4 Probabilités - définitions élémentaires

4.1 expériences aléatoires

On appelle **expérience aléatoire** une expérience dont les issues (les résultats) ne sont pas déterminés à l'avance.

L'ensemble, souvent noté Ω, de toutes les issues possibles est appelé **univers** ou **espace d'échantillonnage** de l'expérience.

exemples :

- On jette un dé à six faces, il y a six issues possibles : $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Un fabricant contrôle les produits sortis de ses chaînes : il y a deux issues possibles, ou bien le produit est sans défaut et peut être vendu, ou bien le produit présente des défauts et va être jeté : $\Omega = \{$ conforme , non conforme $\}$.
- On choisit un nombre entier positif : $\Omega = \mathbb{N}$. A la différence des exemples précédents, Ω est ici infini. On parle là d'infini discret (les valeurs possibles sont toutes isolées).
- On choisit un point dans le plan. Là, $\Omega = \mathbb{R}^2$, et l'univers est aussi infini, mais cette fois-ci on parle d'infini continu.

4.2 événements

Un sous-ensemble, ou partie, de Ω est appelé un **événement**. L'ensemble des événements est donc l'ensemble noté $\mathcal{P}(\Omega)$ des parties de Ω.

En particulier Ω et $\emptyset$ sont appelés **événement certain** et **événement impossible**.

Un ensemble qui ne contient qu'une seule issue est un **événement élémentaire**.

exemple : dans l'expérience du dé, « on obtient 1 » est un événement élémentaire, « on obtient un nombre impair » ou « on obtient un nombre inférieur ou égal à 4 » sont deux événements (non élémentaires).

4.3 opérations sur les événements

A et B sont deux événements. Alors :

L'événement contraire de A est son complémentaire dans Ω, noté $\bar{A}$ ou $\Omega - A$, et se comprend « A n'est pas réalisé ».

La réunion de A et B est $A \cup B$ et se comprend « A ou B (ou les deux) sont réalisés ».

L'intersection de A et B est $A \cap B$ et se comprend « A et B sont réalisés simultanément ».

exemple : Dans l'expérience du dé, si $A = \{1, 3, 5\}$ = « on obtient un nombre impair » et $B = \{1, 2, 3, 4\}$ = « on obtient un nombre inférieur ou égal à 4 », alors $\bar{A} = \{2, 4, 6\}$ =« on obtient un nombre pair », $\bar{B} = \{5, 6\}$ =« on obtient un nombre strictement supérieur à 4 », $A \cup B = \{1, 2, 3, 4, 5\}$ = « on obtient un nombre impair ou un nombre inférieur ou égal à 4 », et $A \cap B = \{1, 3\}$ = « on obtient un nombre impair inférieur ou égal à 4 ».

Deux événements sont **incompatibles** s'ils ne peuvent se produire simultanément, i.e si leur intersection $A \cap B$ est vide.

Bien sûr, un événement et son contraire sont toujours incompatibles.

4.4 loi de probabilité

On peut associer à une expérience aléatoire et à son univers une **probabilité** qui permet de quantifier le fait qu'un événement est « probable » ou est « peu probable ».

Une probabilité est une application p de $\mathcal{P}(\Omega)$ dans $[0; 1]$ telle que $p(\Omega) = 1$ et telle que si A et B sont deux événements incompatibles, $p(A \cup B) = p(A) + p(B)$.

Comme conséquence de cette définition, on a donc les propriétés suivantes :

$$0 \leq p(A) \leq 1 \text{ pour tout événement } A$$

$$p(\emptyset) = 0, \quad p(\Omega) = 1$$

$$p(\bar{A}) = 1 - p(A) \text{ pour tout événement } A$$

$$p(A \cup B) + p(A \cap B) = p(A) + p(B) \text{ pour tous événements } A, B.$$

remarque : pour chaque univers, on peut bien sûr imaginer plusieurs lois de probabilités différentes. Dans l'expérience de la fabrication d'objets, on peut imaginer qu'un chaîne de fabrication fonctionnant bien ait pour probabilité $p($« conforme »$) = 0,95$ et $p($« non conforme »$) = 0,05$, alors qu'une chaîne moins efficace ait des probabilités associées $p($« conforme »$) = 0,75$ et $p($« non conforme »$) = 0,25$.

On doit donc définir non seulement l'univers Ω, mais aussi la loi de probabilité p dont on le munit. Pour être rigoureux on parle donc en toute rigueur d'un **espace probabilisé** (Ω, p).

4.5 le cas particulier des univers finis

Pour étudier un phénomène à l'aide des probabilités, on a besoin de connaître la loi de probabilité p, qui est une fonction de $\mathcal{P}(\Omega)$ dans $[0; 1]$, donc a priori on a besoin de connaître sa valeur sur chaque sous-ensemble de Ω. Mais en fait, quand Ω est fini, la connaissance de p sur chaque événement élémentaire suffit : si $A \subset \Omega$ est un événement quelconque, A est fini et on peut écrire $A = \{a_1, a_2, \ldots, a_k\}$, donc $p(A) = p(\{a_1\}) + p(\{a_2\}) + \ldots + p(\{a_k\})$.

Un cas encore plus particulier mais fondamental est le cas de l'équiprobabilité : sur un univers fini, on dit que la loi est **équiprobable** si tous les événements élémentaires ont la même probabilité. Dans ce cas, la probabilité de chaque événement élémentaire est simplement $1/\text{card}(\Omega)$, et la probabilité d'un événement A est :

$$\boxed{p(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{\text{nombre d'éléments de } A}{\text{nombre d'éléments de } \Omega}.}$$

exemples : dans des expérience de tirage au sort (pile ou face, dé, jeu de cartes, ...), sans précisions supplémentaires on supposera que le jeu n'est pas truqué, ce qui revient à dire que la loi est équiprobable : tous les événements élémentaires ont la même probabilité (une chance sur deux de faire pile, une chance sur deux de faire face ; une chance sur six de tirer 1, une chance sur six de tirer 2, etc... ; une chance sur 32 de tirer chacune des cartes du paquet).

4.6 le cas particulier des probabilités infinies discrètes

Ω infini est dit **discret** si on peut énumérer ses éléments, i.e si on peut écrire $\Omega = \{x_1, x_2, \ldots, \}$. Typiquement, cela correspond à des expériences dont le résultat est un entier naturel. Comme dans le cas précédent, on obtient la probabilité d'un événement quelconque comme somme (éventuellement infinie) des événements élémentaires qui le composent.

exemple : On considère la probabilité de désintégration des atomes d'un composé radioactif durant un intervalle de temps de longueur t fixé. Ici, $\Omega = \mathbb{N}$, et on montrera en exercice que $p_n = \dfrac{\Lambda^n t^n}{n!} e^{-\Lambda t}$. Quelle est la probabilité d'obtenir moins de 5 désintégrations ?

remarque : on n'a jamais équiprobabilité sur un ensemble infini discret. En effet, supposons les probabilités élémentaires p_i ($i \in \mathbb{N}$) toutes égales à un même α. Alors $p(\Omega) = \sum_{i=0}^{+\infty} \alpha$ ne peut valoir que 0 (si $\alpha = 0$) ou $+\infty$ (si $\alpha > 0$), mais en aucun cas 1.

4.7 le cas particulier des probabilités continues

Pour étudier les probabilités sur des univers continus infinis (par exemple : choix d'un nombre au hasard dans $[0; 1]$; durée de vie d'une voiture dans $[0; +\infty[$, ...) on va comme dans le cas fini partir d' « événements de base » qui permettent de reconstituer tous les événements, donc de calculer toutes les propriétés. Mais ici le problème est un peu plus délicat. En effet, en général, avec un univers continu la probabilité de chaque événement élémentaire est nulle, et cette information ne permet pas de déterminer la valeur de $p(A)$ pour tout événement A.

Pour ce qui suit on prend pour Ω un intervalle de $\mathbb{R}$ (par exemple, $[0; 1]$, ou $[0; +\infty[$, ou $\mathbb{R}$ lui-même...).

Ces événements de base vont ici être les segments $[a, b]$. Dans la plupart des cas, les événements qui nous intéressent pourront être décrits comme réunion, intersection, complémentaires, ... de segments et on pourra donc déduire ainsi leur probabilité de celles de ces segments grâce aux règles de calcul sur les probabilités.

Donc dans le cas des probabilités continues on associe à chaque probabilité une **densité** de probabilité qui est une fonction intégrable et positive, telle que $\displaystyle\int_\Omega f = 1$. Et la probabilité p est caractérisée par le fait que pour tout événement A, $p(A) = \displaystyle\int_A f$. En particulier, $p([a; b]) = \displaystyle\int_a^b f$ pour tout segment $[a; b]$.

exemple : le cas le plus simple est celui de la probabilité uniforme sur $[0; 1]$, qui correspond à l'expérience « on choisit au hasard un nombre compris entre 0 et 1, sans privilégier aucune valeur ».

Alors la densité correspondante est $f = 1$, et la probabilité d'obtenir un nombre entre a et b (pour $0 \leq a \leq b \leq 1$) est égale à $p([a, b]) = \displaystyle\int_a^b 1 = b - a$.

Ainsi, avec $a = 0$ et $b = 1$, la probabilité est 1 : le choix d'un nombre entre 0 et 1 donne à coup sûr un nombre entre 0 et 1 !

Au contraire, si $a = b$, on constate que la probabilité de choisir un nombre a donné à l'avance est nulle.

Si $a = 0,25$ et $b = 0,75$: on a une chance sur deux que le nombre choisi soit dans l'intervalle $[a; b]$ de longueur $1/2$.

4.8 probabilités conditionnelles

Soit (Ω, p) un espace probabilisé, et A un événement de probabilité non nulle.

On appelle « probabilité que B soit réalisé sachant que A l'est », ou plus simplement « probabilité de B sachant A », la quantité

$$\boxed{p(B|A) = \frac{p(A \cap B)}{p(A)}.}$$

Ainsi $p(\Omega|A)$ est toujours égal à 1.

exemple : on lance deux dés bien équilibrés. Quelle est la probabilité que la somme des résultats soit strictement supérieure à 10 sachant que l'un des dés a donné 6.

"somme > 10" ={$(6,6);(6,5);(5,6)$}, "l'un des dés donne 6" est de cardinal 11, et l'intersection est de cardinal 3, donc la probabilité est $3/11$.

Connaissant $p(A|B)$, on aimerait parfois connaître $p(B|A)$. C'est souvent possible en écrivant de deux manières différentes $p(A \cap B)$ à l'aide des définitions de $p(A|B)$ et de $p(B|A)$:

$$\boxed{p(A \cap B) = p(A)p(B|A) = p(B)p(A|B).}$$

exemple : 48 des 53 étudiants de TI ont eu la moyenne en automatique, et 14 des 26 étudiants $MCPC$. Quelle est la probabilité qu'un étudiant ayant eu la moyenne soit en TI ?

On a $p(S|MCPC) = 14/26$ et $p(S|TI) = 48/53$. De plus $p(S) = 62/79$ et $p(TI \cap S) = 48/79$. Donc $p(TI|S) = (48/79)/p(S) = (48/79) \times (79/62) = 48/62 = 24/31$.

La formule ci-dessus peut s'exprimer sous la forme plus directement utilisable suivante :

$$\boxed{p(A|B) = \frac{p(B|A) \times p(A)}{p(B)}.}$$

Dans les cas un peu plus compliqués, on peut avoir besoin de la formule de Bayes.

Considérons donc des événements incompatibles A_1, A_2, ..., A_n, et un événement B qui ne peut se produire que si l'un des A_i se produit, les $p(B|A_i)$ étant connus. On cherche la probabilité pour que, B s'étant produit, A_k en soit la cause.

Commençons par remarquer que $p(B) = p(A_1 \cap B) + \ldots + p(A_n \cap B)$; comme $p(A_k \cap B) = p(A_k)p(B|A_k)$, on obtient la formule des **probabilités totales :**

$$\boxed{p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \ldots + p(B|A_n)p(A_n).}$$

Alors en écrivant $p(A_k|B) = p(A_k \cap B)/p(B) = p(B|A_k)p(A_k)/p(B)$, et en remplaçant $p(B)$ par la formule précédente, on obtient la **formule de Bayes :**

$$\boxed{p(A_k|B) = \frac{p(A_k)p(B|A_k)}{\sum_{i=1}^{n} p(A_i)p(B|A_i)}}$$

exemple : Un test de dépistage d'une maladie rare touchant une personne sur 10000 semble efficace : il détecte 99% des personnes infectées, avec seulement $0,5\%$ de « faux positifs ». Quelle est la probabilité qu'une personne dont le test est positif soit effectivement malade ? $p(M|P) = p(P|M)p(M)/(p(P|M)p(M) + p(P|\bar{M})p(\bar{M})) \simeq 1,94\%$.

4.9 événements indépendants

On dit que deux événements A et B sont **indépendants** quand l'un des deux est de probabilité nulle, ou bien, quand les deux sont de probabilité non nulle, si le fait de savoir que l'un est réalisé n'influe pas sur la probabilité que l'autre le soit. Autrement dit deux événements de probabilité non nulle sont indépendants quand $p(B|A) = p(B)$ (ou de manière équivalente quand $p(A|B) = p(A)$).

Comme $p(B|A) = \dfrac{p(A \cap B)}{p(A)}$, cela équivaut à la

proposition : deux événements sont indépendants si et seulement si $p(A \cap B) = p(A)p(B)$.

remarque : ne pas confondre les deux notions d'événements indépendants et d'événements incompatibles ! Deux événements incompatibles ne sont **jamais** indépendants (sauf si les deux sont de probabilités nulle). En effet, si A et B sont incompatibles et que l'on sait que A est réalisé, justement, B ne peut pas se produire...il n'y a donc pas indépendance.