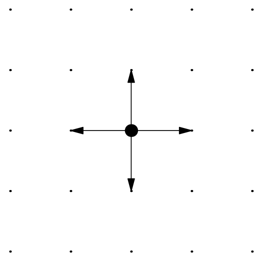
1 Mouvement brownien

Si l'on observe sous un microscope des grains de pollen en suspension dans l'eau, on s'aperçoit qu'ils sont animés d'un mouvement continu et qui paraît totalement désordonné. Nous sommes en présence d'un phénomène *régi par le hasard*, d'un *phénomène aléatoire* (du latin : alea, coup de dé). Ce phénomène est appelé *mouvement brownien*, du nom du botaniste écossais Robert BROWN (1773-1858) qui le découvrit au début du dix-neuvième siècle.

La première théorie en a été élaborée au début du vingtième siècle par Albert EINSTEIN et Marian von SMOLUCHOWSKI (physicien autrichien, 1872-1917).

1.1 Un modèle discret du mouvement brownien

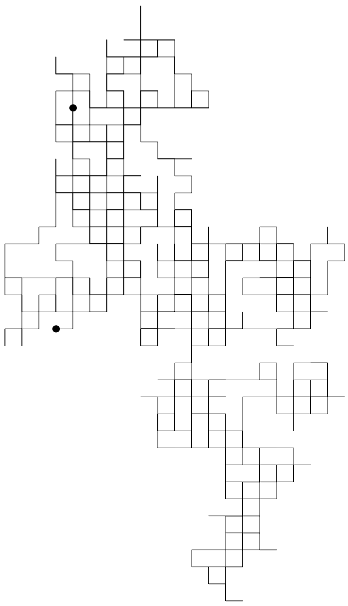
Nous allons construire un modèle (grossier) discret du mouvement brownien à l'aide d'une simulation.



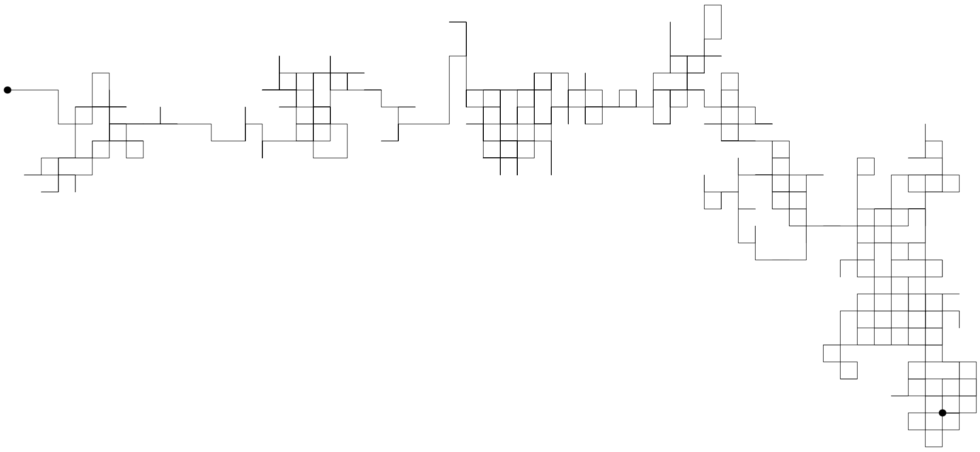
Observons une particule qui se déplace dans le réseau plan infini ci-contre :

L'*expérience aléatoire* (der Zufallsversuch) consiste en ceci : à l'instant 0, la particule se trouve en $(0;0)$; elle choisit au hasard un des quatre points cardinaux N, O, S, E et se déplace d'une maille dans la direction choisie ; la particule choisit à nouveau un des quatre points cardinaux, se déplace, et ainsi de suite, n fois.

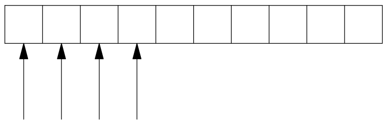
Voici un exemple d'une trajectoire de longueur $n = 1\,000$:



Voici un deuxième exemple, également de longueur $n = 1\,000$.



Cette trajectoire est un des *résultats possibles* de notre expérience aléatoire. On parle encore d'une *épreuve*, ou d'une *éventualité*. La totalité, l'ensemble de toutes les éventualités de notre expérience est notée Ω. C'est l'univers des possibles. Calculons le cardinal de Ω : combien d'épreuves différentes peuvent être produites par notre expérience, combien de trajectoires différentes y a-t-il ? Pour résoudre ce problème, il est nécessaire de trouver une bonne représentation, ou schématisation d'une éventualité. Remarquons qu'il y a autant de trajectoires que de « mots » de n lettres que l'on peut former à l'aide des quatre lettres N, O, S, E ; il suffit donc de compter ces mots. Voici le schéma d'un mot de longueur $n = 10$:



4 possibilités

Pour la première case, il y a quatre possibilités ; pour la seconde, quatre également ; on peut donc remplir les deux premières cases de 4^2 façons différentes. Le cardinal de Ω est donc : $\#\Omega = 4^n = 2^{2n}$. Aucune de ces éventualités n'est privilégiée par rapport aux autres ; chacune a la même chance d'apparaître au cours de l'expérience aléatoire : $\frac{1}{2^{2n}}$. Le nombre $\frac{1}{2^{2n}}$ mesure donc le *degré de possibilité* d'être réalisée d'une trajectoire ; on dit que la *probabilité* d'une trajectoire est $\frac{1}{2^{2n}}$. Ce nombre est infiniment petit, ce qui correspond au fait qu'il est quasiment *impossible* de voir apparaître une trajectoire bien particulière.

Une trajectoire donnée est donc imprévisible et plutôt difficile à décrire ; observer une particule, c'est concentrer son attention non sur les aspects tout à fait individuels de la trajectoire qu'on voit se réaliser, mais sur des aspects plus généraux, plus simples à décrire et à observer. On passe pour ainsi dire
du *niveau microscopique* au *niveau macroscopique*, ou encore (pour utiliser le langage des biologistes) : du *génotype* au *phénotype*.

On pourrait par exemple s'intéresser aux *événements aléatoires* suivants :

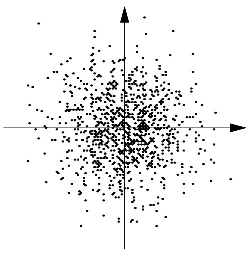
A la particule finit par s'arrêter dans le premier quadrant ;

B la particule finit par s'arrêter dans le demi-plan supérieur.

Ces événements définissent des parties, des sous-ensembles de Ω. Toute trajectoire dont la dernière composante (la position d'arrêt) est à coordonnées positives appartient à cette partie A, *réalise* cet événement A.

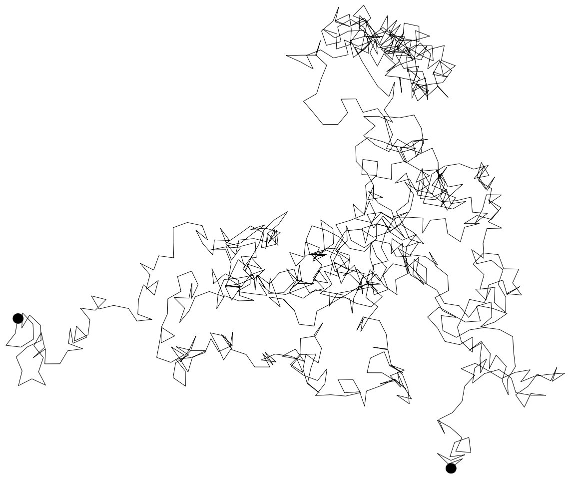
Pour des raisons de symétrie, il est clair que $\frac{1}{4}$ des éléments de Ω réalisent cet événement ou, comme on dit aussi : un quart des *cas possibles* sont *favorables* à cet événement. Le nombre $\frac{1}{4}$ mesure donc le degré de possibilité de notre événement A ; on dit encore que la *probabilité* de l'événement A est $\frac{1}{4}$. On note alors : $P(A) = \frac{1}{4}$. Une réflexion analogue permet d'écrire : $P(B) = \frac{1}{2}$.

Mais alors, si nous répétons *un grand nombre de fois* notre expérience aléatoire, nous pouvons nous attendre à observer qu'*environ* un quart des particules s'arrêtent dans le premier quadrant. La figure suivante représente les positions finales de 1 000 particules. Il y a exactement 242 particules dans le premier quadrant, ce qui correspond très bien à la probabilité de $\frac{1}{4}$.



1.2 Un modèle plus fin

Pour rendre notre simulation d'un *processus de diffusion* plus réaliste, on peut renoncer à la discrétisation du choix de la direction de déplacement. Voici la représentation graphique d'un mouvement brownien où le choix de la direction de déplacement s'est fait de manière continue, non discrète :



Et voici un autre exemple : comme dans la figure précédente, la particule a fait 1 000 pas de 2 mm. Ce qui fait un chemin de deux mètres !

Cependant, la distance qui sépare le point de départ et le point d'arrivée est nettement plus petite et soumise à de fortes fluctuations aléatoires.

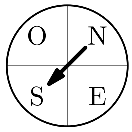
La *distance* D_n qui sépare point de départ et point d'arrivée est une grandeur variable, qui dépend du hasard ; on dit que c'est une *variable aléatoire* associée à notre expérience aléatoire. La moyenne pondérée des valeurs de la variable aléatoire D_n^2 appelée *espérance mathématique* et notée $E(D_n^2)$, permet alors de *prédire* la *moyenne* d'un grand nombre de valeurs observées de D_n. On peut montrer que $E(D_n^2) = n$, et donc que $E(D_n) \leq \sqrt{n}$. Ce qui, dans le cas où $n = 1\,000$, signifie que $E(D_n)$ est proche de 30. Donc : si l'on répète l'expérience aléatoire un très grand nombre de fois, la moyenne des distances observées sera proche de 30 unités de longueur.

2 Quelques expériences aléatoires simples

L'expérience aléatoire précédente était déjà relativement compliquée. En fait, elle est basée sur la répétition d'une expérience aléatoire beaucoup plus simple : le *choix au hasard* d'un des quatre points cardinaux. Pour simuler correctement un mouvement brownien, on doit donc trouver un système qui réalise concrètement ce choix aléatoire, de telle manière qu'aucun des points cardinaux, qu'aucune des quatre directions ne soit privilégiée : les quatre éventualités doivent être *équiprobables*.

Voici quelques *systèmes aléatoires* simples, pour ainsi dire « traditionnels », qui vont nous servir de systèmes-modèles.

2.1 La roue



Le disque a été subdivisé en quatre secteurs de même aire ; si l’on fait tourner la petite aiguille, elle s’arrêtera avec la même probabilité dans chacun des secteurs N, O, S, E.

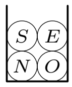
Chacune des éventualités sera choisie avec la même probabilité, à savoir $\frac{1}{4}$.

La réalisation pratique d’une roulette à probabilités uniformes demande beaucoup de soin.

2.2 Les pièces de monnaie

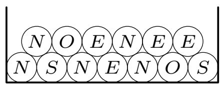
Lançons deux pièces de monnaie (ou lançons la même pièce deux fois de suite). Quatre événements sont possibles : pile-pile, pile-face, face-pile, face-face.
Il y a alors $4! = 4 \cdot 3 \cdot 2 \cdot 1$ bijections possibles entre les ensembles $\Omega_1 = \left\{00, 01, 10, 11\right\}$ et $\Omega_2 = \left\{N, O, S, E\right\}$. Un de ces « codages » permet d’opérer un choix aléatoire parmi les points cardinaux à l’aide de pièces de monnaie.

2.3 L’urne



On marque quatre boules des signes N, O, S, E et on les place dans une urne ; l’expérience aléatoire consiste à tirer une de ces boules.

Nous pouvons supposer qu’aucune des boules ne soit privilégiée par rapport aux autres. Chacune des éventualités se produira avec la même probabilité $\frac{1}{4}$. Remarquons cependant qu’en pratique, on doit prendre beaucoup de précautions pour rendre le tirage à probabilité uniforme.



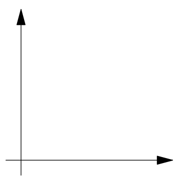
Mais on pourrait arranger les choses autrement :

Est-ce que les quatre événements N, O, S, E sont encore *équiprobables* ?

Complétez le tableau suivant :

événement	N	O	S	E
probabilité	$\frac{5}{13}$			

Si l’on utilisait cette urne pour simuler le mouvement brownien, les directions N et E seraient choisies plus fréquemment ; les particules auraient donc une nette tendance à se diriger vers le NE.



Le graphique suivant montre de nouveau les positions finales de 1 000 particules effectuant une marche aléatoire de 1 000 pas, la direction de déplacement étant choisie suivant les probabilités ci-dessus.

2.4 Les dés

L'expérience aléatoire qui consiste à lancer un dé a six éventualités. Jetons deux dés discernables (ayant par exemple deux couleurs distinctes). L'univers des possibles Ω est de cardinal 36 : $\Omega = \left\{11; 12; \ldots; 66\right\}$. Ces 36 éventualités sont équiprobables ; chacune a la probabilité $\frac{1}{36}$. Considérons les événements suivants :

$$\begin{aligned} A &= \left\{xy \in \Omega \mid x \text{ est pair et } y \text{ est pair}\right\} \\ B &= \left\{xy \in \Omega \mid x \text{ est pair et } y \text{ est impair}\right\} \\ C &= \left\{xy \in \Omega \mid x \text{ est impair et } y \text{ est pair}\right\} \\ D &= \left\{xy \in \Omega \mid x \text{ est impair et } y \text{ est impair}\right\} \end{aligned}$$

Ces quatre événements sont deux à deux *incompatibles*. Comme neuf éventualités sont favorables à A, parmi 36 cas possibles : $P(A) = \frac{9}{36} = \frac{1}{4}$. Cette expérience aléatoire peut donc servir à effectuer un choix aléatoire parmi N, O, S, E en utilisant un codage (arbitraire) ; par exemple : $A \longleftrightarrow N$, $B \longleftrightarrow O$, $C \longleftrightarrow S$, $D \longleftrightarrow E$.

2.5 Le jeu de cartes

Dans un jeu de cartes, on distingue les cartes suivant leur *couleur* et leur *valeur*.

Il y a quatre couleurs : carreau (♦), coeur (♥), trèfle (♣), pique (♠). Les valeurs sont : as, roi, dame, valet, 10, 9, 8, 7, 6, 5, 4, 3, 2. Considérons l'expérience aléatoire qui consiste à choisir au hasard une carte dans un jeu *bien battu* de 52 cartes de sorte que l'hypothèse d'équiprobabilité (dite aussi : hypothèse de Laplace) est vérifiée. Distinguons les quatre événements suivants :

A la carte tirée est ♦
B la carte tirée est ♥
C la carte tirée est ♣
D la carte tirée est ♠

Alors $P(A) = P(B) = P(C) = P(D) = \frac{1}{4}$. Cette expérience aléatoire peut donc servir à opérer un choix aléatoire parmi N, O, S, E en utilisant un codage (arbitraire) ; par exemple : $A \longleftrightarrow N$, $B \longleftrightarrow O$, $C \longleftrightarrow S$, $D \longleftrightarrow E$.

Remarquons encore une fois qu'il y a $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$ codages différents.

3 Le modèle ensembliste de la théorie des probabilités

Nous venons de considérer un certain nombre d'*expériences aléatoires*. Pour décrire une telle expérience, nous avons d'abord introduit l'ensemble de tous les résultats possibles et imaginables de l'expérience ; ces résultats possibles sont appelés *épreuves* ou *éventualités* et leur ensemble, l'*univers des possibles*, est généralement noté Ω.

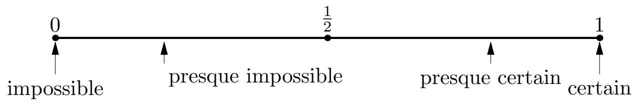
Dans les expériences précédentes, nous avons supposé que chaque éventualité avait la même chance d'apparaître, la même probabilité (les dés ne sont pas pipés, les pièces ne sont pas tordues, les cartes ne sont pas biseautées). On dit qu'on fait une *hypothèse d'équiprobabilité*. Cette hypothèse (qui est souvent une hypothèse de symétrie) conduit donc à associer à chaque éventualité la même probabilité ; en termes plus imagés : on répartit la probabilité 1 *uniformément* sur Ω.

Chaque fois que l'on effectue l'expérience aléatoire, c'est en général moins l'éventualité considérée individuellement qui nous intéresse que le fait qu'elle *réalise* ou non certains *événements* ; un événement est alors identifié, dans notre modèle ensembliste de la théorie des probabilités, à la partie de Ω de toutes les éventualités qui le réalisent.

Le tableau suivant indique la correspondance entre les deux terminologies ensembliste et probabiliste :

terminologie ensembliste	**terminologie probabiliste**
élément de Ω	éventualité ω
partie A de Ω	événement A
$\omega \in \Omega$	l'éventualité ω réalise l'événement A
la partie pleine de Ω	l'événement certain
la partie vide de Ω	l'événement impossible
le cardinal $\#A$ de A	le nombre de cas favorables à A
le cardinal $\#\Omega$ de Ω	le nombre de cas possibles

À chaque événement A de Ω, nous avons pu associer un nombre qui mesure son degré de possibilité, sa *probabilité* :



Cette probabilité se calcule par :

$$P(A) = \frac{\#A}{\#\Omega} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Par définition, la *probabilité uniforme P sur* Ω est l'application qui à chaque partie $A \subset \Omega$ associe le nombre $P(A)$ défini ci-dessus. Formellement :

$$\begin{array}{rrcl} P : & \mathcal{P}(\Omega) & \longrightarrow & [0;1] \\ & A & \longmapsto & \frac{\#A}{\#\Omega} \end{array}$$

Rappelons que $\mathcal{P}(\Omega)$ est l'ensemble de toutes les parties de Ω, donc l'ensemble de tous les événements associés à notre expérience aléatoire.

Quel est le cardinal de $\mathcal{P}(\Omega)$? Si le cardinal de Ω est n, alors le cardinal de $\mathcal{P}(\Omega)$ est 2^n. Pourquoi ?

4 Interprétation statistique d'une probabilité

Quel est l'intérêt « pratique » du nombre $P(A)$ associé à un événement A ? Répétons l'expérience aléatoire un grand nombre de fois et observons la réalisation ou la non-réalisation de A. Si l'on calcule le rapport

$$\frac{\text{nombre d'observations favorables à } A}{\text{nombre d'observations}}$$

nombre appelé : *fréquence relative* de l'événement A, alors il est un fait d'expérience que ce nombre est *en général* proche de $P(A)$. Exprimé autrement : $P(A)$ permet de prédire la fréquence relative des apparitions de A. Nous avons dit : *en général*, car il y a toujours des fluctuations dues au hasard. Mais si le nombre d'observations est grand, infiniment grand, la fréquence relative de A est presque sûrement $P(A)$. Cette observation, ce fait expérimental est parfois appelé : *la loi empirique des grands nombres.*

Comment la nature réalise-t-elle cette infinité de répétitions sur ses systèmes aléatoires, par exemple sur le système qui régit la longueur d'une feuille de chêne ? En en fabriquant une infinité de copies.

La relation entre fréquence relative et probabilité peut aussi être utilisée à des fins d'*estimation.*

Exemple Un petit étang contient un nombre N, inconnu, de poissons. Un biologiste, spécialiste de limnologie, essaie d'évaluer N de la façon suivante : il pêche dans différents endroits de l'étang ; il en sort 20 poissons qu'il marque et qu'il remet vivants dans l'étang. Quelques jours plus tard, le biologiste effectue une nouvelle pêche dans des endroits variés de l'étang. Avant de rejeter à l'eau les poissons pêchés, il note s'ils sont marqués ou non. Il prend ainsi 50 poissons dont 4 sont marqués. On suppose qu'entre les deux pêches, la population de l'étang n'a pas varié et que lors de la seconde pêche il y a équiprobabilité de sortie pour chacun des N poissons. En marquant les poissons, le biologiste a transformé l'étang en système aléatoire ; pêcher un poisson est une expérience aléatoire qu'il répète 50 fois. Le biologiste s'intéresse uniquement à l'événement A : le poisson est marqué. Parmi les 50 observations qu'il recueille, 4 sont favorables à l'événement A. D'après l'interprétation statistique, $P(A)$ est probablement voisin de $\frac{4}{50} = 8\%$. Les 20 poissons marqués représentent donc 8% de la population totale, ce qui implique que $N = 250$:

$$P(A) = \frac{\#A}{N} = \frac{20}{N} \approx \frac{4}{50} \Longrightarrow N \approx 250\,.$$

Les réflexions que l'on vient de faire appartiennent au domaine de la *statistique inductive.*

5 Enchaînements d'expériences aléatoires

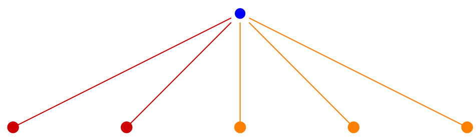
Beaucoup d'expériences aléatoires peuvent être considérées comme *répétitions* ou *successions* de plusieurs expériences plus simples, chaque expérience dépendant d'une certaine manière (qui reste à préciser) de la précédente.

5.1 Tirage avec remise

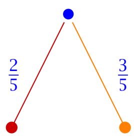
Considérons une urne contenant trois boules oranges et deux boules rouges.

L'expérience aléatoire consiste à tirer successivement, *avec remise* deux boules. Cette expérience est effectivement la répétition d'une même expérience : le tirage au hasard d'une boule dans une urne. La boule étant remise après le tirage, l'urne est ramenée à son état initial.

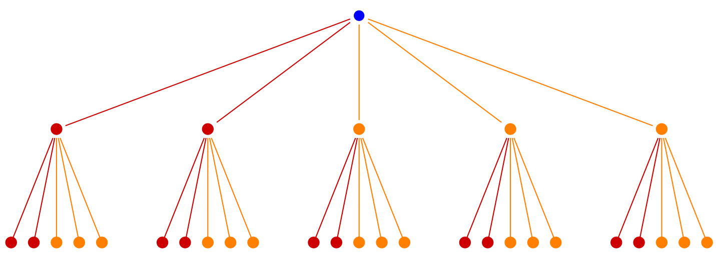
On représente souvent une telle situation par un *arbre*. Le premier tirage sera donc représenté par un arbre à cinq branches ; trois des feuilles sont oranges, deux sont rouges,



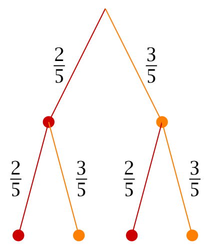
représentation que l'on préfère rendre plus compacte de la manière suivante :



Représentons de la même manière le deuxième tirage :



Ou encore



Il y a quatre chemins dans cet arbre ; ces chemins correspondent aux quatre éventualités 🔴🔴, 🔴🟠, 🟠🔴, 🟠🟠 de notre expérience.

Comment *calculer* la probabilité d'un tel chemin ? Calculons par exemple la probabilité du chemin correspondant au tirage 🟠🔴 : dans $\frac{3}{5}$ des cas, on aura une boule 🟠 lors du premier tirage, et dans $\frac{2}{5}$ de ces $\frac{3}{5}$ des cas, on aura ensuite une boule 🔴 lors du deuxième tirage.

La probabilité est $\frac{2}{5}$ de $\frac{3}{5}$, soit $\frac{2}{5} \otimes \frac{3}{5}$.

On peut donc, en généralisant, établir la règle suivante :

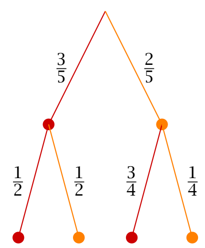
La probabilité d'un chemin est le produit des probabilités des chemins composants.

Soulignons que, dans le cas considéré, les deux expériences sont *indépendantes*.

Complétez le tableau suivant :

événement	🔴🔴	🔴🟠	🟠🔴	🟠🟠
probabilité		$\frac{6}{25}$		

5.2 Tirage sans remise



Reprenons l'urne de l'exemple précédent : on tire successivement, *sans remise*, deux boules.

Après la première expérience, le contenu de l'urne est modifié : si l'on a tiré une boule noire, la nouvelle urne contient deux boules noires et deux boules blanches ; si on a tiré une boule blanche, la nouvelle urne contient trois boules noires et une boule blanche.

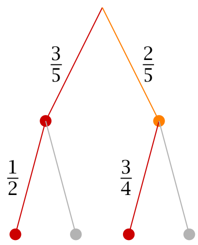
Il y a de nouveau quatre chemins, donc quatre événement s dont les probabilités se calculent comme auparavant : en multipliant les probabilités le long de ce chemin. Ceci conduit à la distribution suivante :

événement	🔴🔴	🔴🟠	🟠🔴	🟠🟠
probabilité	$\frac{3}{10}$			

Dans le cas du tirage sans remise, on dit que les deux expériences sont *dépendantes* : la deuxième expérience dépend du résultat de la première.

Étudions maintenant l'événement E suivant : la deuxième boule tirée est noire. Quelle est sa probabilité ?

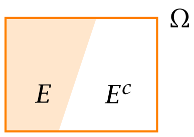
Remarquons que, dans notre arbre, il n'y a que deux chemins qui conduisent à une deuxième boule rouge :



La probabilité de E est alors la *somme* des probabilités des deux chemins : $\frac{1}{2} \cdot \frac{3}{5} \oplus \frac{3}{4} \cdot \frac{2}{5} = \frac{3}{5}$.

La probabilité de l'événement *contraire* ${}^cE$ (obtenir une boule blanche lors du deuxième tirage) peut être calculée de même manière analogue.

Mais il y a une deuxième manière de raisonner : du point de vue ensembliste, ${}^cE$ est le *complémentaire* de l'événement E.



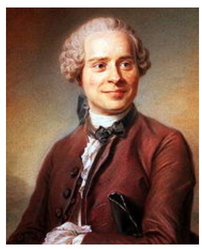
Mais alors il semble plausible de calculer de la manière suivante :

$$P({}^cE) = P(\Omega) - P(E) = 1 - \frac{3}{5} = \frac{2}{5}.$$

Remarquons enfin que notre expérience qui consiste en deux tirages sans remise produit des probabilités qui sont les mêmes que celles correspondant à un seul tirage.

Est-ce un hasard ?

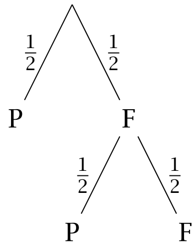
5.3 Croix et pile : l'erreur de d'Alembert



Considérons le jeu (l'expérience aléatoire) suivant : on lance une pièce de monnaie parfaitement symétrique ; si le côté « pile » est au-dessus, le jeu s'arrête et l'on a gagné ; si le côté « face » est au-dessus, on relance la pièce ; si le côté « pile » est au-dessus, le jeu s'arrête et l'on a gagné ; sinon, le jeu s'arrête et l'on a perdu. Quelle est la probabilité de gain ?

D'après d'Alembert (1754), « il y a trois cas possibles dont deux sont favorables ; la probabilité demandée vaut donc deux tiers. »

Voici l'arbre de cette expérience :



La probabilité de gain est donc $\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$.

Où est l'erreur dans le raisonnement de d'Alembert ?