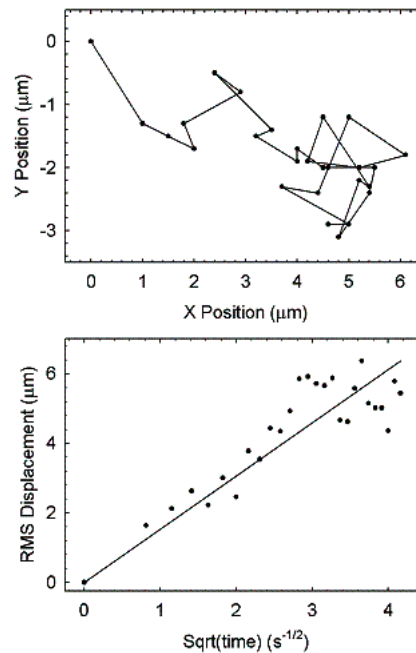


FIG. 2.1 – En haut, trajectoire observée par microscopie d'une bille de polystyrène de diamètre $d = 0,5\mu\text{m}$ en suspension dans de l'eau. En bas, déplacement moyen $\sqrt{\langle x^2 \rangle}$ en fonction de $\sqrt{t}$. Dans V. Protasenko et al., *The chemical Educator*, Vol 10, No 4 (2005).

²L'adjectif "grosse" est à relativiser : typiquement de l'ordre du micron.

³Certains observateurs ont pensé que les particules en mouvement étaient des organismes vivants. L'utilisation de matériaux minéraux a ensuite écarté cette hypothèse. Les observations indiquent que le mouvement brownien ne s'arrête jamais et qu'il augmente lorsque : i) la taille des particules diminue, ii) la viscosité du fluide diminue et iii) la température augmente.

En 1905, Albert Einstein propose une théorie du mouvement brownien basée sur une description hydrodynamique des fluides et sur la théorie cinétique de la chaleur. Il obtient ainsi une équation de la diffusion donnant la densité de probabilité de trouver une particule brownienne en un point donné à un instant donné. Par ailleurs, Marian von Smoluchowski développe en 1906 un modèle de collisions aléatoires qui repose sur une description discrète du mouvement.

Équation de Langevin

En 1908, Paul Langevin⁴ propose un modèle phénoménologique dans le cadre de la mécanique, et donc sur une base déterministe, mais en ajoutant un élément probabiliste : une grosse particule de masse m est soumise à deux forces exercées par l'ensemble des molécules du fluide dans lequel elle est plongée :

- un frottement fluide qui tend à immobiliser la particule : $\vec{f}(t) = -\alpha\vec{v}(t)$, où α est un coefficient positif lié à la viscosité du fluide.
- une force aléatoire $\eta(t)$, appelée bruit blanc gaussien, représentant les chocs aléatoires des molécules du fluide sur la particule dont l'action est de maintenir en mouvement la particule. Elle change donc brutalement sur des temps très courts, typiques des temps de collisions avec les molécules du fluide. Et puisqu'il n'y a pas de direction privilégiée, $\langle\eta(t)\rangle = 0$.

Il y a donc deux échelles de temps dans le modèle.⁵ La force aléatoire $\eta(t)$ varie très rapidement alors que la vitesse de la particule est une variable lente.

En appliquant la relation fondamentale de la dynamique à ce système (à une dimension) on a

$$m \frac{dv(t)}{dt} = -\alpha v(t) + \eta(t). \quad (2.1)$$

L'équation de Langevin est une équation différentielle *stochastique* (c'est même la première qui ait été proposée), c'est-à-dire qu'elle dépend de variables aléatoires. A nouveau, puisqu'il n'y a pas de direction privilégiée, $\langle x \rangle = 0$. Bien sûr, la particule se déplace, et une façon d'évaluer sa position par rapport à l'origine est de calculer son *déplacement carré moyen* $\langle x^2 \rangle$.

Calcul de $\langle x^2 \rangle$

Puisque $v = \dot{x}$, l'équation (2.1) s'écrit en la multipliant par x :

$$mx\ddot{x} = m\left(\frac{d(x\dot{x})}{dt} - \dot{x}^2\right) = -\alpha x\dot{x} + x\eta$$

Prenons la moyenne (moyenne d'ensemble, c'est à dire sur un grand nombre de particules identiques) des deux membres de cette équation. La particule étant en équilibre thermique avec le fluide, son énergie cinétique moyenne est connue : $\frac{1}{2}m\langle\dot{x}^2\rangle = \frac{1}{2}kT$, où T est la température du fluide et k la constante de Boltzmann (théorème d'équipartition de l'énergie). Par ailleurs, en négligeant les faibles

⁴Paul Langevin (1872 - 1946) physicien français. Il est l'auteur de travaux sur le magnétisme, la relativité et les ultrasons.

⁵Ces deux échelles de temps sont bien sûr associées aux deux longueurs caractéristiques du problème : la particule et les molécules du fluide.

corrélations entre la position x et la force aléatoire η , on a $\langle x\eta \rangle = \langle x \rangle \langle \eta \rangle = 0$ et finalement :

$$m \frac{d\langle x\dot{x} \rangle}{dt} = kT - \alpha \langle x\dot{x} \rangle.$$

Curieusement, la force aléatoire n'intervient plus explicitement, mais elle est à l'origine du terme kT , c'est à dire d'une énergie cinétique moyenne non nulle. On obtient donc :

$$\langle x\dot{x} \rangle = \frac{kT}{\alpha} + Ce^{-\gamma t},$$

où $\gamma^{-1} = \frac{m}{\alpha}$ est le temps de relaxation de la vitesse de la particule. Le coefficient de frottement α peut être évalué à partir de la viscosité du fluide et on trouve typiquement $\gamma^{-1} \simeq 10^{-8}s$. Si à $t = 0$, $x = 0$, alors

$$\langle x\dot{x} \rangle = \frac{1}{2} \frac{d\langle x^2 \rangle}{dt} = \frac{kT}{\alpha} (1 - e^{-\gamma t})$$

Et finalement,

$$\langle x^2 \rangle = \frac{2kT}{\alpha} [t - \gamma^{-1}(1 - e^{-\gamma t})].$$

Selon l'échelle de temps d'observation, on note deux comportements :

$$\langle x^2 \rangle \simeq \frac{kT}{m} t^2 \quad \text{si } t \ll \gamma^{-1}, \quad (2.2)$$

$$\langle x^2 \rangle \simeq \frac{2kT}{\alpha} t \quad \text{si } t \gg \gamma^{-1}. \quad (2.3)$$

L'équation (2.2) décrit le *régime balistique*, en $\langle x^2 \rangle \sim t^2$, où sur des temps courts la particule se déplace à la vitesse constante $\sim \sqrt{kT/m}$. En revanche, sur les temps longs, l'équation (2.3) met en évidence le *comportement diffusif*, en $\langle x^2 \rangle \sim t$, illustré par la figure 2.1.

Contrairement au mouvement balistique (mouvement d'un projectile) où la distance parcourue croît proportionnellement au temps, le mouvement brownien est caractérisé par une distance qui croît comme la racine carrée du temps. Asymptotiquement, la particule s'éloigne indéfiniment de l'origine, avec une vitesse qui tend vers 0 et avec une probabilité égale dans toutes les directions.

2.2 Marche aléatoire et retour à l'origine

2.2.1 Premiers résultats

Une particule se déplace sur un axe (marche aléatoire à une dimension) en partant de l'origine. Sa dynamique aléatoire est la suivante : à chaque itération n (temps discrétisé), elle avance ou recule avec la même probabilité⁶ d'une distance

⁶Cette marche aléatoire est donc symétrique, mais on pourrait imaginer une dérive dans une certaine direction.

donnée (une unité). Sa position S_n à l'instant n ne dépend donc que de sa position S_{n-1} à l'instant $n-1$:

$$S_n = S_{n-1} + X_n, \quad n \geq 1, \quad (2.4)$$

où X_n est une variable aléatoire valant ± 1 avec la même probabilité $1/2$ et $\langle X_i X_j \rangle = \delta_{ij}$. La position de la particule à l'instant n est donc

$$S_n = \sum_{i=1}^n X_i.$$

La variable aléatoire S_n peut également être interprétée comme le gain (positif ou négatif) d'un joueur à la $n^{\text{ème}}$ partie d'un jeu équitable (pile ou face) : pile, il gagne 1€, face il perd 1€.

On ne peut bien sûr pas déterminer la trajectoire de la particule de façon certaine, mais on peut espérer calculer des valeurs moyennes (moyennes sur des répétitions d'un grand nombre de trajectoires). Dans la suite, nous décrirons la trajectoire de la particule dans le plan $\{n, s_n\}$ donnant la position s_n de la particule à l'instant n (voir figure 2.2).

Exercice 12 : Distance parcourue [†]

Montrer que la valeur moyenne $\langle S_n \rangle = 0$. Expliquer pourquoi, néanmoins, la particule s'éloignera (en moyenne) de son point de départ. À quelle vitesse ? Peut-on généraliser à deux et trois dimensions ?

Indication : La distance à l'origine n'est pas S_n mais $|S_n| = \sqrt{S_n^2}$.

Sur un chemin donné de n pas, appelons p et q le nombre de pas respectivement vers l'avant et vers l'arrière.⁷ Clairement on a

$$\begin{cases} n = p + q \\ s_n = p - q. \end{cases}$$

Soit N_{n,s_n} le nombre de chemins issus de l'origine allant au point **B** de coordonnées (n, s_n) , c'est-à-dire le nombre de façons d'arriver à la position s_n à

[†]**Réponse :**

Une somme de variables aléatoires de moyenne nulle est encore une variable aléatoire de moyenne nulle, donc $\langle S_n \rangle = 0$. La distance à l'origine n'est pas S_n , mais $|S_n|$, soit $\sqrt{S_n^2}$. Or :

$$S_n^2 = \sum_i X_i^2 + \sum_{i \neq j} X_i X_j.$$

Comme $X_i^2 = 1$ et $\langle X_i X_j \rangle = \langle X_i \rangle \langle X_j \rangle = 0$ (tirages indépendants), il vient $\langle S_n^2 \rangle = n$. La distance à l'origine croît donc comme $\sqrt{n}$.

La généralisation à plusieurs dimensions est évidente : on écrit vectoriellement

$$\vec{S}_n = \vec{X}_1 + \cdots + \vec{X}_n,$$

on calcule le produit scalaire et comme $|\vec{X}_i| = 1$ et que toutes les orientations de $\vec{X}$ sont équiprobables (ceci est vrai aussi bien sur un réseau carré (4 orientations), cubique (6 orientations) ou même dans le cas continu ($\langle \cos \theta \rangle = 0$)), on a encore une valeur moyenne des produits scalaires nulle ($\langle \vec{X}_i \cdot \vec{X}_j \rangle = \langle \cos \theta \rangle$) et donc $\langle |\vec{O}M_n|^2 \rangle = n$.

⁷Attention p et q ne sont pas des probabilités, mais des nombres entiers donnés.

l'instant n :

$$N_{n,s_n} = C_n^p = C_n^q = C_n^{\frac{n+s_n}{2}}. \quad (2.5)$$

Le nombre total de chemins de longueur n est manifestement 2^n . La probabilité p_{n,s_n} que la particule ait la position s_n à l'instant n (le point **B**) est donc

$$p_{n,s_n} = \frac{1}{2^n} N_{n,s_n} = \frac{1}{2^n} C_n^{\frac{n+s_n}{2}}. \quad (2.6)$$

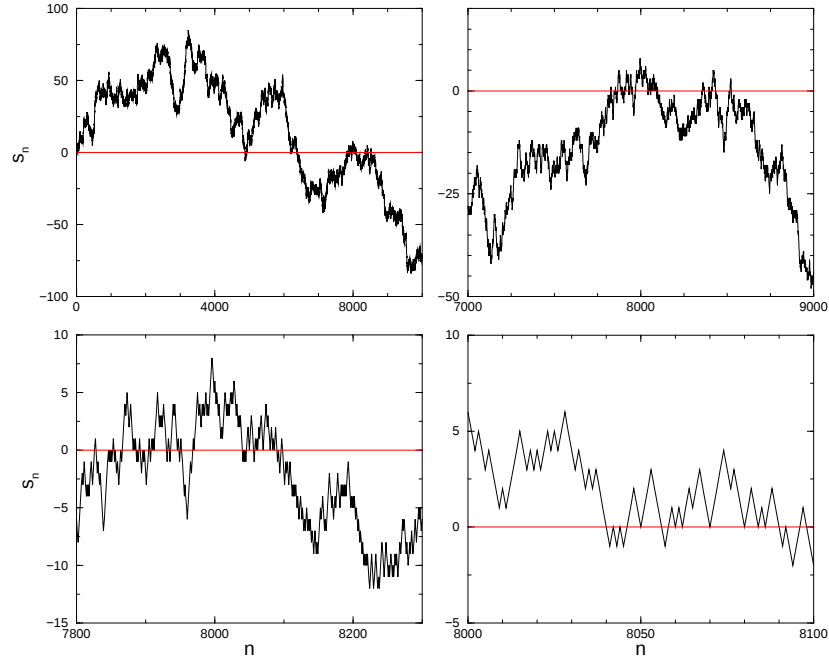


FIG. 2.2 – Marche aléatoire à 1d. Zooms successifs sur différentes portions de la trajectoire (voir les échelles).

Retour à l'origine

Le retour à l'origine à l'instant n correspond à l'événement $s_n = 0$. Dans le cas du jeu de pile ou face, cet événement est la ruine du joueur. Sachant que la particule part de $s_0 = 0$ à l'instant $n = 0$, la probabilité de retour à l'origine ($s_n = 0$) à un instant nécessairement pair, $2n$, est, pour $n \geq 1$

$$u_{2n} = \frac{1}{2^{2n}} C_{2n}^n \quad \text{et} \quad u_0 = 1. \quad (2.7)$$

Remarquons que l'historique de la trajectoire entre les instants 0 et $2n - 1$ n'intervient pas dans cette probabilité. On a évidemment $u_2 = 1/2$. Puis, $u_4 =$

0.375, $u_6 = 0.3125$, $u_{60} \simeq 0.1$... En utilisant la formule de Stirling (voir annexe A.3) on a pour n grand :⁸

$$u_{2n} \simeq \frac{1}{\sqrt{\pi n}}. \quad (2.8)$$

Notons que si chaque position possible entre $-2n$ et $2n$ était équiprobable, la probabilité d'être en 0 à l'instant $2n$ varierait en $1/n$ (elle serait donc plus petite que dans le cas de la marche aléatoire).

2.2.2 Principe de réflexion

Ce principe sera très utile dans la suite : Soit $\mathbf{A}(a, \alpha)$ et $\mathbf{B}(b, \beta)$, avec $b > a \geq 0$, α et β strictement positifs ; le nombre de chemins de $\mathbf{A}$ à $\mathbf{B}$ qui touchent ou traversent l'axe des temps est égal au nombre de chemins qui vont de $\mathbf{A}'(a, -\alpha)$ à $\mathbf{B}$, où $\mathbf{A}'$ est le symétrique de $\mathbf{A}$ par rapport à l'axe des temps (voir figure 2.3).

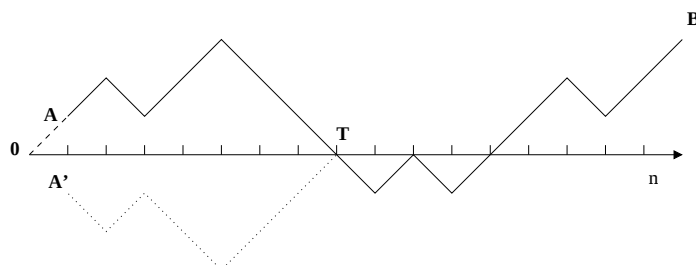


FIG. 2.3 – Marche aléatoire et sa symétrique (en pointillés). Pour le chemin $\mathbf{AB}$, $X_i = +1$ pour $i = 0, 1, 3, 4, 9, 11, 12, 13, 15, 16$ et -1 sinon.

Démonstration : Soit $\mathbf{T}$ le premier point atteint sur l'axe des temps. Par symétrie, il y a bijection entre les chemins de $\mathbf{A}$ à $\mathbf{T}$ et ceux de $\mathbf{A}'$ à $\mathbf{T}$.

Voyons une application du principe de réflexion.

Théorème du vote

Supposons que lors d'une élection, p électeurs votent pour le candidat P et q pour son adversaire Q , avec $p \geq q$. La probabilité pour que durant le vote (ou le dépouillement) le candidat P ait *toujours* plus de voix que Q est

$$Pr = \frac{p - q}{p + q}.$$

Démonstration : En terme de marche aléatoire, cela revient à calculer le nombre de chemins, N_{n, s_n}^+ , entre l'origine et le point $\mathbf{B}(n = p + q, s_n = p - q)$, qui ne touchent pas l'axe des temps (c'est-à-dire tels que $s_i > 0$, pour $i = 1, \dots, n$).

⁸On notera que la formule de Stirling est une très bonne approximation même pour n petit. En effet, l'équation (2.8) donne $u_2 \simeq 0.56$, $u_4 \simeq 0.4$ et $u_6 = 0.33$.

Tous ces chemins passent par le point $\mathbf{A}(1, 1)$ que l'on peut prendre comme nouvelle origine. Le nombre total de chemins, N_t , entre $\mathbf{A}(1, 1)$ et $\mathbf{B}(n, s_n)$ est donc le même qu'entre l'origine et le point $(n-1, s_n-1)$, soit $N_t = N_{n-1, s_n-1}$.

Par ailleurs, $N_t = N_{n, s_n}^+ + N^-$, où N^- est le nombre de chemins entre $\mathbf{A}$ et $\mathbf{B}$ qui touchent ou traversent l'axe des temps. D'après le principe de réflexion, N^- est le nombre de chemins entre $(1, -1)$ et $\mathbf{B}$. En changeant d'origine, N^- est donc le nombre de chemins entre l'origine et le point $(n-1, s_n+1)$, soit $N^- = N_{n-1, s_n+1}$. On en déduit

$$N_{n, s_n}^+ = N_t - N^- = N_{n-1, s_n-1} - N_{n-1, s_n+1},$$

et d'après la formule (2.5),

$$N_{n, s_n}^+ = C_{n-1}^{\frac{n+s_n}{2}-1} - C_{n-1}^{\frac{n+s_n}{2}} = C_{n-1}^{p-1} - C_{n-1}^p.$$

Donc

$$N_{n, s_n}^+ = N_{n, s_n} \frac{p-q}{n} = N_{n, s_n} \frac{p-q}{p+q}.$$

En divisant N_{n, s_n}^+ par le nombre total de chemins N_{n, s_n} , on trouve la probabilité Pr qu'un chemin allant de l'origine au point $\mathbf{B}$ ne touche pas l'axe des temps (la particule ne revient pas à l'origine pendant un temps n).

On voit que dans le cas d'une marche aléatoire on peut écrire $Pr = s_n/n$. Si la marche est symétrique (en moyenne $p = q$) et donc $Pr = 0$: la particule revient certainement à l'origine, si on attend suffisamment longtemps.

2.2.3 Premier retour à l'origine

A la section 2.2.1, nous avons calculé la probabilité u_{2n} que la particule repasse par l'origine au temps $2n$. Soit f_{2n} la probabilité que la particule revienne pour la *première fois* à l'origine à l'instant $2n$. On a

$$u_{2n} = f_2 u_{2n-2} + f_4 u_{2n-4} + \cdots + f_{2n} u_0. \quad (2.9)$$

En effet l'évènement "la particule repasse à l'origine au temps $2n$ " est la réunion de n évènements disjoints : soit elle repasse à l'origine pour la *première fois* au temps 2 et y retourne au temps $2n$ après une durée $2n-2$ (avec éventuellement d'autres passages en 0 entre temps), soit elle repasse à l'origine pour la *première fois* au temps 4, et y retourne au temps $2n$ après une durée $2n-4$ (avec éventuellement d'autres passages en 0 intermédiaires), etc... Les probabilités respectives de ces évènements sont $f_2 u_{2n-2}$, $f_4 u_{2n-4}$, ... On notera que $f_2 = u_2$ et on prendra $f_0 = 0$.

Théorème du retour à l'origine

La probabilité qu'il n'y ait aucun retour à l'origine jusqu'à l'instant $2n$ (compris) est égale à la probabilité de retour à l'instant $2n$. Soit

$$P(S_1 \neq 0, S_2 \neq 0, \dots, S_{2n} \neq 0) = P(S_{2n} = 0) = u_{2n}. \quad (2.10)$$

Les S_i étant tous positifs ou tous négatifs (avec la même probabilité), on a

$$P(S_1 > 0, S_2 > 0, \dots, S_{2n} > 0) = \frac{1}{2}P(S_{2n} = 0) = \frac{1}{2}u_{2n}. \quad (2.11)$$

Par exemple, contrairement à l'intuition, la probabilité que partant de 0, un joueur à pile ou face n'ait pas perdu au bout de 100 lancers (pas de retour à l'origine) est $\frac{1}{2}u_{100} \simeq 0.08$!

Démonstration : Démontrons la relation (2.11). On a

$$P(S_1 > 0, S_2 > 0, \dots, S_{2n} > 0) = \sum_{r=1}^{\infty} P(S_1 > 0, S_2 > 0, \dots, S_{2n} = 2r).$$

(Les termes de la somme avec $r > n$ sont nuls). Le théorème du vote permet d'écrire

$$P(S_1 > 0, S_2 > 0, \dots, S_{2n-1} > 0, S_{2n} = 2r) = \frac{1}{2^{2n}}(N_{2n-1, 2r-1} - N_{2n-1, 2r+1}).$$

Quand on somme sur r , il reste⁹

$$\frac{1}{2^{2n}}N_{2n-1, 1} = \frac{1}{2^{2n}}C_{2n-1}^n = \frac{1}{2}u_{2n}.$$

Calcul de f_{2n}

Raisonnons entre les instants $2n-2$ et $2n$: d'après le théorème du retour à l'origine, la particule n'est pas retournée à l'origine jusqu'à l'instant $2n-2$ compris avec la probabilité u_{2n-2} . Au temps $2n$, soit elle revient à l'origine (c'est donc pour la première fois) avec la probabilité f_{2n} , soit elle n'y revient pas avec la probabilité u_{2n} . Autrement dit,

$$u_{2n-2} = f_{2n} + u_{2n}, \quad n \geq 1. \quad (2.12)$$

En explicitant les u_{2n} , on obtient la probabilité du premier retour¹⁰ au temps $2n$:

$$f_{2n} = \frac{1}{2^{2n}} \frac{1}{2n-1} C_{2n-1}^n = \frac{1}{2n-1} u_{2n}.$$

Puisque u_{2n} varie en $1/\sqrt{n}$, f_{2n} décroît plus rapidement, en $1/n^{\frac{3}{2}}$. Ainsi, $f_2 = 0.5$, $f_4 = 0.125$, $f_6 = 0.0625$...

⁹En particulier, le terme $N_{2n-1, 2n+1} = 0$, car on ne peut pas atteindre la position $2n+1$ en $2n-1$ itérations.

¹⁰En sommant de 0 à n l'équation (2.12) on obtient $(\sum_{k=1}^n f_{2k}) + u_{2n} = 1$. A chaque instant n , soit la particule est déjà retournée au moins une fois à l'origine, soit elle n'y est jamais revenue.

2.2.4 Loi de l'arcsinus

Au jeu de pile ou face, l'intuition suggère que le gain d'un joueur (partant de 0) oscille fréquemment autour de 0. Ainsi, si on interrompt une longue partie, la probabilité que le dernier passage à l'origine soit ancien devrait être faible. Comme nous allons le voir, ce n'est pas le cas. En particulier, nous verrons que la probabilité qu'il n'y ait pas eu de retour à l'origine durant la seconde moitié du jeu est égale à 50%, et ce quelle que soit sa durée!

Dernier passage à l'origine

La probabilité $\alpha_{2k,2n}$ que jusqu'au temps $2n$ (inclus) le dernier passage à l'origine soit arrivé à l'instant $2k \leq 2n$ est

$$\alpha_{2k,2n} = u_{2k}u_{2n-2k}, \quad \text{pour } k = 0, 1 \dots n.$$

Démonstration : L'événement "jusqu'au temps $2n$ (inclus) le dernier passage à l'origine arrive au temps $2k$ " est l'intersection des événements indépendants " $S_{2k} = 0$ " et " $S_{2k+1} \neq 0 \dots S_{2n} \neq 0$ ". Le premier a, par définition, la probabilité u_{2k} et le second u_{2n-2k} . En effet, il n'y a qu'à prendre le point $(2k, 0)$ comme nouvelle origine et appliquer le théorème du retour à l'origine.

La probabilité $\alpha_{2k,2n}$, appelée distribution¹¹ (discrète) de *l'arcsinus* d'ordre n , est symétrique par rapport à n : $\alpha_{2k,2n} = \alpha_{2n-2k,2n}$. Ainsi, en sommant de $2k = n$ à $2k = 2n$, on trouve que la probabilité qu'il n'y ait pas eu de retour à l'origine durant la seconde moitié du jeu est indépendante de n et vaut 50%. À l'aide de l'estimation donnée par la formule (2.8), on montre qu'asymptotiquement (pour k pas trop proche de 0 et $2n$),

$$\alpha_{2k,2n} = \frac{1}{2^{2n}} C_{2k}^k C_{2n-2k}^{n-k} \simeq \frac{1}{\pi} \frac{1}{\sqrt{k(n-k)}}.$$

On voit que la probabilité $\alpha_{2k,2n}$ est la plus grande pour k proche de 0 ou n et qu'elle est minimale en $k = n/2$. Ce résultat est très surprenant : contrairement à ce que l'on pouvait attendre, la particule repasse rarement par l'origine. Elle se déplacera le plus clair de son temps du côté positif (ou négatif). De la même façon, un joueur à pile ou face restera gagnant (ou perdant) pendant longtemps.

Si on pose

$$f(t) = \frac{1}{\pi \sqrt{t(1-t)}} \quad (2.13)$$

On voit que

$$\alpha_{2k,2n} \simeq \frac{1}{n} f(t_k), \quad \text{avec } t_k = k/n$$

À la limite $n \rightarrow \infty$, t_k devient une variable continue, $0 < t < 1$. La figure 2.4 montre une comparaison entre la distribution $\alpha_{2k,2n}$ et la fonction (2.13).

¹¹En effet, on a bien $\sum_{k=0}^n \alpha_{2k,2n} = 1$, le dernier retour à l'origine ayant eu lieu soit en $k = 0$, soit en $k = 1, \dots$ soit en $k = n$.

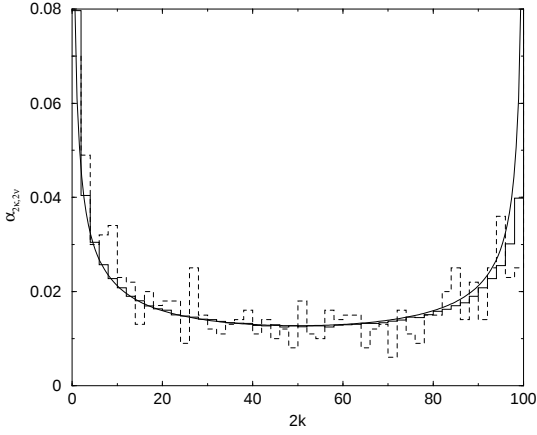


FIG. 2.4 – Probabilité $\alpha_{2k,2n}$: fréquence obtenue pour 10^6 tirages (histogramme en trait plein) et pour 10^3 tirages (histogramme en tireté). La courbe en trait plein est la fonction $f(t)$ donnée par l'équation (2.13).

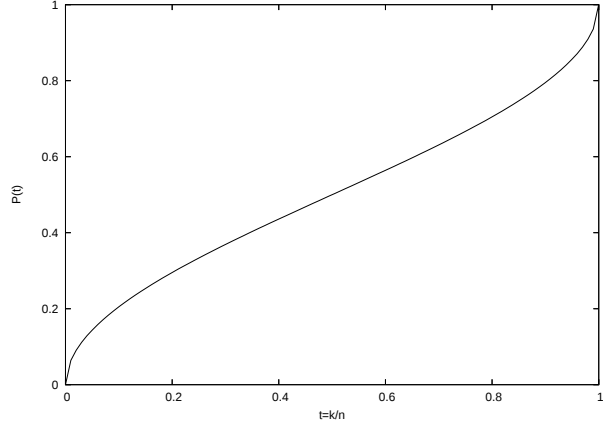


FIG. 2.5 – Loi de l'arcsinus donnant la probabilité que le dernier passage à l'origine ait eu lieu avant l'instant $t = k/n$ (voir l'équation (2.14)).

En intégrant $f(t)$, on obtient la probabilité $P(t)$ pour que le *dernier* passage à l'origine ait lieu *avant* le temps (relatif) $t = k/n$, pour $n \rightarrow \infty$:

$$P(t) = \sum_{\frac{k}{n} < t} \alpha_{2k,2n} \simeq \frac{2}{\pi} \arcsin \sqrt{t}. \quad (2.14)$$

La loi de l'arcsinus $P(t)$ est représentée sur la figure 2.5. Imaginons un jeu de pile ou face commencé, à raison d'un coup par seconde, il y a un an (disons le 1 janvier). La dernière égalisation (passage à l'origine) se ferait avant le 9 janvier dans 10% des cas. Autrement dit, depuis cette date, le même joueur serait toujours gagnant.

2.3 Problème classique de ruine

On considère deux joueurs. Le capital total des deux joueurs est fixé à a et z est le capital initial du premier joueur. A chaque essai ce joueur a une probabilité p de gagner 1€ et q de le perdre (et réciproquement pour son adversaire). Le jeu consiste en une série d'essais qui se termine quand le capital du joueur est devenu 0 (ruine) ou a (victoire). Nous allons calculer la probabilité de ruine et le temps moyen de jeu.

En terme de marche aléatoire, on reconnaît une *marche avec barrières absorbantes* en $x = 0$ et $x = a$. Si la marche est symétrique ($p = q$) et le capital infini ($a \rightarrow \infty$), on retrouve le problème déjà traité du premier passage à l'origine.

Soit q_z la probabilité de ruine du premier joueur de capital initial z (et ce sur un nombre fini, mais “grand” d'essais) et p_z sa probabilité de victoire. On

verra que¹² $p_z + q_z = 1$.

2.3.1 Calcul de q_z

Après le premier essai, le nouveau capital du joueur est soit $z + 1$ avec la probabilité p , soit $z - 1$ avec la probabilité q , on a donc en général

$$q_z = pq_{z+1} + qq_{z-1}, \quad \text{pour } 1 \leq z \leq a-1, \quad (2.15)$$

en imposant les conditions aux limites évidentes¹³ :

$$q_0 = 1 \quad \text{et} \quad q_a = 0. \quad (2.16)$$

Dans un premier temps, supposons $p \neq q$. Manifestement, l'équation (2.15) admet deux solutions particulières $q_z = 1$ et $q_z = (\frac{q}{p})^z$. La solution générale est alors

$$q_z = A + B\left(\frac{q}{p}\right)^z.$$

Les conditions (2.16) déterminent A et B . Soit¹⁴

$$q_z = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^z}{\left(\frac{q}{p}\right)^a - 1} \quad \text{et} \quad p \neq q. \quad (2.17)$$

On voit facilement¹⁵ que si $p = q = \frac{1}{2}$,

$$q_z = 1 - \frac{z}{a} \quad \text{et} \quad p = q. \quad (2.18)$$

La probabilité de victoire p_z du premier joueur est la probabilité de ruine de son adversaire. Il suffit donc de remplacer p , q et z par q , p et $a - z$ dans les formules (2.17) et (2.18) :

$$p_z = \frac{\left(\frac{q}{p}\right)^z - 1}{\left(\frac{q}{p}\right)^a - 1}, \quad \text{si } p \neq q, \quad (2.19)$$

$$p_z = \frac{z}{a}, \quad \text{si } p = q. \quad (2.20)$$

On vérifie facilement que $p_z + q_z = 1$. La probabilité p_z est représentée sur la figure 2.6. Comme on le voit, il suffit que p soit un peu supérieure à q (par exemple $p = 1.1q$) pour nettement avantager le joueur.

¹²Ce résultat à démontrer n'est pas évident, il montre que le jeu a un temps fini.

¹³En effet, pour $z = 1$, le premier essai implique $q_1 = pq_2 + q$. De même, pour $z = a - 1$, le premier essai donne $q_{a-1} = qq_{a-2}$. Les conditions aux limites (2.16) permettent d'utiliser la relation (2.15) pour $z = 1$ et $z = a - 1$. La relation (2.15) est également valable pour la probabilité de victoire p_z .

¹⁴Il est évident que cette solution est unique comme on le voit en partant de $z = 0$ et $z = 1$ et en utilisant, de manière récursive l'expression (2.15) de q_z .

¹⁵Puisque $q_z = A + Bz$ est une solution de (2.15).

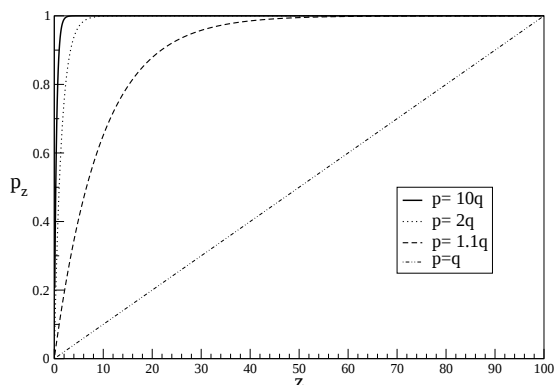


FIG. 2.6 – Probabilité de victoire p_z en fonction du capital initial z (équations (2.19) et (2.20)), pour un capital total $a = 100$ et différentes valeurs du rapport p/q .

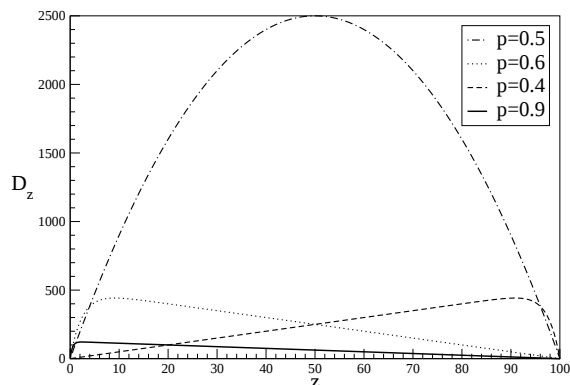


FIG. 2.7 – Durée moyenne de jeu D_z en fonction du capital initial z , pour un capital total $a = 100$ et différentes valeurs de p .

Exercice 13 : Gain moyen[†]

Calculer le gain moyen attendu par le joueur à la fin de la partie.

Dans le cas d'un adversaire infiniment riche ($a \rightarrow \infty$), on a selon les cas :

$$\begin{aligned} q_z &= 1, \quad \forall z \text{ si } p \leq q, \\ q_z &= \left(\frac{q}{p}\right)^z, \quad \text{si } p > q. \end{aligned}$$

Dans le cas $p \leq q$, le joueur est donc certain de perdre quel que soit son capital initial (fini).

En terme de marche aléatoire (à 1d), la probabilité qu'une particule partant de l'origine atteigne la position z est égale à 1 si $p \geq q$ et à $(\frac{p}{q})^z$ sinon (pour changer d'origine, p et q ont été permutées).

2.3.2 Calcul de la durée moyenne de jeu

Soit D_z la durée moyenne de jeu lorsque le premier joueur à un capital initial z . Comme précédemment, au premier essai le capital du joueur est soit $z + 1$ avec la probabilité p , soit $z - 1$ avec la probabilité q , mais il faut dans les deux cas compter un temps de plus, soit $D_{z+1} + 1$ et $D_{z-1} + 1$. On a donc

$$D_z = pD_{z+1} + qD_{z-1} + 1, \quad \text{pour } 1 \leq z \leq a - 1, \quad (2.21)$$

avec les conditions aux limites $D_0 = D_a = 0$.

[†]**Réponse** : En fin de partie, le joueur a gagné $a - z$ avec la probabilité p_z ou perdu z avec la probabilité q_z . Soit $\langle G \rangle = ap_z - z$.

Si $p \neq q$, la recherche d'une solution particulière linéaire en z donne $D_z = z/(q-p)$. Par ailleurs, la différence entre deux solutions de l'équation (2.21) vérifie l'équation (2.15), dont on sait que les solutions sont de la forme $A+B(\frac{q}{p})^z$. La durée moyenne D_z s'écrit donc

$$D_z = \frac{z}{q-p} + A + B\left(\frac{q}{p}\right)^z.$$

Et les conditions aux limites imposent

$$D_z = \frac{z}{q-p} - \frac{a}{q-p} \frac{1 - \left(\frac{q}{p}\right)^z}{1 - \left(\frac{q}{p}\right)^a}. \quad (2.22)$$

Si $p = q$, On montre que $-z^2$ est une solution particulière de l'équation (2.21), donc $D_z = -z^2 + A + Bz$. Et finalement, en tenant compte des conditions aux limites,¹⁶

$$D_z = z(a-z).$$

C'est une durée beaucoup plus longue que celle à laquelle on s'attendrait. Par exemple si les deux joueurs ont chacun 10€, et mise 1€ par partie, en moyenne il faudra faire 100 parties pour que l'un des joueurs soit ruiné. Il est instructif de comparer ce résultat à la marche aléatoire représentée sur la figure 2.2. Comme on le voit sur figure 2.7, le temps le plus long est naturellement obtenu lorsque les deux joueurs ont le même capital et que $p = q$. En dehors de ces conditions, la durée de la partie décroît rapidement.

Dans le cas d'un adversaire infiniment riche ($a \rightarrow \infty$), on a selon les cas :

$$\begin{aligned} D_z &= \frac{z}{q-p}, & \text{si } p < q \\ D_z &\rightarrow \infty, & \text{si } p \geq q. \end{aligned}$$

Là encore le résultat est contre-intuitif.

2.4 Diffusion

2.4.1 Passage à la limite continue

A la section 2.2, nous avons utilisé la formule de Stirling pour passer à la limite continue. De même, on peut calculer la limite lorsque $n \rightarrow \infty$ de la probabilité p_{n,s_n} que la particule soit à la position s_n (proche de l'origine) à l'instant n (voir la formule (2.6)) :

$$p_{n,s_n} = \frac{1}{2^n} C_n^{\frac{n+s_n}{2}} \rightarrow 2 \frac{1}{\sqrt{2\pi n}} e^{-\frac{s_n^2}{2n}} \quad (2.23)$$

quand $n \rightarrow \infty$. Le facteur 2 qui apparaît au second membre de la formule (2.23) tient au fait qu'elle n'est vraie qu'un cas sur deux : quand la parité de n est égale à celle de s_n , sinon la probabilité est nulle.

¹⁶On pourrait également poser $q = p + \epsilon$ et faire un développement limité au second ordre en ϵ dans l'équation (2.22).

Application du théorème de la limite centrale

On retrouve bien sûr ce résultat immédiatement avec le théorème de la limite centrale, puisque S_n est défini comme la somme de n variables aléatoires indépendantes X_k valant ± 1 avec la même probabilité. Clairement $\langle x \rangle = 0$ et $\sigma = 1$. On en déduit à la limite des n grands que

$$p_{n,s_n} \simeq \frac{1}{\sqrt{2\pi n}} e^{-\frac{s_n^2}{2n}}.$$

Exercice 14 : Distance parcourue $II^\dagger$

Montrer qu'à la limite des n grands, $\langle |s_n| \rangle = \sqrt{\frac{2}{\pi}} \sqrt{n} \neq \sqrt{\langle s_n^2 \rangle} = \sqrt{n}$.

Jusqu'à présent, nous avons considéré une marche aléatoire symétrique. Si on suppose maintenant que les probabilités de se déplacer de 1 et -1 sont respectivement p et $q = 1 - p$, on obtient¹⁷

$$p_{n,s_n} \simeq \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{4npq}} e^{-\frac{(s_n - n(p-q))^2}{8npq}}.$$

Limite continue

La particule se déplace sur un réseau (position discrète), nous allons maintenant effectuer le passage à la limite continue. Posons $x = hs_n$ et $t = n\tau$, où h est le pas en distance et τ le pas en temps. On va les faire tendre vers 0. Définissons une vitesse moyenne

$$c = p \frac{h}{\tau} - q \frac{h}{\tau} = \frac{(p-q)h}{\tau}. \quad (2.24)$$

On peut écrire

$$\frac{(s_n - n(p-q))^2}{8npq} = \frac{(x - ct)^2}{2h^2 \frac{t}{\tau} (1 - \frac{\tau^2 c^2}{h^2})}.$$

Pour que cette expression ait un sens, il faut que le rapport $\frac{h^2}{2\tau} = D$ ait une limite finie¹⁸ quand $h \rightarrow 0$ et $\tau \rightarrow 0$. On trouve alors

$$p_{n,s_n} \simeq f(x,t) dx \simeq \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2Dt}} e^{-\frac{(x-ct)^2}{4Dt}} dx, \quad (2.25)$$

où le pas h a été assimilé à dx . La constante D s'appelle le *coefficient de diffusion*. Par exemple, pour une molécule de dioxygène dans l'air (à l'ambiante),

[†]**Réponse** : On utilisera les intégrales données dans l'annexe A.2.

¹⁷En effet, X est alors une variable aléatoire de moyenne $\langle X \rangle = p - q$ et de variance $Var(X) = (p+q) - (p-q)^2 = 4pq$. Pour le jeu de pile ou face, cela revient à jouer avec une pièce imparfaite ou truquée.

¹⁸La définition (2.24) de la vitesse moyenne implique que $p - q \sim h$, si on veut qu'elle reste finie.

$D = 1.8 \cdot 10^{-5} \text{ m}^2\text{s}^{-1}$. La distance moyenne parcourue par la molécule en 1s est donc d'environ $\sqrt{D} \simeq 5 \text{ mm}$ (à comparer à sa vitesse de quelques centaines de mètres par seconde en régime balistique).

La probabilité que la position de la particule soit entre x et $x + dx$ au temps t est donc donnée par $f(x, t)dx$. On vérifie que $\lim_{t \rightarrow 0} f(x) = \delta(x)$, comme il se doit (à $t = 0$, la particule est sûrement en $x = 0$). La distribution $f(x, t)$ est représentée sur la figure 2.8. Comme on le voit, le processus de diffusion se manifeste par un élargissement de la courbe $f(x, t)$ en fonction du temps : la variance croît en $2Dt$. Le déplacement du pic correspond à la dérive à la vitesse c .

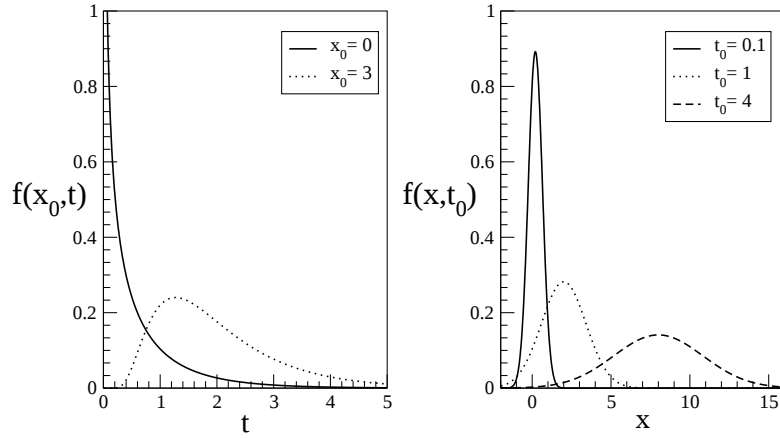


FIG. 2.8 – Distribution $f(x, t)$ (formule 2.25) avec $D = 1$. et $c = 2$. pour différentes positions x_0 (à gauche) et différents temps t_0 (à droite).

Exercice 15 : Distance parcourue III[†]

Montrer que $\langle x \rangle = ct$ (dérive à la vitesse c) et $\langle x^2 \rangle = c^2 t^2 + 2Dt$. On retrouve donc que la distance à l'origine croît en $\sqrt{t}$ (et non pas en t) lorsque la vitesse de dérive c est nulle. En comparant ce résultat à l'équation (2.3), on voit que le coefficient de diffusion s'écrit $D = kT/\alpha$ (relation d'Einstein).

[†]Réponse : En effet,

$$\langle X \rangle = \int_{-\infty}^{\infty} x \frac{1}{2\sqrt{\pi Dt}} e^{-\frac{(x-ct)^2}{4Dt}} dx = ct$$

et

$$\langle X^2 \rangle = \int_{-\infty}^{\infty} x^2 \frac{1}{2\sqrt{\pi Dt}} e^{-\frac{(x-ct)^2}{4Dt}} dx = c^2 t^2 + 2Dt.$$

où l'on a effectué le changement de variable $y = (x - ct)/\sqrt{4Dt}$ et on utilise la valeur de $I(p) = \int_{-\infty}^{\infty} x^p e^{-x^2} dx$: $I(0) = \sqrt{\pi}$, $I(1) = 0$ et $I(2) = \sqrt{\pi}/2$.

2.4.2 Équation de la diffusion

Historiquement, l'équation de la diffusion a été démontrée par Adolph Fick¹⁹ en 1855 sur une base phénoménologique. Partant de l'hypothèse que le flux de matière est proportionnel au gradient de concentration (première loi de Fick) et en exprimant la conservation du nombre de particules on trouve l'équation de la diffusion (deuxième loi de Fick) donnant l'évolution de la concentration d'un soluté dans une solution. En 1905, Albert Einstein obtient la même loi en raisonnant sur la probabilité $P(x, t)$ de trouver une particule brownienne en x à l'instant t . Nous allons obtenir le même résultat en utilisant une marche aléatoire asymétrique ($p \neq q$) de pas $\pm h$, puis en passant à la limite continue.

L'événement "la particule se trouve en x à l'instant $t + \tau$ ", associé à la probabilité $P(x, t + \tau)$, résulte de deux événements disjoints : elle se trouvait en $x - h$ ou en $x + h$ à l'instant t . On a donc

$$P(x, t + \tau) = pP(x - h, t) + qP(x + h, t).$$

où p est la probabilité de faire un pas $+h$ et $q = 1 - p$ celle de faire un pas $-h$. Soit

$$P(x, t + \tau) - P(x, t) = p(P(x - h, t) - P(x, t)) + q(P(x + h, t) - P(x, t)).$$

Si on se limite au premier ordre en τ dans le membre de gauche et au second ordre en h dans celui de droite, il reste

$$\tau \frac{\partial P(x, t)}{\partial t} = p \left(-h \frac{\partial P(x, t)}{\partial x} + \frac{h^2}{2} \frac{\partial^2 P(x, t)}{\partial x^2} \right) + q \left(h \frac{\partial P(x, t)}{\partial x} + \frac{h^2}{2} \frac{\partial^2 P(x, t)}{\partial x^2} \right).$$

Donc

$$\frac{\partial P(x, t)}{\partial t} = \frac{h}{\tau} \frac{\partial P(x, t)}{\partial x} (q - p) + \frac{h^2}{2\tau} \frac{\partial^2 P(x, t)}{\partial x^2}.$$

Et avec les notations précédentes

$$\frac{\partial P(x, t)}{\partial t} = -c \frac{\partial P(x, t)}{\partial x} + D \frac{\partial^2 P(x, t)}{\partial x^2}. \quad (2.26)$$

C'est l'équation de la diffusion (on retrouve la seconde loi de Fick si $p = q$ et $c = 0$). Il s'agit d'une forme simple de l'équation de Fokker-Planck.

On peut vérifier que l'équation (2.25) est la solution de cette équation différentielle telle que $P(0, 0) = 1$ (à $t = 0$, la particule est en $x = 0$). On peut aussi le démontrer directement en utilisant la fonction caractéristique

$$\phi(u, t) = \int e^{iux} P(x, t) dx.$$

L'équation différentielle (2.26) devient

$$\frac{\partial \phi}{\partial t} = icu \phi(u, t) - Du^2 \phi(u, t),$$

¹⁹Adolph Fick (1829-1901) physicien allemand.

qui s'intègre immédiatement en

$$\phi(u, t) = \phi(u, 0) e^{(icu - Du^2)t}.$$

En prenant la transformé de Fourier inverse on a donc

$$P(x, t) = \frac{1}{2\pi} \int \phi(u, 0) e^{(icu - Du^2)t} e^{-iux} du.$$

En supposant $P(x, 0) = \delta(x)$, alors²⁰ $\phi(u, 0) = 1$ et on retrouve l'équation (2.25) :

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x-ct)^2}{4Dt}}.$$

Généralisation pour $d > 1$

On peut facilement généraliser l'équation de la diffusion pour des dimensions de l'espace $d > 1$ en remplaçant la dérivée seconde d'espace par le Laplacien. On trouve dans le cas symétrique ($p = q$ et donc $c = 0$)

$$P(\mathbf{r}, t) = \frac{1}{(4\pi Dt)^{\frac{d}{2}}} e^{-\frac{r^2}{4Dt}},$$

où $\mathbf{r}$ est le vecteur position de la particule. La probabilité $Pr(r)\mathbf{dr}$ de trouver la particule dans le domaine $[\mathbf{r}, \mathbf{r} + d\mathbf{r}]$ quel que soit l'instant t est $d\mathbf{r} \int P(\mathbf{r}, t) dt$. L'intégration sur t , (en posant $\frac{r^2}{4Dt} = u$) donne

$$Pr(r)\mathbf{dr} = \mathbf{dr} \frac{1}{\pi^{\frac{d}{2}} r^{d-2} 4D} \Gamma\left(\frac{d}{2} - 1\right),$$

qui n'est finie que pour $d \geq 3$. En particulier, pour $d = 3$, $Pr(r)\mathbf{dr} = \mathbf{dr} \frac{1}{\pi r 4D}$. Ainsi, à une dimension, on a vu qu'il est certain que la particule repasse par l'origine (ou que le joueur perde) pourvu qu'on attende suffisamment longtemps. Cette propriété reste vraie à $2d$, mais n'est plus vérifiée pour $d \geq 3$: par exemple, la probabilité de retour à l'origine pour $d = 3$ est $p \simeq 0.3405 < 1$.

2.4.3 Processus de Wiener

Plutôt que d'effectuer le passage à la limite continue d'une marche aléatoire discrète, on peut s'intéresser directement à un processus stochastique dont la variation, à chaque intervalle de temps $dt \rightarrow 0$, est une grandeur continue distribuée selon une certaine loi de probabilité. En particulier, une variable aléatoire X suit *un processus de Wiener* si :

- pendant un intervalle de temps $dt \rightarrow 0$, sa variation dX est distribuée selon une loi gaussienne de paramètres $m = 0$ et $\sigma = \sqrt{dt}$.
- les variations de X sont indépendantes sur deux intervalles de temps consécutifs (propriété de Markov, voir section 2.5).

²⁰En effet, $\delta(x) = \frac{1}{2\pi} \int e^{-iux} du$.

En décomposant le temps t en N intervalles de durée dt , la variation de X sur un temps $t = Ndt$ sera donc donnée par :

$$\Delta X = X(t) - X(0) = \sum_{i=1}^N dX_i,$$

où dX_i est la variation de X au i^{eme} intervalle de temps. D'après les deux propriétés du processus de Wiener la variation ΔX suit une loi gaussienne de paramètres $m = 0$ et $\sigma = \sqrt{t}$ (somme de variables aléatoires gaussiennes indépendantes). Remarquons que si la variance de dX ne variait pas en dt la variation ΔX tendrait soit vers 0 soit vers ∞ .²¹ Comme il se doit, l'incertitude sur X , mesurée par sa variance, augmente avec le temps.

Processus de Wiener généralisé

Une variable aléatoire S suit un *processus de Wiener généralisé* si sa variation dS pendant un intervalle de temps dt s'écrit :

$$\boxed{dS = a dt + b dX}, \quad (2.27)$$

où a et $b > 0$ sont des constantes et X une variable aléatoire suivant un processus de Wiener. Le premier terme du second membre de l'équation (2.27) décrit une dérive ("drift" en anglais) *déterministe* et constante de paramètre a , l'espérance de variation du processus par unité de temps. Le second terme est une variable *aléatoire* distribuée sur une gaussienne de variance $b^2 dt$. Ainsi, la variation ΔS de S sur un temps $t = Ndt$ est distribuée sur une loi normale de paramètres $m = at$ et $\sigma = b\sqrt{t}$. La figure 2.9 montre clairement la dérive en at et les fluctuations autour de cette tendance moyenne contrôlée par le paramètre b .

Plus généralement, les coefficients a et b peuvent dépendre de la variable S et du temps t (on suppose alors qu'ils restent constants sur un intervalle de temps dt). On parle dans ce cas de *processus d'Itô*.

Lemme d'Itô

Le lemme d'Itô²² (1951) est aux fonctions de variables aléatoires ce que le théorème de Taylor est aux fonctions de variables déterministes. Nous adoptons ici une approche heuristique pour le démontrer.

Soit S une variable aléatoire qui suit un processus d'Itô dont la variation est donnée par :

$$dS = a(S, t)dt + b(S, t)dX. \quad (2.28)$$

Dans la suite on omettra les arguments de a et b .

²¹Supposons que la variance de la distribution normale de dX soit $(dt)^\alpha$, la variation ΔX sur un temps $t = Ndt$ serait alors distribuée sur une gaussienne de variance $Ndt^\alpha = tdt^{\alpha-1}$. Le comportement de ΔX à la limite continue ($dt \rightarrow 0$) impose $\alpha = 1$. En effet si $\alpha > 1$ alors $\Delta X \rightarrow 0$ et si $\alpha < 1$, $\Delta X \rightarrow \infty$ quel que soit t .

²²Kiyosi Itô (1915-), mathématicien japonais.

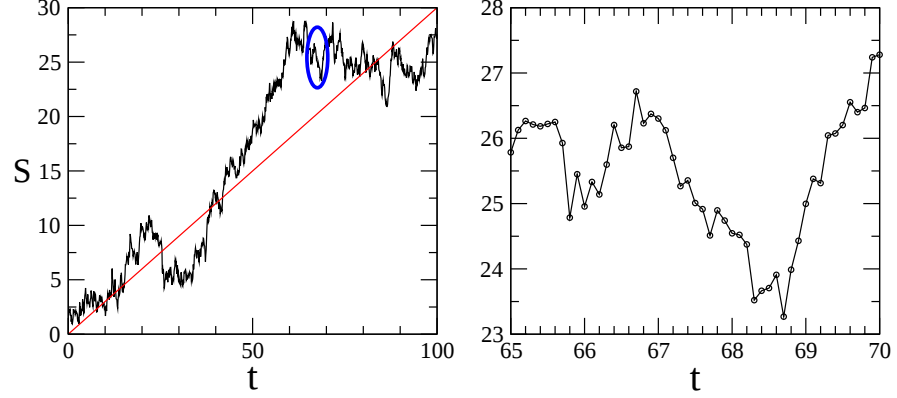


FIG. 2.9 – Processus de Wiener de paramètres $a = 0.3$ et $b = 1.5$ avec $dt = 0.1$. Sur la figure de gauche est tracée la droite d'équation $S = 0.3t$ donnant l'évolution moyenne. La figure de droite représente un zoom de la figure de gauche.

Soit $f(S, t)$ une fonction continue de S et de t . Son développement en série de Taylor s'écrit :

$$df = \frac{\partial f}{\partial S} dS + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} dS^2 + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} dt^2 + \frac{\partial^2 f}{\partial t \partial S} dt dS + \dots \quad (2.29)$$

D'après l'équation (2.28), la variation dS dépend de dt explicitement, mais également implicitement à travers dX . Par définition du processus de Wiener, puisque $dX^2 \sim dt$ (la variance de dX), $dS^2 = (adt + b dX)^2$ contient un terme d'ordre dt qui domine les deux autres en dt^2 et $dt^{\frac{3}{2}}$ à la limite $dt \rightarrow 0$. Ainsi,

$$\lim_{dt \rightarrow 0} dS^2 = b^2 dt. \quad (2.30)$$

Curieusement, dS^2 perd son caractère stochastique pour devenir une variable déterministe à la limite continue. En remplaçant dS et dS^2 par leurs expressions (2.28) et (2.30) en fonction de dt dans l'équation (2.29), on obtient :

$$df = \frac{\partial f}{\partial S} (adt + b dX) + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (b^2 dt) + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} dt^2 + \frac{\partial^2 f}{\partial t \partial S} dt (adt + b dX) + \dots$$

On ne conservant que les termes dominants d'ordre $dX \sim \sqrt{dt}$ et dt , on a le lemme d'Itô :

$$\boxed{df = \left(\frac{\partial f}{\partial S} a + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} b^2 \right) dt + \frac{\partial f}{\partial S} b dX.} \quad (2.31)$$

Ainsi, comme S , la fonction $f(S, t)$ suit un processus d'Itô, c'est-à-dire :

$$df = a'(S, t) dt + b'(S, t) dX,$$

où les coefficients a' et b' sont donnés par l'équation (2.31).

Les deux variables aléatoires S et f sont corrélées et dépendent de la même variable aléatoire dX . En éliminant dX , il est donc possible de construire une troisième variable G , en fonction de S et f , dont la variation dG soit complètement déterministe pendant la durée dt . Comme nous le verrons, cette remarque est fondamentale dans le modèle de Black-Scholes.

2.4.4 Marche aléatoire gaussienne et vols de Lévy

On peut bien sûr généraliser la marche aléatoire gaussienne étudiée dans la section précédente à des marches aléatoires dont le déplacement à chaque itération est donné par une certaine distribution de probabilité. Par exemple, la figure 2.10 montre deux marches aléatoires obtenues par simulation avec la loi normale et la loi de Cauchy représentées sur la figure 1.9. On voit clairement l'effet de cette loi large qui autorise des déplacements de grandes tailles par rapport à ceux de la loi gaussienne. Ces sauts de grandes amplitudes sont appelés des vols de Lévy.²³

Comme on peut le voir sur la figure 2.11, la différence entre ces deux lois de probabilité est encore plus spectaculaire dans le cas d'une marche aléatoire plane (dans un espace de dimension $d = 2$). On pourra également comparer la marche aléatoire gaussienne simulée de la figure 2.11 avec la trajectoire observée expérimentalement de la figure 2.1.

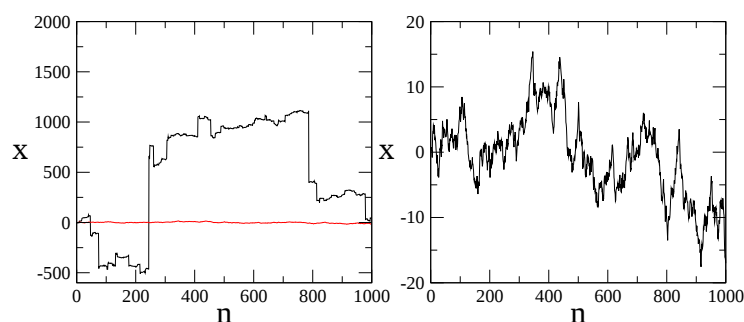


FIG. 2.10 – Marche aléatoire à 1d de 1000 pas distribués sur la loi de Cauchy (à gauche) et sur la loi normale (à droite) représentées sur la figure 1.9 (remarquer la différence d'échelles!). Pour comparaison, la trajectoire gaussienne de la figure de droite est représentée sur la figure de gauche.

²³Paul Lévy (1886-1971), mathématicien français.

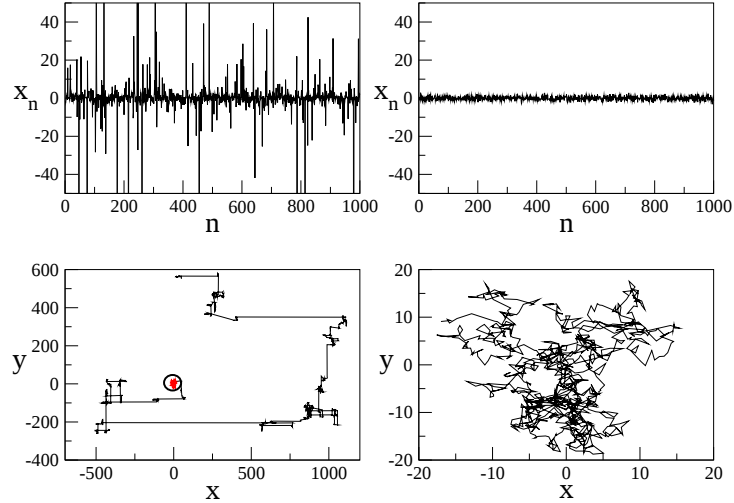


FIG. 2.11 – Marche aléatoire à 2d de 1000 pas distribués sur la loi de Cauchy (à gauche) et sur la loi normale (à droite) représentées sur la figure 1.9. Les figures du haut montrent la variable aléatoire dans les deux cas et celles du bas une trajectoire dans le plan $\{x, y\}$ (remarquer la différence d'échelles!). La trajectoire gaussienne de la figure de droite est représentée à l'intérieur du cercle, centré à l'origine, de la figure de gauche.

2.5 Chaîne de Markov

Reprenons la marche aléatoire étudiée à la section 2.2.1. Les sauts successifs X_n sont des variables aléatoires *indépendantes*, en revanche, les positions successives de la particule (la fortune du joueur) ne sont pas indépendantes : la position S_n au temps n dépend de la position S_{n-1} au temps $n-1$ par la relation (2.4). Le point important est que la position à un instant donné n ne dépend *que* de la position à l'instant antérieur $n-1$ et non pas de toute l'histoire de la particule aux temps $0, 1, 2, \dots, n-1$: *Le futur ne dépend du passé que par l'intermédiaire du présent*. Cette marche aléatoire est appelée une chaîne de Markov.²⁴

Les chaînes de Markov, qui jouent un rôle fondamental en finance, représentent une forme de causalité intermédiaire entre les processus complètement indépendants et les processus avec histoire complète (la réalisation à l'instant n dépend de celles à tous les instants antérieurs $0, 1, 2, \dots$).

²⁴Andrei Andreïevitch Markov (1856-1922), mathématicien russe.

2.5.1 Propriété de Markov

D'une façon générale, on considère une variable aléatoire X pouvant prendre des valeurs (discrètes ou réelles) qu'on appelle états. Un processus stochastique est décrit par les valeurs $x_0, x_1, x_2, \dots$ aux instants ordonnés $t_0, t_1, t_2, \dots$ et par une distribution de probabilité initiale $P(x_0, t_0)$ pour la première valeur x_0 de la variable aléatoire X . On dira que ce processus stochastique est *une chaîne de Markov* si, $\forall t_n$, les *probabilités jointes*, $P(x_0, t_0 ; x_1, t_1 ; \dots ; x_n, t_n)$ que X vaille x_0 à l'instant t_0 , x_1 à t_1 ...et x_n à t_n , satisfont la propriété

$$P(x_0, t_0 ; x_1, t_1 ; \dots ; x_n, t_n) = W(x_n, t_n | x_{n-1}, t_{n-1}) \dots W(x_2, t_2 | x_1, t_1) W(x_1, t_1 | x_0, t_0) P(x_0, t_0), \quad (2.32)$$

où $W(x_n, t_n | x_{n-1}, t_{n-1})$ est la *probabilité conditionnelle de transition* pour que le système étant dans l'état x_{n-1} à l'instant t_{n-1} passe dans l'état x_n à l'instant t_n . Les tirages successifs de la variable aléatoire X ne sont donc pas indépendants.

Il faut comprendre en quoi la relation (2.32) est particulière. Dans le cas le plus général, on aurait

$$\begin{aligned} P(x_0, t_0 ; x_1, t_1 ; \dots ; x_n, t_n) &= W(x_n, t_n | x_{n-1}, t_{n-1} ; x_{n-2}, t_{n-2} ; \dots ; x_0, t_0) \cdot \\ &\quad P(x_0, t_0 ; x_1, t_1 ; \dots ; x_{n-1}, t_{n-1}) \\ &= W(x_n, t_n | x_{n-1}, t_{n-1} ; x_{n-2}, t_{n-2} ; \dots ; x_0, t_0) \cdot \\ &\quad W(x_{n-1}, t_{n-1} | x_{n-2}, t_{n-2} ; \dots ; x_0, t_0) \cdot \\ &\quad P(x_0, t_0 ; x_1, t_1 ; \dots ; x_{n-2}, t_{n-2}). \end{aligned}$$

Soit, en itérant, une formule analogue à (2.32), mais où $W(x_n, t_n | x_{n-1}, t_{n-1} ; \dots ; x_0, t_0)$ y serait remplacée par $W(x_n, t_n | x_{n-1}, t_{n-1})$: la valeur de X à t_n ne dépend²⁵ que de la valeur à t_{n-1} .

La donnée des nombres $W(x_n, t_n | x_{n-1}, t_{n-1})$ et de la loi $P(x_0, t_0)$ déterminent donc complètement l'évolution du système, c'est-à-dire permettent de connaître la probabilité $W(x_n, t_n | x_0, t_0)$ d'avoir $X = x_n$ à l'instant t_n sachant que $X = x_0$ à t_0 pour tout n . Voici quelques exemples de chaîne de Markov :

- les mouvements déterministes : les probabilités W et P dépendent alors du temps et sont certaines.
- le mouvement brownien : a priori, à une échelle de temps microscopique, les chocs entre particules dépendent des événements antérieurs, mais à une échelle macroscopique (celle de l'observation) les particules perdent la mémoire des collisions passées.
- les variations d'un actif financier : l'hypothèse markovienne, fréquemment utilisée en finance, est pourtant discutable. Nous y reviendrons.

2.5.2 Équation de Chapman-Kolmogorov

Une chaîne de Markov est dite *homogène* (dans le temps) si les probabilités $W(x_n, t_n | x_{n-1}, t_{n-1})$ ne dépendent pas de l'instant t_n , mais seulement de l'intervalle de temps $t_n - t_{n-1}$.

²⁵Si les tirages de la variable aléatoire X étaient indépendants, $W(x_n, t_n | x_{n-1}, t_{n-1})$ ne dépendrait pas de x_{n-1} et de t_{n-1} .

Considérons une chaîne de Markov homogène dans l'espace continu à 1d, d'après l'équation (2.32), on a

$$P(x_1, t_1 ; x_2, t_2 ; x_3, t_3) = W(x_3, t_3 | x_2, t_2) W(x_2, t_2 | x_1, t_1) P(x_1, t_1). \quad (2.33)$$

En intégrant l'équation (2.33) par rapport à x_2 , il reste

$$P(x_1, t_1 ; x_3, t_3) = W(x_3, t_3 | x_1, t_1) P(x_1, t_1) = P(x_1, t_1) \int W(x_1, t_1 | x_2, t_2) W(x_2, t_2 | x_3, t_3) dx_2.$$

Et en divisant par $P(x_1, t_1)$, on obtient la probabilité d'arriver en x_3 au temps t_3 quand on est parti de x_1 au temps t_1 :

$$W(x_3, t_3 | x_1, t_1) = \int W(x_3, t_3 | x_2, t_2) W(x_2, t_2 | x_1, t_1) dx_2. \quad (2.34)$$

C'est l'équation de Chapman-Kolmogorov. Elle dit simplement que, dans un processus sans mémoire, la probabilité d'aller de 1 en 3 est la somme de toutes les probabilités (indépendantes) d'aller de 1 en 2 puis de 2 en 3. Le temps t_2 ne joue aucun rôle, comme on le verra plus loin sur des exemples.

Si maintenant on intègre l'équation (2.34) par rapport à x_1 , on obtient

$$P(x_3, t_3) = \int W(x_3, t_3 | x_2, t_2) P(x_2, t_2) dx_2.$$

Soit en échangeant les indices $3 \rightarrow 2$ et $2 \rightarrow 1$:

$$P(x_2, t_2) = \int W(x_2, t_2 | x_1, t_1) P(x_1, t_1) dx_1. \quad (2.35)$$

En pratique, les probabilités P et W d'un processus stochastique doivent vérifier les équations (2.34) et (2.35) pour qu'il soit markovien.

Voyons quelques exemples importants de processus markovien.

Processus de Wiener (gaussien)

Nous avons déjà abordé les processus de Wiener section 2.4.3. La probabilité de transition pour un tel processus homogène et non-stationnaire est donnée par une loi gaussienne :

$$W(x_2, t_2 | x_1, t_1) = \frac{1}{\sqrt{4\pi D(t_2 - t_1)}} e^{-\frac{(x_2 - x_1)^2}{4D(t_2 - t_1)}}. \quad (2.36)$$

Alors nécessairement

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}. \quad (2.37)$$

La probabilité de transition (2.36) vérifie bien l'équation de Chapman-Kolmogorov (2.34) qui s'écrit après le changement de variables, $z = x_2 - x_1$, comme un produit de convolution :

$$\int dz P_1(z) P_2[(x_3 - x_1) - z] = P_3(x_3 - x_1),$$

où P_1 et P_2 sont deux gaussiennes centrées de variances $\sigma_1^2 = 2D(t_2 - t_1)$ et $\sigma_2^2 = 2D(t_3 - t_2)$ respectivement. La fonction caractéristique d'une gaussienne centrée d'écart type σ étant $e^{-\frac{\sigma^2 u^2}{2}}$, le produit ordinaire des transformées de Fourier de P_1 et P_2 donne

$$e^{-\frac{(\sigma_1^2 + \sigma_2^2)u^2}{2}},$$

qui est bien la fonction caractéristique d'une gaussienne centrée de variance $\sigma_1^2 + \sigma_2^2 = 2D[(t_2 - t_1) + (t_3 - t_2)] = 2D(t_3 - t_1)$, c'est-à-dire $P_3(x_3 - x_1) = W(x_3, t_3 | x_1, t_1)$. On peut montrer de même que les probabilités (2.36) et (2.37) vérifient bien l'équation (2.35).

On peut montrer que les trajectoires obtenues lors d'un processus gaussien sont continues, mais non dérivables.

Processus de Poisson

On a alors, avec des n_i entiers :

$$W(n_2, t_2 | n_1, t_1) = \frac{(t_2 - t_1)^{n_2 - n_1}}{(n_2 - n_1)!} e^{-(t_2 - t_1)}.$$

où $n_2 > n_1$. On a donc

$$P(n, t) = \frac{t^n}{n!} e^{-t}.$$

Pour démonstration l'équation de Chapman-Kolmogorov (2.34) on pourra utiliser le changement de variables $k = n_2 - n_1$ et le binôme de Pascal

$$[(t_3 - t_2) + (t_2 - t_1)]^{n_3 - n_1} = \sum_{k=0}^{n_3 - n_1} \frac{(n_3 - n_1)!}{k!(n_3 - n_1 - k)!} (t_3 - t_2)^{n_3 - n_1 - k} (t_2 - t_1)^k.$$

Processus de Cauchy

Dans ce cas :

$$W(x_2, t_2 | x_1, t_1) = \frac{t_2 - t_1}{\pi} \frac{1}{(t_2 - t_1)^2 + (x_2 - x_1)^2}.$$

Et aussi

$$P(x, t) = \frac{t}{\pi} \frac{1}{t^2 + x^2}.$$

On vérifie encore à l'aide des transformées de Fourier que l'équation de Chapman-Kolmogorov est satisfaite et que t_2 ne joue aucun rôle.

Dans un processus de Cauchy, les sauts pouvant être de grandes amplitudes, les trajectoires sont discontinues et non dérivables.

2.5.3 Formalisme matriciel

Dans le cas d'une chaîne de Markov *homogène*, on introduit la *probabilité de transition*²⁶ entre l'état x_i et l'état x_j .

$$w_{ij} = Pr(X_{n+1} = x_j | X_n = x_i),$$

Il est pratique d'introduire la *matrice de transition* $\mathcal{W}$ (ou matrice stochastique) dont les éléments sont les probabilités de transition w_{ij} , où i correspond à la ligne et j à la colonne.

$$\mathcal{W} = \begin{pmatrix} w_{11} & w_{12} & w_{13} & \dots \\ w_{21} & w_{22} & w_{23} & \dots \\ \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \dots \end{pmatrix}.$$

Clairement, de l'état i on doit atteindre un certain état j , donc

$$\sum_j w_{ij} = 1.$$

Partant d'une distribution initiale $P(i)$, représentée sous la forme d'un vecteur ligne $\mathcal{P}$, la distribution à l'instant suivant est donnée par le produit matriciel $\mathcal{P}\mathcal{W}$.

Pour simplifier (éviter les matrices infinies), on se limitera au cas où les X_n ont un nombre fini d'états. Reprenons par exemple le problème classique de ruine de la section 2.3. Il s'agissait d'une marche aléatoire à $1d$ avec une probabilité p d'avancer d'un pas et $q = 1 - p$ de reculer d'un pas, avec des parois absorbantes en $x = 0$ et $x = a$ (a entier). La matrice de transition de cette marche aléatoire est

$$\mathcal{W} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 & 0 & 0 \\ 0 & q & 0 & p & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}. \quad (2.38)$$

Un état tel que $w_{i,i} = 1$ est dit *absorbant* (quand on arrive dans l'état i on y reste). Si la matrice contient (au moins) un état absorbant et s'il est possible de passer de tout état non absorbant à un état absorbant, alors la matrice est dite absorbante. Quel que soit l'état initial, l'état final sera dans un état absorbant en un nombre fini de pas et y restera.

En définissant la *probabilité* $w_{ij}^{(m)}$ de passer de l'état i à l'état j en exactement m pas,

$$w_{ij}^{(m)} = Pr(X_{n+m} = x_j | X_n = x_i),$$

²⁶On trouve aussi dans la littérature la convention que w_{ij} représente la transition $j \rightarrow i$.

l'équation de Chapman-Kolmogorov devient

$$w_{ij}^{(2)} = \sum_k w_{ik} w_{kj}.$$

Et plus généralement,²⁷

$$w_{ij}^{(m+n)} = \sum_k w_{ik}^{(m)} w_{kj}^{(n)}.$$

En remplaçant le produit ordinaire des probabilités de tirages indépendants par un produit de matrices, on voit que les $w_{ij}^{(m)}$ sont les éléments de la matrice $\mathcal{W}^m$. Puisqu'en partant de l'état i , on arrive bien quelque part en m pas, un certain état j , on doit avoir $\forall m$,

$$\sum_j w_{ij}^{(m)} = 1. \quad (2.39)$$

Ce formalisme matriciel est très pratique pour étudier le comportement asymptotique ($m \rightarrow \infty$), comme nous allons le voir dans l'exemple suivant.

Chaîne de Markov à deux états

Soit X une variable aléatoire qui ne peut prendre que deux valeurs (On peut penser à la vitesse d'un mobile sur un axe), $x = v$ (état 1) et $x = -v$ (état 2). On appellera x_n les réalisations successives de X avec la loi suivante : quand $x_n = v$, x_{n+1} a la probabilité p de valoir $-v$, et donc $1 - p$ de conserver v ; quand $x_n = -v$, x_{n+1} a la probabilité α de valoir v et $1 - \alpha$ de conserver $-v$. On appellera X_n la variable aléatoire correspondante. On écrira, $\forall n \geq 0$

$$\begin{aligned} Pr(X_{n+1} = -v | X_n = v) &= p = w_{12} \\ Pr(X_{n+1} = -v | X_n = -v) &= 1 - \alpha = w_{22} \\ Pr(X_{n+1} = v | X_n = v) &= 1 - p = w_{11} \\ Pr(X_{n+1} = v | X_n = -v) &= \alpha = w_{21}. \end{aligned}$$

Ce qui définit la matrice de transition

$$\mathcal{W} = \begin{pmatrix} 1-p & p \\ \alpha & 1-\alpha \end{pmatrix}.$$

²⁷Dans le cas où les X_n ont des valeurs réelles, on définira une densité de probabilité conditionnelle $f^{(m)}(y|x)$ telle que

$$f^{(m)}(y|x) dx dy = Pr(y < X_{n+m} < y + dy | x < X_n < x + dx).$$

Comme dans le cas discret on aura

$$f^{(m+n)}(z|x) = \int f^{(m)}(z|y) f^{(n)}(y|x) dy.$$

La fonction $f(y|x)$ est une fonction positive telle que $\int f(y|x) dy = 1$, $\forall x$. On peut enfin considérer que les instants m deviennent continus. On pourra alors intégrer, dériver par rapport à t .

Connaissant $x_0 = \pm v$, les quatre valeurs de $Pr(X_n = \pm v | X_0 = \pm v)$ sont données par les éléments de la matrice $\mathcal{W}^m$. Le calcul donne

$$\mathcal{W}^m = \frac{1}{\alpha + p} \begin{pmatrix} \alpha & p \\ \alpha & p \end{pmatrix} + \frac{(1 - \alpha - p)^m}{\alpha + p} \begin{pmatrix} p & -p \\ -\alpha & \alpha \end{pmatrix},$$

dont la limite quand $m \rightarrow \infty$ est $\mathcal{W}^\infty = \frac{1}{\alpha + p} \begin{pmatrix} \alpha & p \\ \alpha & p \end{pmatrix}$. Les deux lignes étant identiques, la probabilité asymptotique d'avoir la vitesse v (ou $-v$) ne dépend pas de la condition initiale.

On peut montrer que la convergence de $\mathcal{W}^m$ vers $\mathcal{W}^\infty$ et “l'oubli” des conditions initiales vu dans cet exemple est plus général : il s'applique à toutes les matrices *régulières*. Une matrice $\mathcal{W}$ est dite régulière s'il existe un entier r tel que tous les éléments de la matrice $\mathcal{W}^r$ soient strictement positifs.²⁸ Dans ce cas, le système tend vers une distribution d'équilibre $\mathcal{P}_e$, telle que $\mathcal{P}_e \mathcal{W} = \mathcal{P}_e$. La distribution d'équilibre d'une chaîne de Markov régulière est donc un vecteur propre de la matrice de transition $\mathcal{W}$ associé à la valeur propre 1.

²⁸Une matrice $\mathcal{W}$ qui contient un état absorbant ne peut pas être régulière. C'est par exemple le cas de la matrice (2.38).

2.6 Application à la modélisation financière

October is one of the particularly dangerous months to speculate in stocks. The others are July, January, September, April, November, May, March, June, December, August and February

(Mark Twain)

La valeur future d'un actif financier (action, obligation, matière première...) ne peut pas être connue avec certitude, mais une modélisation financière permet d'estimer sa distribution de probabilité à une date ultérieure.²⁹ Comme une particule brownienne soumise aux chocs incessants des molécules d'un fluide, la valeur d'un actif est la résultante de nombreux ordres de vente et d'achat. Il suffit de comparer les variations de l'indice CAC40 représentées sur la figure 2.12 et la marche aléatoire brownienne de la figure 2.2 pour apprécier la ressemblance. Mais comme nous le verrons dans cette section, cette analogie a des limites.

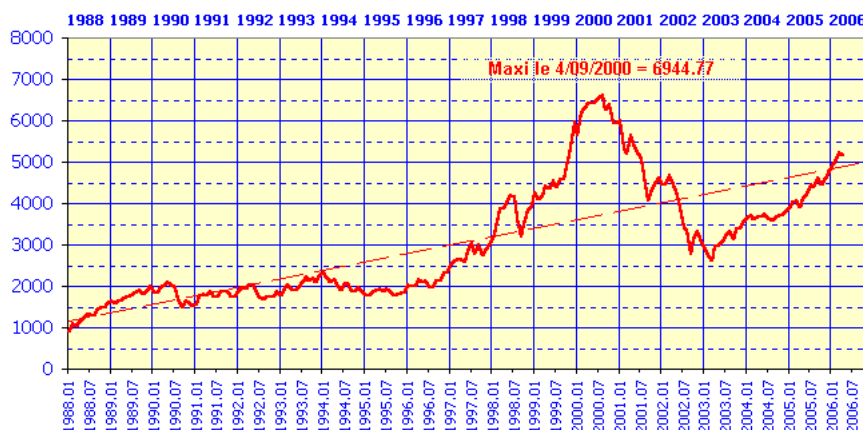


FIG. 2.12 – Variation du CAC40 depuis sa création le 31/12/1987.

Remarquons qu'il n'y a pas d'échelle de temps idéale en finance. Les marchés étant fermés plusieurs heures par jour, le temps physique ne convient pas. Le "trading time" (temps pendant lequel le marché est ouvert) a l'inconvénient de ne pas être uniforme puisque l'activité varie dans la journée et qu'il y a une discontinuité lors de la fermeture du marché... Les analyses de données financières dépendent donc du choix de l'échelle de temps.

²⁹Cette section s'inspire de deux classiques : *The Mathematics of Financial Derivatives : A Student Introduction*, P. Willmott, S. Howison and J. Dewynne (Cambridge University Press, 1997) et *Options, futures et autres actifs dérivés*, John Hull et al. (Pearson Education, 2004).

Des actifs financiers aux produits dérivés

Un actif financier est titre ou un contrat entre un débiteur et un créateur susceptible de rapporter un gain en capital à ce dernier. Il en existe de différents types comme les *obligations*, à revenus fixes (“bond” en anglais), ou les *actions*, à revenus variables (“stocks”).

Ainsi, une *action* est une part du capital d’une société qui donne à son porteur, l’actionnaire, le droit de toucher un *dividende* en fonction des profits réalisés par la société. Les parts ont une valeur qui reflète le point de vue des investisseurs sur les futurs profits d’une entreprise et donc sur les éventuels dividendes. Cette valeur est fixée par le prix auquel les parts sont achetées et vendues sur le marché boursier. Elle représente donc l’effet cumulé des achats et ventes réalisés par les nombreux opérateurs.

Les *produits dérivés* (“dérivatives”) sont des contrats plus sophistiqués qui reposent sur des *actifs sous-jacent* (action, obligation, matière première). Ils permettent de réaliser des opérations boursières en limitant les risques comme une assurance.³⁰

Par exemple, une *option d’achat* (“european call option”) est un produit dérivé asymétrique qui donne le droit - et non l’obligation - d’acheter une quantité donnée d’un actif financier sous-jacent S (“underlying”) à une *date d’échéance*³¹ convenue (“expiry date”), ou à maturité, à un prix fixé à l’avance, le *prix d’exercice* K (“strike” ou “exercice price”). L’évolution attendue de l’actif financier conditionne donc l’achat ou la vente d’options. Par exemple, si l’on pense que la valeur de l’actif augmentera, on achètera des options d’achat (ou on émettra des options de vente (“european put option”). Mais cette assurance contre l’augmentation des cours a un prix. Et puisque la valeur de l’actif à maturité ne peut être connue à l’avance, la question centrale est comment fixer la valeur d’une option V ?

Exemple :

Pour une option d’achat³² dont le prix d’exercice K est 250 à la date d’échéance. Quelle est la valeur C de l’option à maturité et à $t = 0$?

Dans un modèle stochastique, la valeur future de l’actif sera distribuée selon une certaine loi. Pour simplifier, supposons qu’à la date d’échéance, l’actif ne puisse prendre que deux valeurs S avec la même probabilité : $S = 270$ ou $S = 230$.

Si $S = 270$ le souscripteur (“holder”) peut alors acheter l’actif au prix d’exercice $K = 250$ et réalise donc un profit immédiat de 20 (en le revendant au prix du marché). La valeur de l’option à maturité est donc $C = S - K = 20$. Au contraire, si $S = 230$, le souscripteur n’exercera pas son droit d’option et $C = 0$. D’une façon générale, la valeur de l’option (“call”) à maturité est (voir figure 2.13)

$$C = \max(0, S - K)$$

³⁰Les marchands génois lorsqu’ils affrétaient un navire, achetaient une option sur un second. Si le premier n’arrivait pas à bon port, l’assurance permettait d’acheter la cargaison du second à un prix fixé à l’avance. Sinon, l’option n’était pas exercée.

³¹Ou n’importe quand avant une date donnée pour les options américaines.

³²Dans la suite nous ne considérerons que le cas des options d’achat.

et le profit escompté, correspondant à la valeur de l'option est $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 20 = 10$. Cette estimation du gain permet de fixer, dans ce cas simple, le prix de l'option à la signature du contrat à $t = 0$: $C_0 = 10$. Ainsi, dans le cas où $S = 270$, le profit sera $20 - 10 = 10$ et dans le cas où $S = 230$, $0 - 10 = -10$. Soit 100% de perte ou de profit (gain/perte de 10 pour un investissement de 10).

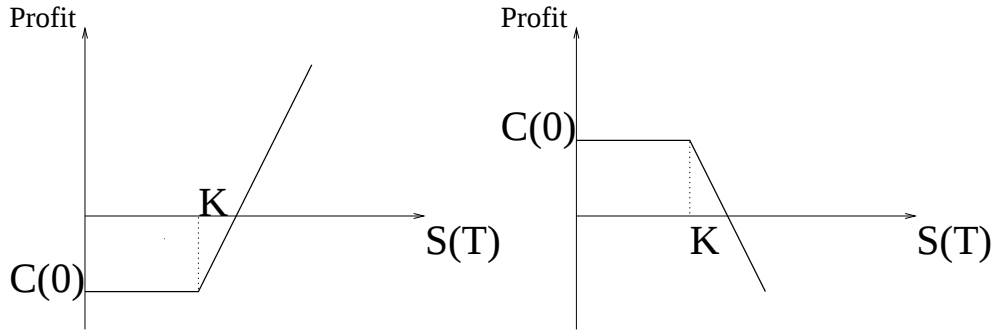


FIG. 2.13 – Profit réalisé par un acheteur (à gauche) et un vendeur (à droite) d'une option d'achat en fonction de la valeur de l'actif sous-jacent à maturité.

Dans cet exemple simple, la distribution de l'actif était supposée uniforme. Avant de proposer une loi de probabilités plus réaliste, voyons de quels paramètres peut dépendre a priori la valeur C d'une option d'achat :

- la valeur de l'actif S_0 à la signature du contrat (plus S_0 est élevée, plus il y a de chance que C le soit à maturité)
- la valeur de l'actif S à maturité (plus S est élevée plus le profit, et donc C , est important)
- le prix d'exercice K (plus K est bas, plus C sera élevée)
- le temps (plus la date d'échéance approche, plus les valeurs de S , et donc de C , sont prévisibles)
- la *volatilité* de l'actif qui mesure les fluctuations (plus elle est grande plus la valeur de S , et donc de C , peuvent fluctuer)
- le taux d'intérêt (qui dépend a priori du temps) que l'on suppose connu sur des temps courts³³ qui permet un gain certain (compte en banque).

Voyons maintenant deux modélisations du prix d'un actif financier et d'une option.

³³Combien suis-je prêt à payer maintenant pour recevoir une somme E à la date T ? si les taux d'intérêts, r , sont constants, l'argent $M(t)$ déposé à la banque, varie comme $dM = rMdt$, soit $M(t) = \alpha e^{rt}$ et à $t = T$, $M(T) = E$, donc $M(t) = Ee^{-r(T-t)}$.

2.6.1 L'approche de Bachelier

En 1900, Louis Bachelier³⁴ propose dans sa thèse la première modélisation financière. Les principales hypothèses de ce modèle sont les suivantes :

- la valeur d'un actif financier varie à *chaque instant* par saut successif $\pm v$ avec la même probabilité $1/2$: "Il semble que le marché, c'est-à-dire l'ensemble des spéculateurs ne doit croire à un instant donné ni à la hausse, ni à la baisse, puisque, pour chaque côté, il y a autant d'acheteurs que de vendeurs".
- les variations de l'actif sont *indépendantes* d'un instant à l'autre (propriété de Markov) : "On peut admettre que la probabilité d'un écart à partir du cours vrai est indépendante de la valeur absolue de ce cours, et que la courbe des probabilités est symétrique par rapport au cours vrai".

On reconnaît bien sûr la marche aléatoire étudiée à la section 2.2. La probabilité $P(s, t)$ que l'actif vaille s € à l'instant t est donc donnée par l'équation de la diffusion (2.26) avec $c = 0$ et s'exprime à l'aide de l'équation (2.25) où le coefficient D , lié à la variance, est appelé par Bachelier "le coefficient d'instabilité". On parlerait aujourd'hui de la *volatilité*.

La valeur d'un actif à un instant donné étant une somme de variables aléatoires distribuées sur une loi de variance finie ($\pm v$ avec la probabilité $1/2$), la probabilité $P(s, t)$ est bien une loi gaussienne d'après le théorème de la limite centrale.

Le modèle de Bachelier permet de calculer la valeur d'une option. En effet, la probabilité pour que la valeur d'un actif S soit supérieure à K à maturité t est, d'après l'équation (2.25) :

$$P(t) = \frac{1}{\sqrt{4\pi Dt}} \int_K^\infty e^{-\frac{(s-s_0)^2}{4Dt}} ds,$$

où s_0 est le prix de l'actif (connu) à $t = 0$. Le gain moyen attendu, si $S > K$ est donc

$$G(t) = \langle S - K \rangle = \frac{1}{\sqrt{4\pi Dt}} \int_K^\infty (s - K) e^{-\frac{(s-s_0)^2}{4Dt}} ds.$$

Avec $K = S_0$, ce bénéfice vaut $\sqrt{Dt/\pi}$, il croît donc en $\sqrt{t}$ ce qui d'après Bachelier est "un des résultats les plus importants de notre étude".

Comme nous le verrons, plusieurs réserves peuvent être émises sur ce modèle. En particulier, cette probabilité est seulement asymptotique à la limite des temps longs.

2.6.2 Le modèle de Black-Scholes

Dans le cadre du modèle de Black-Scholes, on suppose que ce sont les variations relatives $\delta S/S = \delta \ln S$ des actifs qui sont distribuées sur une loi gaussienne. La valeur d'un actif exprimée en devise n'est pas en effet une bonne

³⁴Louis Bachelier (1870-1946) mathématicien français fondateur des mathématiques financières. Les citations suivantes sont tirées de sa thèse : "Théorie de la spéculation", Annales scientifiques de l'École Normale Supérieure, 3ème série, Tome 17 (1900), pp.21-86 que l'on trouve sur internet : <http://www.numdam.org/>

variable puisque la valeur d'une monnaie varie (inflation, croissance, fluctuation sur le marché des changes). En revanche, $\delta S/S$ donne directement le pourcentage de gain ou de perte sur une période donnée. Ensuite, le modèle permet d'estimer la valeur d'une option en construisant un portefeuille sans risque dont l'évolution sur un temps court n'est plus aléatoire...

Le cours d'un actif

Pratiquement tous les modèles donnant le prix d'une option sont basés sur une hypothèse simple concernant l'évolution du prix des actifs : le processus de Wiener.³⁵

D'abord, la variation absolue dS d'un actif n'a pas beaucoup de sens, ce qui est significatif,³⁶ c'est sa variation relative dS/S . Ensuite, l'actif peut évoluer de deux façons : La première, prévisible, correspond à la tendance générale du marché qui peut être à la hausse ou la baisse. Cette variation est typiquement liée au taux d'intérêt.³⁷ La seconde modélise l'aspect aléatoire du à l'accumulation d'effets extérieurs. Et a priori, on peut supposer que l'incertitude sur la valeur relative de l'actif est indépendante de sa valeur S . Ainsi, pendant un intervalle de temps $dt \rightarrow 0$, la variation relative de l'actif sera donnée par l'équation différentielle stochastique :

$$\frac{dS}{S} = \mu dt + \sigma dX. \quad (2.40)$$

Le premier terme correspond donc à une dérive ("drift") déterministe caractérisée par un *taux de croissance* μ (lié au placement sans risque dans une banque). S'il n'y avait que ce terme, on aurait $S(t) = S_0 e^{\mu t}$ avec certitude. Le second terme modélise la composante aléatoire de la variation relative dS/S sous la forme d'un processus de Wiener, c'est à dire une variable aléatoire dX distribuée sur une loi normale centrée ($\langle dX \rangle = 0$) de variance dt , et pondérée par la *volatilité* σ qui caractérise la taille des fluctuations. La variation relative dS/S est donc distribuée sur une loi normale de moyenne $m = \mu dt$ et de variance $\sigma^2 dt$ ou encore $\langle dS \rangle = \mu S dt$ et $Var(dS) = \sigma^2 S^2 dt$. Typiquement, $dt = 1/250$ (il y a 250 jours ouvrables dans l'année) et la volatilité, qui peut être estimée à partir des données passées, vaut $\sigma = 0.05 - 0.4$ par $\sqrt{an}$.

En première approximation, ce modèle décrit très bien les cours des actifs, mais comme nous le verrons, les fluctuations gaussiennes interdisent les grandes fluctuations (krach boursier) pourtant observées en 1929, 1987...

La variable aléatoire S suit donc un processus d'Itô³⁸ avec $a = \mu S$ et $b = \sigma S$ (voir section 2.4.3). D'après le lemme d'Itô, la variation d'une fonction $f(S, t)$

³⁵L'hypothèse d'efficience des marchés justifie le processus stochastique : i) l'histoire passée de l'actif est reflétée par sa valeur présente (propriété de Markov) ii) les marchés réagissent instantanément à tout type d'information.

³⁶Quelle que soit la valeur de l'actif, un investisseur exigera une rentabilité moyenne donnée, par exemple 7%.

³⁷Car plus les taux d'intérêt sont élevés, plus les investisseurs exigent un haut taux de rentabilité.

³⁸Pour améliorer le modèle, on peut supposer que, μ et σ dépendent également du temps t .

est donnée par l'équation (2.31) :

$$df = \left(\frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dX. \quad (2.41)$$

Une première conséquence du lemme d'Itô est la détermination de la distribution de la variable S . En utilisant l'équation (2.41) et en choisissant $f(S) = \ln S$ on a immédiatement :

$$df = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dX.$$

La variation $df = d \ln S$ est donc distribuée normalement et $\ln S = f(t) = \int_0^t df$ suit une gaussienne de moyenne $\ln s_0 + (\mu - \frac{1}{2} \sigma^2)t$ de variance $\sigma^2 t$, où s_0 est la valeur initiale de S . La distribution de l'actif S est donc une loi log-normale (voir section 1.3.2) :

$$P(s) = \frac{1}{\sigma \sqrt{2\pi t}} \frac{1}{s} e^{-\frac{[\ln \frac{s}{s_0} - (\mu - \frac{1}{2} \sigma^2)t]^2}{2\sigma^2 t}} \quad \text{pour } s > 0.$$

Le modèle de Black-Scholes

Le modèle de Black-Scholes a été développé en 1973 par Fisher Black et Myron Scholes d'une part et par Robert Merton d'autre part.³⁹ Il permet de déterminer la valeur d'une option $V(S, t)$ (call ou put) en fonction du temps et du prix de l'actif financier sous-jacent. Il repose sur les hypothèses suivantes :

- La distribution de l'actif est une loi log-normale (sa variation est donc donnée par l'équation (2.40)).
- Le taux d'intérêt r et la volatilité σ de l'actif sont connus en fonction du temps.
- Il n'y a ni coûts de transaction ni d'impôts.
- Il n'y a pas de dividende sur l'actif sous-jacent avant la date de maturité.
- Il n'y a pas d'opportunités d'arbitrage : il n'est pas possible de réaliser un profit certain *sans risque* et sans coût.⁴⁰
- Le marché évolue continûment.
- Il est possible de vendre ou d'acheter n'importe quelle quantité d'actifs (vente à découvert d'un actif non possédé).

Plusieurs de ces hypothèses peuvent être abandonnées. Par exemple, le taux d'intérêt peut être lui aussi une variable aléatoire. Construisons un portefeuille contenant une option et une quantité $-\alpha$ (qui peut être négative ou positive) de l'actif sous-jacent. La valeur du portefeuille est donc :

$$\Pi = V - \alpha S,$$

³⁹Leurs travaux ont été récompensés en 1997 par le "prix de la Banque de Suède à la mémoire d'Alfred Nobel". Le "prix nobel d'économie" d'existant pas, ce pseudo-Nobel est accordé depuis 1969.

⁴⁰L'arbitrage, c'est réalisé un gain sans risque. Par exemple acheter des oranges à Miami et les revendre plus chères à Naples. "Rapidement", le prix des oranges va augmenter à Miami (car la demande augmente) et diminuer à Naples (car l'offre augmente). Ainsi dès qu'elles apparaissent, les opportunités d'arbitrage sont censées être éliminées.

et sa variation en dt est :

$$d\Pi = dV - \alpha dS.$$

La variation dV de la valeur de l'option est donnée par le lemme d'Itô (2.31). Comme nous l'avons remarqué à la section 2.4.3, les deux variables aléatoires S et $V(s, t)$ sont corrélées et dépendent de la même variable aléatoire dX , on peut donc choisir une valeur de α qui élimine le terme en dX de façon à obtenir un portefeuille dont la variation $d\Pi$ est complètement déterministe pendant la durée dt :

$$\begin{aligned} d\Pi &= dV - \alpha dS \\ &= \left(\frac{\partial V}{\partial S} \mu S + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial V}{\partial S} \sigma S dX - \alpha (\mu S dt + \sigma S dX) \\ &= \left(\frac{\partial V}{\partial S} - \alpha \right) \sigma S dX + \left[\mu S \left(\frac{\partial V}{\partial S} - \alpha \right) + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right] dt. \end{aligned}$$

En choisissant $\alpha = \frac{\partial V}{\partial S}$ le terme en dX disparaît. Le paramètre α mesure la corrélation entre les variations de l'option et celles de l'actif. La dynamique de la variable $d\Pi$ est donc complètement déterministe :

$$d\Pi = \left[\frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right] dt.$$

Utilisons à présent l'hypothèse d'absence d'opportunités d'arbitrage (et de coût de transaction). Partant d'un capital Π , il est possible d'obtenir un gain certain et *sans risque* en le plaçant sur un compte en banque dont le taux d'intérêt vaut r . Dans ce cas, l'augmentation en dt est certaine et vaut :

$$d\Pi_0 = r\Pi dt.$$

Si $d\Pi_0 < d\Pi$, il serait possible d'emprunter de l'argent avec le taux d'intérêt r , de l'investir dans le portefeuille et de réaliser un profit certain et *sans risque* (en tenant compte du remboursement de l'emprunt). Au contraire, si $d\Pi_0 > d\Pi$, il suffit de placer le portefeuille sur un compte en banque pour réaliser un profit certain. Dans les deux cas, le profit *sans risque*, *sans coût* et instantané viole l'hypothèse d'absence d'opportunités d'arbitrage. On a donc nécessairement $d\Pi_0 = d\Pi$, soit :

$$r\Pi dt = \left[\frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right] dt.$$

Et puisque $\Pi = V - \frac{\partial V}{\partial S} S$, on obtient l'équation différentielle de Black-Scholes :

$$\boxed{\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0.} \quad (2.42)$$

Remarquons que l'équation de Black-Scholes ne contient pas le taux de croissance μ : la valeur d'une option est indépendante de l'évolution moyenne de l'actif sous-jacent. En revanche la volatilité σ de l'actif joue un rôle central dans l'estimation de la valeur de l'option.

En supposant r et σ constants pendant la durée de vie de l'option, on peut résoudre analytiquement l'équation (2.42) (voir TD) pour *des conditions aux limites particulières dépendantes du produit dérivé considéré*.

Pour une option d'achat C , on a vu que

$$C(S, T) = \max(0, S - K). \quad (2.43)$$

Pour une option de vente ("put") P , on aurait de même :⁴¹

$$P(S, T) = \max(0, K - S). \quad (2.44)$$

Par ailleurs, on peut montrer qu'il existe une *équation de parité* exprimant à chaque instant la valeur d'une option d'achat C en fonction de la valeur d'une option de vente P (voir TD) :

$$S(t) + P(S, t) - C(S, t) = Ke^{-r(T-t)}. \quad (2.45)$$

Finalement, en utilisant les conditions aux limites (2.43) et (2.44), on calcule la valeur d'une option d'achat C et d'une option de vente P respectivement (voir TD) :

$$C(S, t) = SN(d_1) - Ke^{-r(T-t)}N(d_2) \quad (2.46)$$

$$P(S, t) = -SN(-d_1) + Ke^{-r(T-t)}N(-d_2), \quad (2.47)$$

où

$$d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\ln(\frac{S}{K}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}},$$

et

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz,$$

où $N(x)$ est la fonction de répartition de la loi normale centrée et réduite (la probabilité qu'une variable gaussienne soit inférieure à x). On voit une fois de plus que l'approximation gaussienne est au cœur du modèle de Black-Scholes.

Par ailleurs le coefficient $\alpha = \frac{\partial V}{\partial S}$ donnant la quantité relative d'actifs sous-jacents et de produits dérivés est donné par,

$$\frac{\partial C}{\partial S} = N(d_1) \quad (2.48)$$

$$\frac{\partial P}{\partial S} = N(d_1) - 1, \quad (2.49)$$

⁴¹En effet, une option de vente permet de vendre un actif sous-jacent à un prix d'exercice K fixé à l'avance. Si à maturité $K > S$, on exerce l'option qui vaut alors $P(S, T) = K - S$. Dans le cas contraire, on vendra au prix du marché S et alors, $P(S, T) = 0$.

pour une option d'achat et de vente respectivement. Ce coefficient (parfois appelé Δ) est fondamental, puisqu'il permet d'annuler les risques (variables aléatoires) sur un intervalle de temps dt : Il doit donc être ajusté régulièrement.⁴²

Terminons par des critiques essentielles du modèle de Black-Scholes :

- Contrairement à l'hypothèse de Markov, les variations des valeurs boursières peuvent être fortement corrélées sur des temps courts.
- Ces variations sur certaines échelles de temps peuvent être distribuées sur des lois larges. Si leur distribution était purement gaussienne, il n'y aurait jamais de krach (il faudrait attendre l'âge de l'univers pour observer une fluctuation comparable à celle du krach boursier de 1987!).
- Le temps n'est pas strictement continu et le passage à la limite n'est donc pas toujours justifié.

⁴²Les transactions ayant un coût, le prix de cet ajustement doit être pris en compte dans le modèle.