Avant – Propos

Ce cours de **Probabilités** est destiné surtout aux étudiants de L2 Mathématiques (dans le cadre du système L.M.D.), il couvre le programme officiel. Le contenu de ce cours est consacrée au programme du module Probabilités, à savoir :

1) Rappels sur les probabilités.

2) Variables aléatoires à une dimension.

3) Lois de probabilités absolument continues usuelles.

J'espère que ce support aidera les étudiants de deuxième année Mathématiques à assimiler les Probabilités et plus particulièrement les variables aléatoires. Enfin, des erreurs peuvent être relevées, prière de les signaler à l'auteur.

Table des matières

Introduction aux Probabilités

Les probabilités sont une branche des Mathématiques dont l'objet est l'étude des phénomènes aléatoires. Historiquement, il s'agissait essentiellement des jeux de hasard et des problèmes d'espérance de vie pour des calculs de rente. L'approche mathématique de ces problèmes n'apparaît réellement qu'au 17^e siècle avec Pascal et Fermat, puis au siècle suivant avec Bayes et Laplace.
L'élaboration d'un cadre mathématique rigoureux est très récente et elle est due à Kolmogorov, qui a axiomatisé le calcul des probabilités (fondements du calcul des probabilités, 1933) et a permis en particulier l'utilisation de la théorie de la mesure.

Avant de passer à la théorie des probabilités, nous présentons quelques principes de base de l'analyse combinatoire, qui sont extrêmement utiles pour le calcul des probabilités.

0.1 Analyse Combinatoire

Il y a, bien des problèmes en théorie des probabilités qui peuvent être résolus simplement en comptant le nombre de manières différentes selon lesquelles un certain événement peut se réaliser. Par convention on appelle analyse combinatoire la théorie mathématique du dénombrement.

0.1.1 Principe Fondamental de dénombrement

Ce principe de dénombrement (ci-dessous théorème 0.1.1) sera essentiel par la suite. Il établit en gros que si une expérience peut produire m résultats et une autre n, alors il y a $m \times n$ résultats possibles lorsqu'on considère ces deux expériences ensemble.

Théorème 0.1.1. *Supposons qu'il faille réaliser deux expériences. Si l'expérience **1** peut produire l'un quelconque de m résultats et si, pour chacun d'entre eux, il y a n résultats possibles pour l'expérience **2**, alors il existe mn résultats pour les deux expériences prises ensemble.*

Exemple 0.1.1. *Une petite communauté se compose de dix hommes et de leurs fils, chaque homme ayant trois fils. Si un homme et l'un de ses fils doivent être désignés "père et fils exemplaires", combien y a-t-il de choix différents possibles ?*

En considérant le choix du père comme la première expérience et ensuite le choix de l'un de ses fils comme la seconde, nous conclurons d'après le principe fondamental qu'il y a $10.3 = 30$ *choix possibles.*

0.1.2 Principe fondamental généralisé

Lorsqu'il y a plus de deux expériences à réaliser, le principe fondamental peut être généralisé comme suit :

Théorème 0.1.2. *Si* r *expériences doivent être réalisées et sont telles que la première peut produire l'un quelconque de* n_1 *résultats, et si pour chacun d'entre eux il y a* n_2 *résultats possibles pour la* 2^e *expérience, et si pour chaque résultat des deux premières expériences il y en a* n_3 *pour la* 3^e *expérience, et ainsi de suite, il y aura alors au total* $n_1.n_2....n_r$ *résultats pour les* r *expériences prises ensemble.*

Exemple 0.1.2. *Le comité de planification d'un collège est constitué de* 3 *étudiants de première année,* 4 *de deuxième,* 5 *de troisième et* 2 *de dernière année. Un sous-comité de* 4 *étudiants comportant un représentant de chaque classe doit être choisi. Combien peut-on former de sous-comités ?*

Nous pouvons considérer le choix d'un sous-comité comme le résultat combiné de 4 *expériences distinctes, chacune consistant à choisir un unique représentant dans l'une des classes. Par conséquent, en application de la version généralisée du principe fondamental, il y a* $3.4.5.2 = 120$ *sous-comités possibles.*

0.1.3 Permutations

Permutations d'objets distinguables

Combien existe-t-il d'un arrangement et ordonné des lettres a, b et c. Par énumération directe nous en trouvons 6, à savoir : abc, acb, bac, bca, cab et cba. Chaque arrangement est par convention appelé permutation. Il y a ainsi 6 permutations possibles des éléments d'un ensemble de 3 objets. Ce résultat aurait également pu être construit à partir du principe fondamental ; la première lettre de la permutation peut être n'importe laquelle des 3, la deuxième lettre peut ensuite être choisie parmi les 2 restantes tandis que la troisième ne peut plus faire l'objet d'aucun choix. Ainsi, il y a $3 \times 2 \times 1 = 6$ permutations possibles.

Définition 0.1.1. *Une permutation d'un ensemble* X *est une bijection de* X *sur lui-même. Notamment, une permutation de* n *éléments (où* n *est un entier naturel) est une bijection d'un ensemble fini de cardinal* n *sur lui-même.*

Supposons maintenant que nous ayons n objets. Un raisonnement analogue à celui que nous venons d'utiliser ci-dessus établit le théorème suivant :

Théorème 0.1.3. *Le nombre de permutations de n objets est $n!$.*

Exemple 0.1.3. *Un cours de théorie des probabilités est suivi par 6 étudiants(garçons) et 4 étudiantes(filles). Un examen a lieu, puis les étudiants sont classés selon leur note. On suppose exclu que deux étudiants obtiennent la même note.*
• Combien de classements peut-on avoir ?
• Si les garçons sont classés entre eux uniquement et les filles entre elles, combien de classements globaux peut-on avoir ?

• Comme chaque classement correspond à un certain arrangement ordonné de 10 personnes, on voit que la réponse à cette partie du problème est $10! = 3628800$.

• Comme il y a $6!$ classements des garçons entre eux et $4!$ classements des filles entre elles, il résulte par application du principe fondamental qu'il y aura dans ce cas $(6!)(4!) = (720)(24) = 17280$ classements possibles.

Permutations d'objets partiellement indistinguables

Nous allons maintenant nous attacher à déterminer le nombre de permutations dans un ensemble de n objets quand certains de ces objets sont indistinguables les uns des autres. Pour mieux saisir de quoi il s'agit, considérons l'exemple suivant :

Exemple 0.1.4. *Combien d'arrangements différents peut-on former avec les lettres P E P P E R ?*

On remarquera d'abord qu'il existe $6!$ permutations des lettres P_1 $E1$ P_2 P_3 E_2 R lorsque les trois P et les deux E sont distingués les uns des autres. Cependant, considérons l'une quelconque de ces permutations $-$ P_1 P_2 E_1 P_3 E_2 R $-$ par exemple.
Si nous permutons les P entre eux et les E entre eux, l'arrangement résultant sera encore de la même forme P P E P E R. En fait, chacune des $3!$ $2!$ permutations

$$\begin{array}{cccccc@{\qquad}cccccc}
P_1 & P_2 & E_1 & P_3 & E_2 & R & P_1 & P_2 & E_1 & P_3 & E_1 & R \\
P_1 & P_3 & E_1 & P_2 & E_2 & R & P_1 & P_3 & E_2 & P_2 & E_1 & R \\
P_2 & P_1 & E_1 & P_3 & E_2 & R & P_2 & P_1 & E_2 & P_3 & E_1 & R \\
P_2 & P_3 & E_1 & P_1 & E_2 & R & P_2 & P_3 & E_2 & P_1 & E_1 & R \\
P_3 & P_1 & E_1 & P_2 & E_2 & R & P_3 & P_1 & E_2 & P_2 & E_1 & R \\
P_3 & P_2 & E_1 & P_1 & E_2 & R & P_3 & P_2 & E_2 & P_1 & E_1 & R
\end{array}$$

est de la forme P P E P E R. Par conséquent il y aura $6!/(3!.2!) = 60$ arrangements possibles des lettres P E P P E R.

Plus généralement, grâce au même raisonnement que celui utilisé dans l'exemple 0.1.4, on établit le théorème suivant :

Théorème 0.1.4. *Il y a*

$$\frac{n!}{n_1!.n_2!...n_r!}$$

permutations différentes de n objets parmi lesquels n_1, sont indistinguables entre eux, n_2 autres entre eux également, ..., n_r entre eux.

Exemple 0.1.5. *Parmi les 10 participants à un tournoi d'échec, on compte 4 russes,3 américains, 2 anglais et un brésilien. Si dans le classement du tournoi on ne peut lire que la liste des nationalités des joueurs mais pas leur identité, à combien de classements individuels différents une telle liste correspond-elle ?*

Il y a

$$\frac{10!}{4!.3!.2!.1!} = 12600$$

classements possibles.

0.1.4 Combinaisons

Nous serons souvent intéressés à déterminer le nombre de groupes de r objets qu'il est possible de former sans répétition à partir d'un total de n objets. Par exemple, combien de groupes de 3 objets peut-on construire en tirant parmi les 5 objets A, B, C, D et El Pour y répondre, on peut raisonner comme suit : puisqu'il y a 5 façons de choisir le premier objet, 4 de choisir ensuite le deuxième et 3 de choisir le dernier, il y a donc 5.4.3 façons de composer des groupes de 3 objets en tenant compte de l'ordre dans lequel ces objets sont choisis. Cependant, un triplet donné, par exemple le triplet constitué des objets A, B et C, apparaîtra 6 fois. En effet, chacune des permutations ABC, ACB, BAC, BCA, CAB et CBA sera distinguée lorsqu'on tient compte de l'ordre. Il en résulte que le nombre total de groupes pouvant être formés est

$$\frac{5.4.3}{3.2.1} = 10.$$

Plus généralement, $n.(n-1)...(n-r+1)$ représente le nombre de manières de choisir un groupe de r objets parmi n lorsqu'on tient compte de l'ordre. Comme chaque groupe de r objets sera distingué $r!$ fois dans ce dénombrement, le nombre de groupes de r objets pris dans un ensemble de n sera

$$\frac{n.(n-1)...(n-r+1)}{r!} = \frac{n!}{(n-r)!r!}.$$

Définition 0.1.2. *E étant un ensemble à n éléments, on appelle combinaison de r éléments de E toute collection non ordonnée de r éléments distincts de E, ie toute partie de E à r éléments.*
On note $\mathcal{C}_n^r$ le nombre de combinaisons de r éléments parmi n. On a

$$\mathcal{C}_n^r = \frac{n!}{(n-r)!r!},$$

Théorème 0.1.5. *$\mathcal{C}_r^n$ est le nombre de combinaisons de r objets pris parmi n, ou encore le nombre de groupes de taille r si, dans le choix, l'ordre n'est pas considéré comme significatif.*

Exemple 0.1.6. *On veut former un comité comprenant 3 des 20 personnes d'un groupe. Combien y a-t-il de ces comités ?*

Il y a $\mathcal{C}_{20}^3 = \dfrac{20.19.18}{3.2.1} = 1140$ *comités possibles.*

Exemple 0.1.7. *A partir d'un groupe de 5 hommes et de 7 femmes, combien de comités différents composés de 2 hommes et de 3 femmes peut-on former ? Qu'en est-il si 2 des femmes s'entendent mal et refusent de siéger simultanément au comité ?*

Comme il y a $\mathcal{C}_5^2$ groupes possibles de 2 hommes et $\mathcal{C}_7^3$ groupes possibles de 3 femmes, il y a selon le principe fondamental $\mathcal{C}_5^2.\mathcal{C}_7^3 = 350$ comités de 2 hommes et 3 femmes.

Considérons maintenant le cas où deux des femmes refusent de siéger ensemble au comité. Comme il y aura $\mathcal{C}_2^0.\mathcal{C}_5^2$ groupes possibles de trois femmes ne contenant aucune des deux ennemies en question et $\mathcal{C}_2^1.\mathcal{C}_5^2$ groupes contenant exactement l'une des deux, il y aura par conséquent $\mathcal{C}_2^1.\mathcal{C}_5^2 + \mathcal{C}_2^1.\mathcal{C}_5^2 = 30$ groupes de 3 femmes ne contenant pas les deux ennemies à la fois. Puisqu'il y a $\mathcal{C}_5^2$ façons de choisir les 2 hommes, il sera possible au total de composer $30.\mathcal{C}_5^2 = 300$ comités différents.

L'identité suivante entre grandeurs combinatoires est très utile :

Théorème 0.1.6.

$$\mathcal{C}_n^r = \mathcal{C}_{n-1}^{r-1} + \mathcal{C}_{n-1}^r, \quad 1 \leq r \leq n. \tag{1}$$

Preuve. L'équation 1 peut être démontrée analytiquement mais aussi grâce à l'argument combinatoire suivant : considérons un groupe de n objets et fixons notre attention sur l'un d'entre eux en particulier, appelons-le objet 1. Il y a alors $\mathcal{C}_{n-1}^{r-1}$ combinaisons de taille r qui contiennent l'objet 1 (puisque chaque combinaison de ce genre est formée en choisissant $r-1$ objets parmi les $n-1$ restants). Il y a également $\mathcal{C}_{n-1}^r$ combinaisons de taille r ne contenant pas l'objet 1. Comme il y a au total $\mathcal{C}_n^r$ combinaisons de taille r, 1 se trouve vérifiée.

Les nombres $\mathcal{C}_r^n$ sont souvent appelés coefficients binomiaux en raison de leur rôle dans le théorème du binôme.

Théorème 0.1.7.

$$(x+y)^n = \sum_{k=0}^{n} \mathcal{C}_n^k x^k y^{n-k}. \tag{2}$$

Preuve. Par un argument d'analyse combinatoire. Considérons le produit suivant :

$$(x_1 + y_1)(x_2 + y_2)...(x_n + y_n)$$

En le développant on obtient une somme de 2^n termes, chaque terme étant un produit de n facteurs. Chacun des 2^n termes de la somme contiendra à son tour soit le facteur x_i, soit y_i et ceci pour tout $i = 1, 2, ..., n$. Par exemple :

$$(x_1 + y_1)(x_2 + y_2) = x_1x_2 + x_1y_2 + y_1x_2 + y_1y_2$$

Combien maintenant de ces 2^n termes de la somme auront-ils k facteurs en x et $(n-k)$ en y.Comme chaque terme constitué par k des x_i et $(n-k)$ des y_i correspond au choix d'un groupe de k des n valeurs $x_1, x_2, ..., x_n$, il y aura $\mathcal{C}_n^k$ de ces termes. Par conséquent, en posant $x_i = x$, $y_i = y$ pour $i = 1, 2, ..., n$ nous voyons que :

$$(x+y)^n = \sum_{k=0}^{n} \mathcal{C}_n^k x^k y^{n-k}.$$

0.1.5 Arrangement

Arrangement sans répétition

Soit Ω un ensemble avec $card(\Omega) = n$. Considèrons un sous-ensemble E de Ω de taille $p(p \leq n)$, la disposition est ordonnée et sans répétition. On dit qu'on a un arrangement sans répétition de p éléments parmi n. Le nombre de $p-$arrangements d'un ensemble à n éléments est :

$$A_n^p = \frac{n!}{(n-p)!}.$$

Définition 0.1.3. *Soient Ω un ensemble fini de cardinal n et p un entier naturel.*
Un p-arrangement de Ω (ou p-arrangement sans répétition de Ω, ou encore arrangement sans répétition de n éléments pris k à k) est une application injective de $\{1, 2, ..., k\}$ dans Ω.

Réaliser un arrangement sans répétition des éléments de Ω, c'est déterminer un $p-$uplet $(x_1, ..., x_p)$ d'éléments de Ω deux à deux distincts. C'est aussi définir une application injective d'un ensemble E à p éléments dans Ω à n éléments.

Arrangement avec répétition

Un arrangement avec répétition, en mathématiques, se produit lorsqu'on range dans un certain ordre p objets, choisis parmi n objets discernables, chaque objet pouvant être répété. On peut représenter ces différents arrangements par des p-uplets. Par exemple, quand on tire successivement avec remise p boules dans une urne contenant n boules numérotées de 1 à n, on peut représenter ces tirages par des p-uplets de boules ou par des applications de $\{1, 2, ..., p\}$ dans l'ensemble des boules.

Le nombre d'arrangements avec répétition de p éléments d'un ensemble fini de cardinal n est égal à n^p.

C'est aussi le nombre d'applications d'un ensemble à p éléments vers un ensemble à n éléments.

Définition 0.1.4. *Étant donnés un ensemble fini* Ω *et un entier naturel* p*, un arrangement avec répétition de* p *éléments de* Ω *est un* p*-uplet d'éléments de* Ω *(un élément de la puissance cartésienne* Ω^p*), autrement dit : une application de* $\{1, 2, ..., p\}$ *dans* Ω*.*

0.2 Probabilité sur un espace fini, événements

Définition 0.2.1. ***Expérience aléatoire***
Une expérience aléatoire est une expérience dont on ne peut pas prévoir le résultat de façon certaine. On ne s'intéresse qu'aux expériences aléatoires dont on peut indiquer l'ensemble des résultats possibles.

Définition 0.2.2. *On appelle univers l'ensemble des résultats possibles d'une expérience aléatoire. Il est noté Ω.*

Définition 0.2.3. ***Evénement***
Un événement (ou évènement) lié à une expérience aléatoire est un sous-ensemble des résultats possibles pour cette expérience (c'est-à-dire un certain sous-ensemble de l'univers lié à l'expérience).

Exemple 0.2.1. *On effectue deux jets successifs d'un dé :*
$\Omega = \{(i,j),\ (i,j) \in \{1,2,3,4,5,6\}^2\}$.

Remarque 0.2.1. Si on répète la même expérience aléatoire d'univers Ω, on pourra choisir comme univers Ω^n dans le cas de n répétitions, et $\Omega^{\mathbb{N}}$ si on la répète indéfiniment.

Notation : On note $\mathcal{A}$ l'ensemble des événements relatifs à l'expérience aléatoire. On a donc par définition $A \subset \mathcal{P}(\Omega)$.

On souhaite pouvoir définir certaines opérations sur les événements, et on utilise la correspondance suivante entre vocabulaires probabiliste et ensembliste :

Terminologie probabiliste	**Terminologie ensembliste**	Ω
Evénement certain	Ensemble entier	Ω
Evénement impossible	Ensemble vide	$\emptyset$
Evénement élémentaire	Singleton	ω
Evénement contraire de A	Complémentaire de A	$\overline{A}$
A ou B	Réunion de A et B	$A \cup B$
A et B	Intersection de A et B	$A \cap B$
A implique B	A inclus dans B	$A \subset B$
A et B incompatibles	A et B disjoints	$A \cap B = \emptyset$
ω réalise A	ω appartient à A	$\omega \in A$

On exige que $\mathcal{A}$ contienne l'ensemble vide et l'univers, et qu'il soit stable pour les opérations de réunion et d'intersection finies ou dénombrables, ainsi que par passage au complémentaire.

Définition 0.2.4. *Soit Ω un ensemble. On dit que $\mathcal{A} \subset \mathcal{P}(\Omega)$ est une tribu (ou une σ-algèbre) si :* **i)** $\Omega \in \mathcal{A}$ *et* $\emptyset \in \mathcal{A}$ *;*

ii) $A \in \mathcal{A} \Longrightarrow \overline{A} \in \mathcal{A}$ *;*

iii) $(\forall i \in I \subset \mathbb{N}, A_i \in \mathcal{A}) \Longrightarrow \bigcup\limits_{i \in I} A_i \in \mathcal{A}$ *;*

iv) $(\forall i \in I \subset \mathbb{N}, A_i \in \mathcal{A}) \Longrightarrow \bigcap\limits^{i \in I} A_i \in \mathcal{A}$*,*
c'est-à-dire que $\mathcal{A}$ soit une tribu de parties de Ω.

Définition 0.2.5. *On appelle événement élémentaire tout singleton (partie constituée d'un seul élément) de la tribu $\mathcal{A}$.*

Exemple 0.2.2. *On lance deux fois un dé à six faces. On pose :*
$\Omega = \{(i, j),\ (i, j) \in \{1, 2, 3, 4, 5, 6\}^2\}$ *et* $\mathcal{A} = \mathcal{P}(\Omega)$.
On considère A : "la somme obtenue est supérieure ou égale à 11".
On a : $A = (5, 6); (6, 5); (6, 6)$ et A est un événement ($A \in \mathcal{A}$).
On considère B : "la somme obtenue est divisible par 3 et par 4".
On a : $B = (6, 6)$ et B est un événement élémentaire.

Définition 0.2.6. *Le couple $(\Omega; \mathcal{A})$, où Ω est l'univers et $\mathcal{A}$ une tribu de parties de Ω, est appelé espace probabilisable.*

0.2.1 Probabilité

Il s'agit d'affecter à chaque événement A un poids $\mathbf{P}(A)$ indiquant sa <<chance>> d'être réalisé si l'on effectue l'expérience aléatoire.

Définition 0.2.7. *Étant donné un espace probabilisable $(\Omega; \mathcal{A})$, on appelle probabilité sur $(\Omega; \mathcal{A})$ toute application $\mathbf{P} : \mathcal{A} \to [0, 1]$ satisfaisant aux trois axiomes suivants :*

1. $\forall A \in \mathcal{A}$, $\mathbf{P}(A) \geq 0$ *(positivité).*
2. $\mathbf{P}(\Omega)= 1$ *(totalité).*

3. *Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments deux à deux disjoints,*

$$\mathbf{P}(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n) \quad \textit{(Additivité).}$$

Définition 0.2.8. *Soit $A \in \mathcal{A}$.*
Si $\mathbf{P}(A) = 0$, on dit que A est un événement impossible.
Si $\mathbf{P}(A) = 1$, on dit que A est un événement certain.

Définition 0.2.9. *Deux événements A et B sont indépendants sous $\mathbf{P}$ si*

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B).$$

Propriété 0.2.1. *Toute probabilité* $\mathbf{P}$ *possède les propriétés suivantes :*
1. $\mathbf{P}(\emptyset) = 0$.
2. $\forall (A,B) \in \mathcal{A}^2 \ A \subset B \Rightarrow \mathbf{P}(A) \leq \mathbf{P}(B)$.
3. $\forall A \in \mathcal{A}$, $\mathbf{P}(A) \in [0,1]$ *et* $\mathbf{P}(A) = 1 - \mathbf{P}(\overline{A})$.
4. $\forall A, B \in \mathcal{A}$, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(A \cap B)$.

Preuve.

1) On pose $(A_n)_{n\in\mathbb{N}}$ une partition de Ω, d'éléments deux à deux disjoints, tel que $A_0 = \Omega$ et $A_i = \emptyset$, pour tout $i \geq 1$.
D'après l'additivité

$$\begin{array}{rcl} \sum_{i=0}^{+\infty} \mathbf{P}(A_i) & = & \mathbf{P}(\Omega) \qquad\qquad\qquad = \quad 1 \\ & = & \mathbf{P}(A_0) + \sum_{i=1}^{+\infty} \mathbf{P}(\emptyset) \ = \ 1 + \sum_{i=1}^{+\infty} \mathbf{P}(\emptyset). \end{array}$$

Par conséquent, $\sum_{i=1}^{+\infty} \mathbf{P}(\emptyset) = 0$, et puisque $\mathbf{P}(\emptyset) \geq 0$, donc $\mathbf{P}(\emptyset) = 0$.

2) Si $A \subset B$, alors $B = A \cup (B \cap \overline{A})$ et cette réunion est disjointe.
D'après l'additivité, on a $\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(B \cap \overline{A})$ et comme $\mathbf{P}(B \cap \overline{A}) \geq 0$, on en déduit $\mathbf{P}(B) \geq \mathbf{P}(A)$.

3) On prend la partition de $\Omega = A \cup \overline{A}$.
D'après l'additivité on a :

$$\begin{array}{rcl} \mathbf{P}(\Omega) & = & 1 \\ & = & \mathbf{P}(A) + \mathbf{P}(\overline{A}). \end{array}$$

En effet $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$.

4) On a les décompositions suivantes en unions disjointes

$$\begin{array}{rcl} A \cup B & = & (A \cap \overline{B}) \cup (A \cap B) \cup (\overline{A} \cap B), \\ A & = & (A \cap \overline{B}) \cup (A \cap B), \\ B & = & (A \cap B) \cup (\overline{A} \cap B). \end{array}$$

En utilisant l'additivité on en déduit :

$$\begin{array}{rcl} \mathbf{P}(A \cup B) & = & \mathbf{P}(A \cap \overline{B}) + \mathbf{P}P(A \cap B) + \mathbf{P}(\overline{A} \cap B) \\ & = & [\mathbf{P}(A) - \mathbf{P}(A \cap B)] + \mathbf{P}(A \cap B) + [\mathbf{P}(B) - \mathbf{P}(A \cap B)] \\ & = & \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) \end{array}$$

0.3 Exemples de calcul probabiliste

Exemple 0.3.1. *On effectue une partie de pile ou face en trois coups. Quelle est la probabilité d'obtenir pile aux premier et troisième lancers.*
On peut modéliser cette expérience en prenant $\Omega = \{f, p\}^3$. *La pièce étant supposée symétrique, nous n'avons à priori pas de raison de supposer que l'un des 8 triplets de résultats*

possibles soit favorisé ou défavorisé par rapport aux autres. Nous choisirons donc $\mathbf{P}$ *de sorte que tous les événements élémentaires aient même probabilité (hypothèse d'équiprobabilité), soit :*

$$\forall \omega \in \Omega, \ \ \mathbf{P}(\omega) = \frac{1}{Card(\Omega)} = \frac{1}{2^3}.$$

L'événement B *dont on veut calculer la probabilité s'écrit :*

$$B = \{(p, f, p); (p, p, p)\}.$$

D'où

$$\mathbf{P}(B) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}.$$

Exemple 0.3.2. *On fait remplir un questionnaire à 20 questions binaires. Quelle est la probabilité qu'un candidat répondant au hasard obtienne au moins 16 bonnes réponses ?*
On choisit ici :

$$\Omega = \{oui, non\}^{20}, \ \ \mathcal{A} = \mathcal{P}(\Omega).$$

Si le candidat répond complètement au hasard, on peut considérer que chacune des 2^{20} *grilles de réponses possibles a la même probabilité d'apparaître (hypothèse d'équiprobabilité sur* Ω*). Pour tout* $B \subset \Omega$*, on a alors :*

$$\mathbf{P}(B) = \frac{CardB}{Card\Omega}.$$

En particulier pour $B =$ *obtention d'au moins 16 bonnes réponses,*

$$\mathbf{P} = \frac{\mathcal{C}_{20}^{16} + \mathcal{C}_{20}^{17} + \mathcal{C}_{20}^{18} + \mathcal{C}_{20}^{19} + \mathcal{C}_{20}^{20}}{2^{20}} \simeq 0,006.$$

1

Rappels sur les Probabilités Conditionnelles

Comment doit-on calculer la probabilité que l'on attribue à un événement lorsque l'on dispose d'une information conditionnelle ? Le concept de probabilité conditionnelle permet de répondre à cette question.

Par exemple, supposons que nous jetions deux dés et que chacun des 36 événements élémentaires ait la même probabilité de survenir, soit $\frac{1}{36}$. Supposons encore que nous puissions observer le premier dé, qui donne un 3. Sur la base de cette information (condition) ; quelle est la probabilité que la somme des deux dés est 8 ?
Notons l'événement $A =$" La somme des deux dés est 8 " et $B =$ " Le premier dé est un 3", notre objectif, c'est de chercher la probabilité de A sachant B c.-à-d. la probabilité que la somme des deux dés est 8 sachant que le premier dé est un 3.
Pour calculer cette probabilité on procède comme suit : le dé initial étant un 3, il ne peut plus avoir que 6 événements dans notre expérience, à savoir : $(3, 1)$, $(3, 2)$, $(3, 3)$, $(3, 4)$, $(3, 5)$ et $(3, 6)$. Puisque chacun de ces évènements a originellement la même probabilité d'apparaître, ils auront encore des probabilités égales. Autrement dit, étant donné que le premier dé est un 3, la probabilité conditionnelle de chacun des évènements $(3, 1)$, $(3, 2)$, $(3, 3)$, $(3, 4)$, $(3, 5)$ et $(3, 6)$ devient $\frac{1}{6}$, tandis que la probabilité (conditionnelle) des 30 autres événements de l'ensemble fondamental devient 0. Finalement la probabilité cherchée est-elle $\frac{1}{6}$.

1.1 Probabilités Conditionnelles : Définition et Propriètés

Définition 1.1.1. *Soit B un événement tel que $\mathbf{P}(B) \neq 0$. Pour tout événement observable A, on définit la probabilité de A sachant B notée $\mathbf{P}(A|B)$ par :*

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}. \tag{1.1}$$

$\mathbf{P}(A|B)$ est appelée probabilité conditionnelle de l'événement A sous la condition B.

Exemple 1.1.1. *Une pièce de monnaie est lancée deux fois, sachant qu'une pièce de monnaie dispose deux cotés Pile (noté P) et Face (noté F), alors l'ensemble fondamental est $\Omega = \{(F, F), (F, P), (P, F), (P, P)\}$, quelle est la probabilité conditionnelle que les deuxième jet amène face sachant que le premier est face ?*

En désignant par $B =$ " Le premier jet est face ", et $A =$ " Le deuxième jet est face ", la probabilité voulue est donnée par :

$$\begin{aligned}
\mathbf{P}(A|B) &= \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \\
&= \frac{\mathbf{P}(\{(F, F)\})}{\mathbf{P}(\{(F, F), (F, P)\})} \\
&= \frac{\frac{1}{4}}{\frac{1}{2}} \\
&= \frac{1}{2}.
\end{aligned}$$

Remarque 1.1.1. Remarquons que pour l'instant, il ne s'agit que d'un jeu d'écriture. On a simplement défini un réel $\mathbf{P}(A|B)$ pour que :

$$\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\mathbf{P}(B).$$

Exemple 1.1.2. *Une urne contient r boules rouges et v boules vertes. On tire deux boules l'une après l'autre, sans remise. Quelle est la probabilité d'obtenir deux rouges ? Notons A et B les événements :*
$B =$ "rouge au 1^{er} tirage", $A =$ "rouge au 2^{e} tirage".
Un espace probabilisé (Ω, F, P) modélisant correctement cette expérience devrait vérifier :

$$\begin{aligned}
\mathbf{P}(B) &= \frac{r}{r+v}, \\
\mathbf{P}(A|B) &= \frac{r-1}{r-1+v}.
\end{aligned}$$

On déduit :

$$\mathbf{P}(deux\ rouges) = \mathbf{P}(A \cap B) = \mathbf{P}(A|B)\mathbf{P}(B) = \frac{r}{r+v} \times \frac{r-1}{r-1+v}.$$

Propriété 1.1.1.

1. $\mathbf{P}(\emptyset|B) = 0$, $\mathbf{P}(\Omega|B) = 1$ *et si* $B \subset A$ *alors* $\mathbf{P}(A|B) = 1$.

2. *Si les* A_i *sont deux à deux disjoints :*

$$\mathbf{P}(\overset{n}{\underset{i=1}{\cup}} A_i \mid B) = \sum_{i=1}^{n} \mathbf{P}(A_i \mid B).$$

3. *Pout tous* $A, B \in \Omega$, $\quad \mathbf{P}(\overline{A} \mid B) = 1 - \mathbf{P}(A \mid B)$.

4. *Pour tous* $A_1, A_2, B \in \Omega$, *si* $A_1 \subset A_2$, $\mathbf{P}(A_1 \mid B) \leq \mathbf{P}(A_2 \mid B)$.

5. *Pour tous* $A_1, A_2, B \in \Omega$.

$$\mathbf{P}(A_1 \cup A_2 \mid B) = \mathbf{P}(A_1 \mid B) + \mathbf{P}(A_2 \mid B) - \mathbf{P}(A_1 \cap A_2 \mid B).$$

Proposition 1.1.1. ***(Règle des conditionnements successifs)***

Si $A_1, A_2, ..., A_n$ *sont* n *événements tels que* $\mathbf{P}(A_1 \cap A_2 ... \cap A_{n-1}) \neq 0$, *on a :*

$$\mathbf{P}(A_1 \cap A_2 ... \cap A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2 \mid A_1)\mathbf{P}(A_3 \mid A_1 \cap A_2) \times ... \times \mathbf{P}(A_n \mid A_1 \cap A_2 ... \cap A_{n-1}).$$

Proposition 1.1.2. ***(Conditionnement par les cas possibles)***

i) *Si* $B \in \Omega$ *tel que* $\mathbf{P}(B) \neq 0$ *et* $\mathbf{P}(\overline{B}) \neq 0$, *on a*

$$\forall A \in \Omega, \qquad \mathbf{P}(A) = \mathbf{P}(A \mid B)\mathbf{P}(B) + \mathbf{P}(A \mid \overline{B})\mathbf{P}(\overline{B}).$$

ii) *Si* $B_1, ..., B_n$ *est une partition finie de* Ω *en événements de probabilités non nulles,*

$$\forall A \in \mathcal{A}, \ \ \mathbf{P}(A) = \sum_{i=1}^{n} \mathbf{P}(A|B_i)\mathbf{P}(B_i).$$

iii) *Si* $(B_i)_{i \in \mathbb{N}}$ *est une partition de* Ω *telle que* $\forall i \in \mathbb{N}$, $\mathbf{P}(B_i) \neq 0$ *:*

$$\forall A \in \mathcal{A}, \ \ \mathbf{P}(A) = \sum_{i=0}^{+\infty} \mathbf{P}(A|B_i)\mathbf{P}(B_i).$$

Preuve.
Il suffit de vérifier **(iii)**, les deux premières propriétés se démontrant de façon analogue. Comme $(B_i)_{i\in\mathbb{N}}$ est une partition de Ω,

$$A = A \cap \Omega = A \cap \left(\underset{i\in N}{\cup} B_i\right) = \underset{i\in N}{\cup} (A \cap B_i)\,,$$

et cette réunion est disjointe car les B_i étant deux à deux disjoints, il en est de même pour les $(A \cap B_i)$. Par conséquent de l'additivité :

$$\mathbf{P}(A) = \sum_{i=0}^{+\infty} \mathbf{P}(A \cap B_i) = \sum_{i=0}^{+\infty} \mathbf{P}(A|B_i)\mathbf{P}(B_i).$$

Lorsqu'on a une partition de Ω en n hypothèses ou cas possibles B_i et que l'on sait calculer les $\mathbf{P}(B_i)$ et les $\mathbf{P}(A|B_i)$, on peut se poser le problème inverse : calculer $\mathbf{P}(B_j|A)$ à l'aide des quantités précédentes. La solution est donnée par la formule suivante quelquefois appelée (abusivement) formule des probabilités des causes.

1.2 Formule de Bayes

1.2.1 Formule des Probabilités Totales

Soient A et B deux événements quelconques. Nous pouvons écrire A sous la forme $A = (A \cap B) \cup (A \cap \overline{B})$. Comme évidemment $A \cap B$ et $A \cap \overline{B}$ deux ensembles disjoints, on peut écrire :

$$\begin{aligned} \mathbf{P}(A) &= \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \overline{B}) \\ &= \mathbf{P}(A \mid B)\mathbf{P}(B) + \mathbf{P}(A \mid \overline{B})\mathbf{P}(\overline{B}) \\ &= \mathbf{P}(A \mid B)\mathbf{P}(B) + \mathbf{P}(A \mid \overline{B})(1 - \mathbf{P}(B)). \end{aligned} \tag{1.2}$$

L'équation 1.2 appelée formule des probabilités totales, peut être interprétée de la façon suivante : la probabilité de l'événement A est une moyenne pondérée de la probabilité conditionnelle de A lorsque B est apparu et de la probabilité conditionnelle du même A lorsque B n'est pas apparu, les poids étant les probabilités des événements conditionnants.

Il existe de nombreuses situations où il est difficile de calculer directement la probabilité d'un événement mais où il est par contre possible de la calculer connaissant ses probabilités conditionnelles si certains événements sont réalisés. Quelques exemples illustrent cette démarche.

1.2.2 Introduction à la formule de Bayes

Formellement, la formule de Bayes est simplement dérivée de (1.1) et (1.2). L'expression générale de cette formule sera donnée plus tard. Les exemples suivants vont néanmoins servir à donner une idée intuitive du champ d'application de cette formule importante.

Exemple 1.2.1. *Un étudiant répond à une question à choix multiple. Deux cas qui se présente , $A =$ "Il connait la réponse" , $\overline{A} =$ " Il devine la réponse " . Soit q la probabilité que l'étudiant connaisse la réponse et donc $(1-q)$ qu'il la devine. On suppose que l'étudiant qui devine répondra correctement avec une probabilité $1/m$, où m est le nombre de réponses possibles. Quelle est la probabilité qu'un étudiant connaisse la réponse sachant qu'il a correctement répondu à la question ?*

Solution *: Soit B l'événement " L'étudiant répond correctement à la question", alors on cherche $\mathbf{P}(A \mid B)$:*

$$\begin{aligned} \mathbf{P}(A \mid B) &= \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \\ &= \frac{\mathbf{P}(B \mid A)\mathbf{P}(A)}{\mathbf{P}(B \mid A)\mathbf{P}(A) + \mathbf{P}(B \mid \overline{A})\mathbf{P}(\overline{A})}. \end{aligned}$$

De plus, $\mathbf{P}(A) = q$, $\mathbf{P}(\overline{A}) = 1 - q$, $\mathbf{P}(B \mid \overline{A}) = 1/m$ et c'est clair que $\mathbf{P}(B \mid A) = 1$, car $A \subset B$.

D'où

$$\mathbf{P}(A \mid B) = \frac{q}{q + \frac{1-q}{m}} = \frac{qm}{qm + 1 - q}.$$

Exemple 1.2.2. *Un laboratoire d'analyse de sang assure une fiabilité de 95% la détection d'une certaine maladie lorsqu'elle est effectivement présente. Cependant, le test indique aussi un résultat faussement positif pour 1% des personnes réellement saines. Si 0.5% de la population porte effectivement la maladie quelle est la probabilité qu'une personne soit vraiment malade sachant le test est positif ?*

Solution *: Soit D l'événement " La personne soumise au test est porteuse de la maladie" et E l'événement " Le résultat du test est positif". Notre objectif c'est de calculer $\mathbf{P}(D \mid E)$. On a*

$$\begin{aligned} \mathbf{P}(D \mid E) &= \frac{\mathbf{P}(D \cap E)}{\mathbf{P}(E)} \\ &= \frac{\mathbf{P}(E \mid D)\mathbf{P}(D)}{\mathbf{P}(E \mid D)\mathbf{P}(D) + \mathbf{P}(E \mid \overline{D})\mathbf{P}(\overline{D})} \\ &= \frac{0.95 \times 0.005}{0.95 \times 0.005 + 0.01 \times 0.995} \\ &= 0.323. \end{aligned}$$

1.2.3 Formules des probabilités totales généralisée

L'équation 1.2 peut être généralisée de la manière suivante : supposons que $B_1, B_2, \ldots, B_n$ est un système exhaustif d'événements et tels que

$$\bigcup_{i=1}^{n} B_i = \Omega.$$

Dans ce cas-là, on peut écrire :

$$A = \bigcup_{i=1}^{n} (A \cap B_i),$$

Et en utilisant le fait que les événements $A \cap B_i, \quad i = 1, \ldots, n$ sont disjoints, on obtient la propriètè suivante :

$$\begin{aligned} \mathbf{P}(A) &= \sum_{i=1}^{n} \mathbf{P}(A \cap B_i) \\ &= \sum_{i=1}^{n} \mathbf{P}(A \mid B_i)\mathbf{P}(B_i) \end{aligned} \tag{1.3}$$

1.2.4 Formules de Bayes généralisée

Supposons que A s'est réalisé et que nous cherchions à déterminer la probabilité que l'un des B_j se soit aussi réalisé. On déduit de l'équation (1.3) dans le théorème suivant :

Théorème 1.2.1. ***(Théorème de Bayes)***
Soit $B_1, B_2, \ldots, B_n$, deux à deux disjoints et A un événement réalisable, alors :

$$\begin{aligned} \mathbf{P}(B_j \mid A) &= \frac{\mathbf{P}(A \cap B_j)}{\mathbf{P}(A)} \\ &= \frac{\mathbf{P}(A \mid B_j)\mathbf{P}(B_j)}{\sum_{i=1}^{n} \mathbf{P}(A \mid B_i)\mathbf{P}(B_i)}. \end{aligned} \tag{1.4}$$

L'équation 1.4 *est appelée formule de Bayes.*

Preuve. Il suffit d'appliquer la proposition (1.1.2).

Exemple 1.2.3. *Un avion est porté disparu. On pense que l'accident a pu arriver aussi bien dans n'importe laquelle de trois régions données i, $(i = 1, 2, 3)$. Notons par $(1 - \alpha_i)$ la probabilité qu'on découvre l'avion dans la région i s'il y est effectivement. Les valeurs α_i représentent donc la probabilité de manquer l'avion lors des recherches, on peut l'attribuer à diverses causes d'ordre géographique ou à la végétation à la région. Quelle est la probabilité que l'avion se trouve dans la $i-$ème région si les recherches dans la région 1 n'ont rien donné, $i = 1, \quad 2, \quad 3$?*

Solution *Soit* R_i, $i = 1, 2, 3$ *les événements " L'avion est tombé dans la région i". Soit aussi A l'événement " les recherche dans la région 1 n'ont rien donné ".*
On utilise la formule de Bayes pour $i = 1$ *:*

$$\begin{aligned}
\mathbf{P}(R_1 \mid A) &= \frac{\mathbf{P}(A \cap R_1)}{\mathbf{P}(A)} \\
&= \frac{\mathbf{P}(A \mid R_1)\mathbf{P}(R_1)}{\sum_{i=1}^{3} \mathbf{P}(A \mid R_i)\mathbf{P}(R_i)} \\
&= \frac{(\alpha_1)\frac{1}{3}}{(\alpha_1)\frac{1}{3} + (1)\frac{1}{3} + (1)\frac{1}{3}} \\
&= \frac{\alpha_1}{\alpha_1 + 2}.
\end{aligned}$$

Pour $j = 1, 2$

$$\begin{aligned}
\mathbf{P}(R_j \mid A) &= \frac{\mathbf{P}(A \cap R_j)}{\mathbf{P}(A)} \\
&= \frac{(1)\frac{1}{3}}{(\alpha_1)\frac{1}{3} + (1)\frac{1}{3} + (1)\frac{1}{3}} \\
& \frac{1}{\alpha_1 + 2}.
\end{aligned}$$

1.3 Travaux dirigès

Exercice 1.3.1. *Soit un espace probabilisé* $(\Omega, \mathcal{A}, \mathbf{P})$ *et considérons* n *événements* $A_1, A_2, ..., A_n$ *avec* $n \geq 3$*, tels que* $\mathbf{P}(A_1 \cap A_2 \cap ... \cap A_n) > 0$*.*
Démontrer la formule :

$$\mathbf{P}(A_1 \cap A_2 ... \cap A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2 \mid A_1)\mathbf{P}(A_3 \mid A_1 \cap A_2) \times ... \times \mathbf{P}(A_n \mid A_1 \cap A_2 ... \cap A_{n-1}).$$

Exercice 1.3.2. *Dans une entreprise deux ateliers fabriquent les mêmes pièces. L'atelier 1 fabrique en une journée deux fois plus de pièces que l'atelier 2. Le pourcentage de pièces défectueuses est* 3% *pour l'atelier 1 et* 4% *pour l'atelier 2. On prélève une pièce au hasard dans l'ensemble de la production d'une journée. Déterminer :*
1) *la probabilité que cette pièce provienne de l'atelier 1 ;*
2) *la probabilité que cette pièce provienne de l'atelier 1 et est défectueuse ;*
3) *la probabilité que cette pièce provienne de l'atelier 1 sachant qu'elle est défectueuse.*

Exercice 1.3.3. *L'inspecteur chargé d'une enquête criminelle est à un certain stade convaincu à* 60% *de la culpabilité d'un suspect. Une nouvelle pièce à conviction permet soudain d'affirmer que le criminel est gaucher. Or* 7% *des individus dans la population sont gauchers. Comment l'inspecteur doit-il réapprécier la culpabilité du suspect, s'il se trouve que le suspect est gaucher ?*

Exercice 1.3.4. *Une compagnie d'assurance répartit ses clients en trois classes* R_1*,* R_2 *et* R_3 *: les bons risques, les risques moyens, et les mauvais risques. Les effectifs de ces trois classes représentent* 20% *de la population totale pour la classe* R_1*,* 50% *pour la classe* R_2 *et* 30% *pour la classe* R_3*. Les statistiques indiquent que les probabilités d'avoir un accident au cours de l'année pour une personne de l'une de ces trois classes sont respectivement de* 0.05*,* 0.15 *et* 0.30*.*
1) *Quelle est la probabilité qu'une personne choisie au hasard dans la population ait un accident dans l'année ?*
2) *Si M.Martin n'a pas eu d'accident cette année, quelle est la probabilité qu'il soit un bon risque ?*