1 Éléments de la théorie des probabilités

1.1 Expérience stochastique, événement aléatoire

Une expérience est dite *stochastique* ou *aléatoire* s'il est impossible de prévoir son résultat. En principe, on admet qu'une expérience stochastique peut être répétée indéfiniment dans des conditions identiques ; son résultat peut donc varier d'une réalisation à l'autre.

Exemples:

1. *on jette un dé et on observe le résultat obtenu ;*
2. *on jette un dé et on s'intéresse au nombre de jets qui sont nécessaires jusqu'à ce que la face 6 apparaisse pour la première fois ;*
3. *on mesure la durée de bon fonctionnement d'un équipement technique choisi au hasard parmi un grand nombre d'équipements identiques.*

Un ensemble Ω qui est constitué par tous les résultats possibles d'une expérience aléatoire est appelé un *espace d'échantillonnage* ou *espace fondamental*. L'ensemble fondamental peut être
– fini (premier exemple)
– infini dénombrable (deuxième exemple)
– infini non dénombrable (dernier exemple)
Un espace d'échantillonnage fini ou infini dénombrable est appelé espace d'échantillonnage discret ; s'il est infini non dénombrable, espace d'échantillonnage continu.

Un *événement* est un sous-ensemble A de l'espace d'échantillonnage Ω, c'est-à-dire qu'il est un ensemble des résultats possibles. Si en particulier A se réduit à un seul élément on parle d'*événement élémentaire*. Parmi les événements particuliers nous avons Ω lui-même, qui est l'événement *sûr* ou *certain* et l'ensemble vide $\emptyset$ qui est l'événement *impossible*.

1.2 Operations sur les événements

Dans la mesure où des événements sont des ensembles, il est clair que les propositions relatives aux événements peuvent être traduites dans le language de la théorie des ensembles et inversement. En particulier, nous avons une algèbre des événements qui correspond à l'algèbre des ensembles.

Ainsi si A et B sont des événements

- $A+B$ est l'événement soit A, soit B, soit les deux ;
- $A \cdot B$ est l'événement à la fois A et B ;
- $\overline{A}$ est l'événement non A ;
- $A - B$ est l'événement A mais pas B.

Si $A \cdot B = \emptyset$, alors A et B sont dits *mutuellement exclusifs* ou *incompatibles.* Cela signifie qu'ils ne peuvent pas être réalisés simultanément.

On dit qu'un événement A *implique* un autre événement B si chaque fois que A est réalisé B l'est aussi. On appelle *partition* de Ω toute famille finie ou dénombrable d'événements $A_1, A_2, \ldots, A_n, \ldots$, différents de $\emptyset$ et deux à deux incompatibles, telle que $A_1 + A_2 + \cdots + A_n + \cdots = \Omega$.

Les opérations sur les événements définies ci-dessus ont les propriétés suivantes, qui sont tout à fait analogues aux règles bien connues des opérations sur les ensembles :

$$\left.\begin{array}{lrcl} - & A\cdot(B+C) & = & (A\cdot B)+(A\cdot C) \\ - & A+(B\cdot C) & = & (A+B)\cdot(A+C) \end{array}\right\} \text{distributivité}$$

$$\left.\begin{array}{lrcl} - & \overline{A\cdot B} & = & \overline{A}+\overline{B} \\ - & \overline{A+B} & = & \overline{A}\cdot\overline{B} \end{array}\right\} \text{lois de Morgan}$$

$$- \quad \overline{\overline{A}} = A$$

Exemple:

- *Une unité de production utilise deux machines automatiques dont chacune peut tomber en panne au cours d'une période de temps donnée. On considère les événements suivants :*
  - A_1 *: la première machine fonctionne ;*
  - A_2 *: la deuxième machine fonctionne ;*
  - B_k *: exactement k machines fonctionnent* $(k = 0, 1, 2)$ *;*
  - C *: au moins une machine fonctionne.*

Alors :

$$B_0 = \overline{A}_1 \cdot \overline{A}_2; \qquad B_1 = (\overline{A}_1 \cdot A_2) + (A_1 \cdot \overline{A}_2)$$
$$B_2 = A_1 \cdot A_2 \quad \text{et} \quad C = A_1 + A_2$$

Problèmes résolus

1. On tire une carte au hasard d'un paquet de 52 cartes à jouer. Décrire l'espace fondamental dans les deux cas
   (a) on ne tient pas compte des couleurs,

(b) on tient compte des couleurs.

Si
- l'événement tirage d'un roi est A
- l'événement tirage d'un trèfle est B
- décrire les événements :

$$A + B,\ A \cdot B,\ A + \overline{B},\ \overline{A},\ \overline{A} + \overline{B},\ A - B,\ \overline{A} - \overline{B}.$$

2. Trois joueurs A, B et C, jettent une pièce à tour de rôle. Le premier qui obtient pile gagne. L'ensemble fondamental S de cette expérience peut être décrit comme suit :

$$S = \{1,\ 01,\ 001, \ldots,\ 000\ldots\}$$

   (a) Donner une interprétation des points de S.

   (b) Décrire les événements suivants en termes de ces points :
   - premier événement : A : A gagne
   - deuxième événement : B : B gagne
   - troisième événement : $(\overline{A + B})$.

   On admettra que A joue d'abord, puis B et enfin C.

3. Une ville de 100000 habitants compte trois journaux locaux : I, II, et III. Les proportions de lecteurs pour ces journaux sont :

I :	10%	I et II :	8%	et I, II et III : 1%
II :	30%	I et III :	2%	
III :	5%	II et III :	4%	

   Ces proportions nous indiquent par exemple que 8000 personnes lisent à la fois les journaux I et II.

   (a) Trouver le nombre de personnes ne lisant qu'un journal.

   (b) Combien de personnes lisent au moins deux journaux ?

   (c) II est un quotidient du soir, tandis que I et III sortent le matin. Combien de personnes lisent-elles au moins un journal du matin plus celui du soir ?

   (d) Combien de personnes lisent-elles un journal du matin seulement et le journal du soir ?

1.3 Probabilité

1.3.1 Loi empirique des grands nombres

Si nous observons la fréquence de réalisation d'un événement A au cours d'une longue suite de répétitions de l'expérience Ω à laquelle il est lié, nous constaterons expérimentalement que cette fréquence N_A/N fluctue de moins en moins lorsque le nombre N de répétitions de l'expérience augmente. Ce résultat expérimental que nous appellerons la *loi empirique des grands nombres* conduit donc à abstraire de l'expérience la notion de *probabilité statistique* d'un événement A : cette probabilité statistique de A qui est la limite de la fréquence N_A/N lorsque N augmente indéfiniment, correspond certainement à l'intuition première que nous avons de la notion de probabilité.

Voici encore un autre aspect de cette loi empirique des grands nombres : si nous observons plusieurs longues suites de répétitions d'une expérience Ω, nous pourrons constater expérimentalement que les fréquences d'un événement A lié à Ω calculées sur les diverses suites sont approximativement les mêmes.

Les fréquences $F(A) = N_A/N$ de même que leurs valeurs asymptotiques possèdent les propriétés évidentes suivantes :

1. pour tout $A : F(A) \geq 0$ et $F(\Omega) = 1$,
2. $F(\sum_1^n A_j) = \sum_1^n F(A_j)$ pour toute suite finie d'événements deux à deux incompatibles.

Nous retiendrons ces propriétés (en étendant la seconde à des familles dénombrables afin de permettre des passages à la limite naturels) pour définir *axiomatiquement* les probabilités sans en faire reposer la définition sur la loi empirique des grands nombres. La théorie mathématique rigoureuse qui sera ensuite développée s'est avérée à même de rendre compte des phénomènes aléatoires.

1.3.2 Espace probabilité

Le but du calcul des probabilités est d'associer une mesure à certains événements pour caracteriser leur fréquence. Comment définir la probabilité d'un événement ? Nous voulons qu'elle soit d'autant plus grande que A soit plus probable. À chaque événement de Ω nous associons un nombre réel $P(A)$. La valeur $P(A)$ est une mesure de chance de réalisation de l'événement A lors de l'expérience stochastique considérée. Une *loi de probabilité* ou une *probabilité* sur un ensemble fondamental Ω est une fonction P définie sur les sous-ensembles de Ω de telle sorte que les *axiomes* suivants soient satisfaits :

- (1) $0 \leq P(A) \leq 1$ pour tout A
- (2) $P(\Omega) = 1$
- (3) Pour tout nombre d'événement exclusifs $A_1, A_2, \ldots$ $P(A_1 + A_2 + \ldots) = P(A_1) + P(A_2) + \ldots$

1.3.3 Propriétés élémentaires d'une loi de probabilité

On peut aisément démontrer d'autres propriétés importantes de P.

- $P(\overline{A}) = 1 - P(A)$ (règle du complément), donc $P(\emptyset) = 0$
- $P(A) < P(B)$ si A implique B (règle d'inclusion) et $P(A-B) = P(A)-P(B)$.
- $P(A+B) = P(A) + P(B) - P(AB)$ (règle d'additivité) si A et B sont deux événements quelconques.

Plus généralement si A_1, A_2, A_3 sont trois événements quelconques

$$\begin{aligned} P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cdot A_2) - P(A_1 \cdot A_3) - \\ - P(A_2 \cdot A_3) + P(A_1 \cdot A_2 \cdot A_3) \end{aligned}$$

il est possible de généraliser à n événements.

Remarque: le fait, que la probabilité d'un événement est égal 1 ou 0 n'implique pas nécessairement que cet événement est certain ou impossible.

1.4 Évaluation d'une loi de probabilité

Les axiomes introduits précédemment ne permettent pas d'évaluer numériquement la fonction P associée à une expérience stochastique donnée.

On peut se faire d'une loi de probabilité l'image intuitive, si la loi définit une façon de répartir sur l'ensemble Ω une masse unité. Si $P(A)$ désigne la masse portée par A, on voit que (les axiomes) sont alors remplis.

1.4.1 Loi uniforme discrète. Analyse combinatoire

Pour une classe importante d'expériences stochastiques on peut raisonnablement admettre l'existence de certaines propriétés de symétrie. Soit $\Omega = \{e_1, e_2, \ldots, e_N\}$ un ensemble fini, composé d'un nombre fini N d'éléments. N est appelé *le cardinal* ($|\Omega|$) ou *l'effectif* de Ω. Les propriétés de symétrie impliquent que tous les résultats possibles de l'expérience ont la même chance d'apparition, en d'autres termes qu'ils sont équiprobables. Alors : $P(e_1) = P(e_2) = \cdots = P(e_N) = \dfrac{1}{N}$ où e_i est un événement élémentaire, et à un événement A constitué de k événements élémentaires la probabilité : $P(A) = \dfrac{|A|}{|\Omega|} = \dfrac{k}{N}$ (fréquence de A). Il est facile de vérifier que la fonction $P(A)$ ainsi définie obéit aux axiomes. Le calcul de $|A|$ fait souvent appel à *l'analyse combinatoire*. Notons que le concept de la loi de probabilité uniforme a longtemps servi de définir la notion de probabilité, ce qui se traduit par la formule bien connue :

$$P(A) = \frac{n_A}{n}$$

où n est le nombre de 'cas possible' (lors de la réalisation d'une expérience stochastique) et n_A est le nombre de 'cas favorables' (favorables à la réalisation de l'événement A).

Exemple: *Si on jette deux fois un dé équilibré, quelle est la probabilité que la somme des deux faces soit au moins égale à 10 ?*

Réponse: On a ici affaire à un espace probabilisé comprenant 36 événements élémentaires, à savoir tous les couples (i, j) où $1 \leq i \leq 6$, $1 \leq j \leq 6$. Si on définit une loi uniforme sur Ω, il en résulte que

$$P \text{ (somme de 2 faces} \geq 10) = 6/36 = \frac{1}{6}.$$

1.4.2 Analyse combinatoire

Lorsqu'on a affaire à une loi uniforme discrète, la probabilité d'un événement A, $P(A)$, est déterminée par le quotient des nombres d'éléments des ensembles A et Ω. Si ces nombres sont trop grands pour être comptés directement, on fait appel à l'analyse combinatoire.

- Soit un ensemble de n objets tous différents. On appelle *arrangements* sans répétition de ces n objets k à k les listes de k objets, rangés dans un ordre déterminé choisis parmi ces n objets. Le nombre d'arrangements sans répétition de n éléments pris k à k est : $A_n^k = \dfrac{n!}{(n-k)!}$.
- Soit un ensemble de n objets tous différents. Chaque liste de ces n objets rangé dans un ordre déterminé est une *permutation* sans répétition de ces n objets. Le nombre des permutations de n éléments est $p_n = n!$.
- Soit un ensemble de n objets tous différents. On appelle combinaisons sans répétition de ces n objets k à k les ensembles (les collections) de k objets choisis parmi ces n objets.
  Le nombre des combinaisons sans répétitions de n éléments k à k est :
  $C_n^k = \binom{n}{k} = \dfrac{n!}{k!(n-k)!}$
- Le nombre des permutations de n éléments avec $n_1, n_2, \ldots, n_k$ répétitions est :
  $p_n^{n_1, n_2, \ldots, n_k} = \dfrac{n!}{n_1! n_2! \ldots n_k!}$.

Le nombre des arrangements avec répétitions de n éléments k à k est : $A_n^{*k} = n^k$.
Le nombre des combinaisons avec répétitions de n éléments k à k est :

$$C_n^{*k} = \binom{n+k-1}{k}.$$

On peut résoudre de nombreux problèmes liés à la physique statistique en utilisant par l'analyse combinatoire. En voilà trois examples. Tous les trois problèmes peuvent être dérivés de l'arrangement de n billes dans N boîtes dans des conditions différentes.

Des systèmes divers des molécules de gaz peuvent être appliqués dans la statistique de Maxwell–Boltzmann.

Exemple: *Nous avons n particules toutes différentes. Mettons cettes n particules sur N niveaux. Quelle est la probabilité qu'exactement k particules soient sur un niveau donné ?*

Réponse: Utilisons nos connaissances de la théorie de la probabilité classique.

Nombre totale des cas : N^n, soit le nombre des arrangements avec répétition de N objects n à n. (Selon l'hypothèse de statistique de Maxwell–Boltzmann toutes les positions sont équiprobables). Nombre des événements favorables $\binom{n}{k} \cdot (N-1)^{n-k}$, car de n objects nous pouvons choisir à $\binom{n}{k}$ façons et nous pouvons arranger la reste $(n-k)$ objets à $(N-1)^{n-k}$ façons sur $(N-1)$ niveaux. Donc la probabilité cherchée est $p_k = \binom{n}{k} \cdot (N-1)^{n-k}/N^n$. Dans le cas des autres particules comme p.e. les photons, nous applicons le statistique Bose-Einstein. Dans ce model on ne peut pas distinguer les particules.

Exemple: *Mettons n particules non distinguables sur N niveaux de l'énergie. Quelle est la probabilité que exactement k particules soient à un niveaux donné.*

Réponse: Nombre de tous les cas : $\binom{N+n-1}{n}$ soit le nombre des combinations avec répétition de N éléments à n.

Les différents arrangements sont équiprobables. Le nombre des cas favorables est égale au nombre des arrangements de la reste $n-k$ particules sur les $N-1$ niveaux.

Nous pouvons calculer le nombre des combinations avec répétition de nouveau :

$$p_k = \frac{\binom{N+n-k-2}{n-k}}{\binom{N+n-1}{n}}.$$

Le model Bose-Einstein n'est pas valable généralement. Dans le cas de certains particules qu'on ne peut pas distinguer il fait aussi prendre en consideration le principe de Pauli, c'est à dire au maximum une particule occupe un niveau. Ce phenomène est bien décrit par la statistique Fermi–Dirac.

Exemple: *Mettons n particules non distinguables sur N niveaux d'une telle façon qu'il y ait une particule au maximum à chaque niveau. Quelle est la probabilité qu'il y ait exactement une particule sur un niveau choisi au hasard ?*

Réponse: Nombre totale des cas : $\binom{N}{n}$ soit le nombre des choix de n niveaux où nous mettons les particules. Nombre des cas favorables $\binom{N-1}{n-1}$, le nombre de possibilité d'arranger la reste $n-1$ particules sur les $N-1$ niveaux. Donc la probabilité cherchée $p = \frac{\binom{N-1}{n-1}}{\binom{N}{n}}$.

Exemple: *Un emballage comprend six articles dont on sait qu'exactement deux sont défectueux. On prélève la moitié des articles sans remettre un article déjà choisi. Quelle est la probabilié que cet échantillon de taille trois comprenne au moins deux bons articles ?*

Réponse: La probabilité de cet événement est donnée par $P(A) = P(A_2) + P(A_3)$ où $A_k = \{\text{exactement } k \text{ bons articles}\}$. $A = A_2 + A_3$ et A_2, A_3 sont mutuellement exclusifs. D'après le modèle de la loi uniforme discrète, il vient

$$P(A) = \frac{\binom{4}{2}\binom{2}{1}}{\binom{6}{3}} + \frac{\binom{4}{3}\binom{2}{0}}{\binom{6}{3}} = \frac{4}{5}$$

1.4.3 Lois uniformes continues

La notion de loi uniforme s'étend au cas d'un ensemble fondamental infini non dénombrable. Si l'on maintient l'hypothèse exigeant que chaque événement élémentaire ait la même probabilité de réalisation, celle-ci ne peut qu'être nulle en vertu de l'axiome, bien que chaque événement élémentaire peut clairement être réalisé. Pour définir $P(A)$, on ne peut donc plus utiliser les événements élémentaires dont A est composé.

On peut encore définir une loi uniforme par exemple dans les 3 cas suivants :

- Ω est une courbe de longueur finie (par exemple un cercle). Pour $A \subset \Omega$, on note $\mu(A)$ la longueur de A.
- Ω est une surface d'aire finie (par exemple une sphère). Pour $A \subset \Omega$, on note $\mu(A)$ l'aire de A.
- Ω est une région de l'espace de volume fini (par exemple un cube). Pour $A \subset \Omega$, on note $\mu(A)$ le volume de A.

En considérant Ω comme un solide homogène (c'est-à-dire de densité constante) de masse 1, la loi de probabilité correspondante P est dite la loi uniforme sur Ω. La probabilité $P(A)$ ne dépend ni de la position de A, ni de sa forme, mais seulement de sa grandeur.

On a : $P(A) = \dfrac{\mu(A)}{\mu(\Omega)}$.

En particulier, si A n'a qu'un nombre fini de points, $P(A) = 0$.

C'est à cette loi uniforme que l'on fait allusion lorsque l'on parle, sans préciser, d'un point choisi au hasard dans Ω.

Exemple: *Ω étant une boule de centre 0 et de rayon R (ensemble des points de l'espace dont la distance à 0 est inférieure à R), on choisi au hasard un point M dans Ω. Indiquer, selon les diverses valeurs du nombre x, la probabilité $P(OM < x)$.*

Réponse:

$$\begin{array}{lll} P(OM < x) = 0 & \text{si} & x \leq 0, \\ P(OM < x) = \left(\dfrac{x}{R}\right)^3 & \text{si} & 0 \leq x \leq R \\ P(OM < x) = 1 & \text{si} & R \leq x. \end{array}$$

Problème: Ω étant un disque de centre 0, on choisi un point M au hasard sur Ω et on construit le carré de centre 0 ayant M pour milieu d'un des côtés.

Quelle est la probabilité pour que ce carré soit contenu dans le disque Ω ?