Chapitre 1 — Analyse combinatoire

1. Introduction

L'analyse combinatoire est une branche des mathématiques qui étudie comment compter les objets. Elle fournit des méthodes de dénombrements particulièrement utiles en théorie des probabilités. Les probabilités dites combinatoires utilisent constamment les formules de l'analyse combinatoire développées dans ce chapitre.

2. Arrangements

2.1. Définition

Etant donné un ensemble E de n objets, on appelle arrangements de p objets toutes suites ordonnées de p objets pris parmi les n objets. Le nombre d'arrangements de p objets pris parmi n est noté : A^p_n.

Remarque : On a nécessairement $1 \geq p \geq n$ et $n,p \in N$

Deux arrangements de p objets sont donc distincts s'ils diffèrent par la nature des objets qui les composent ou par leur ordre dans la suite.

Exemples :

1. Une séquence d'ADN est constituée d'un enchaînement de 4 nucléotides [A (Adénine), C (Cytosine), G (Guanine) et T (Thymine)]. Il existe différents arrangements possibles de deux nucléotides ou dinucléotides avec p=2 et n=4.

2. Le nombre de mots de 5 lettres (avec ou sans signification) formés avec les 26 lettres de l'alphabet correspond au nombre d'arrangements possibles avec p=5 et n=26.

3. Le tiercé dans l'ordre lors d'une course de 20 chevaux constitue un des arrangements possibles avec p=3 et n=20.

Dans les exemples précédents, l'ordre des éléments dans la suite est essentiel. Ainsi pour le deuxième exemple, le mot NICHE est différent du mot CHIEN.

Mais dans les deux premiers exemples, une base ou une lettre de l'alphabet peut se retrouver plusieurs fois alors que dans le troisième exemple, les trois chevaux à l'arrivée sont forcément différents. Il faut donc distinguer le nombre d'arrangements avec répétition et le nombre d'arrangements sans répétition (arrangements au sens strict).

2.2. Arrangements avec répétitions

Lorsqu'un objet peut être observé plusieurs fois dans un arrangement, le nombre d'arrangement avec répétition de p objets pris parmi n, est alors :

$A^p_n = n^p$ avec $1 \geq p \geq n$

Voici pourquoi :

Pour le premier objet tiré, il existe n manières de ranger l'objet parmi n.

Pour le second objet tiré, il existe également n possibilités d'arrangements car le premier objet fait de nouveau parti des n objets. On parle de tirage avec remise.

Ainsi pour les p objets tirés, il y aura $n \times n \times n \times n \times \ldots \times n$ (p fois) arrangements possibles, soit

$A^p_n = n \times n \times n \times \ldots \times n = n^p$

Exemples :

1. Concernant l'exemple de la séquence d'ADN, le nombre de dinucléotides attendus si l'on fait l'hypothèse qu'une base peut être observée plusieurs fois dans la séquence (ce qui correspond effectivement à la réalité) est donc : $A^2_4 = 4^2 = 16$ dinucléotides possibles

Les 16 dinucléotides identifiables dans une séquence d'ADN sont :

AA	AC	AG	AT	CA	**CC**	CG	CT
GA	GC	**GG**	GT	TA	TC	TG	**TT**

2.3.1 Arrangements sans répétition

Lorsque chaque objet ne peut être observé qu'une seule fois dans un arrangement, le nombre d'arrangements sans répétition de p objets pris parmi n est alors :

$A^p_n = \frac{n!}{(n-p)!}$ avec $1 \geq p \geq n$

Voici pourquoi : Pour le premier objet tiré, il y a n manières de ranger l'objet parmi n.

Pour le second objet tiré, il n'existe plus que n-1 manières de ranger l'objet car le premier objet ne peut plus être pris en compte. On parle de tirage sans remise.

Ainsi pour les p objets tirés parmi n, si $1 \geq p \geq n$, il y aura :

$A^p_n = n(n-1)(n-2)...(n-p+1)$ (p produits)

De plus $A^p_n = n\,(n-1)(n-2)...(n-p+1)$

d'où $A^p_n = \frac{n!}{(n-p)!}$

Rappel : Si $n \in N^*$, on appelle factorielle n, notée n! , le produit des n premiers entiers : $1 \times 2 \times 3 \times ..... \times p \times (p+1) \times ... \times (n-1) \times n = n!$

0!=1 par convention car 0! n'est en principe pas définie.

Dès que n dépasse la dizaine, n! se compte en millions. Il est bon de connaître la formule d'approximation suivante (« formule de Stirling »):

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

Exemple: $A^2_4 = 4^2 = 16$

Concernant l'exemple de la séquence d'ADN, le nombre de dinucléotides attendu dans une séquence si l'on fait l'hypothèse qu'une base n'est observée qu'une seule fois est donc : $A^2_4 = \frac{4!}{(4-2)!}$ = 12 dinucléotides possibles

Sous cette contrainte, les 12 dinucléotides possibles sont :

~~AA~~	AC	AG	AT	CA	~~CC~~	CG	CT
GA	GC	~~GG~~	GT	TA	TC	TG	~~TT~~

Ceci correspond aux 16 arrangements possibles avec répétition ($A^p_n = n^p$) auxquels sont soustraits les 4 dinucléotides (n) résultant de l'association d'une même base.

3. Permutations

3.1. Permutations sans répétition

Etant donné un ensemble E de n objets, on appelle permutations de n objets distincts toutes suites ordonnées de n objets ou tout arrangement n à n de ces objets. Le nombre de permutations de n objets est noté : $P_n = n!$

La permutation de n objets constitue un cas particulier d'arrangement sans répétition de p objets pris parmi n lorsque p=n

Ainsi le nombre de permutations de n objets est : $A^n_n = \frac{n!}{(n-n)!} = n!$

Exemple:

Le nombre de manières de placer 8 convives autour d'une table est : $P_8 = 8!$ 40 320 possibilités

3.2. Permutations avec répétitions

Dans le cas où il existerait plusieurs répétitions k d'un même objet parmi les n objets, le nombre de permutations possibles des n objets doit être rapporté aux nombres de permutations des k objets identiques.

Le nombre de permutations de n objets est alors : $P_n = \frac{n!}{k!}$

En effet, les permutations de k objets identiques sont toutes identiques et ne comptent que pour une seule permutation.

Exemple:

Considérons le mot « CELLULE ». Le nombre de mots possibles (avec ou sans signification) que l'on peut écrire en permutant ces 7 lettres est : $P_7 = \frac{7!}{2!3!}$ = 420 mots possibles

En considérant deux groupes de lettres identiques : L (3 fois) et E (2 fois).

4. Combinaisons

4.1. Définition

Si l'on reprend l'exemple de la séquence d'ADN, à la différence des arrangements où les dinucléotides AC et CA formaient deux arrangements distincts, ces derniers ne formeront qu'une seule combinaison. Pour les combinaisons, on ne parle plus de suite ni de série puisque la notion d'ordre des objets n'est plus prise en compte. On parle alors de tirages avec ou sans remise.

4.2. Combinaisons sans remise

Étant donné un ensemble E de n objets, on appelle combinaisons de p objets tout ensemble de p objets pris parmi les n objets sans remise.

Le nombre de combinaisons de p objets pris parmi n est noté : C_n^p

Remarque : On a nécessairement $1 \geq p \geq n$ et $n,p \in N^*$ Si $n<p$, alors $C_n^p = 0$

Exemples : (1) La formation d’une délégation de 5 personnes parmi un groupe de 50 constitue une combinaison avec p=5 et n=50.

Pour l’exemple, les objets tirés sont clairement distincts.

Le nombre de combinaisons de p objets pris parmi n et sans remise est :

$C_n^p = \frac{n!}{p!(n-p)!}$ notée (^p_n) avec $1 \geq p \geq n$

Voici pourquoi :

Pour calculer ce nombre, on utilise le principe de la division.

• Il y a A_n^p manières de tirer p objets parmi n en les ordonnant soit $A_n^p = \frac{n!}{(n-p)!}$

• Une fois les p objets tirés, il y a p! manières de les ordonner.

• Il y a donc A_n^p / p! manières de tirer p objets parmi n sans les ordonner.

$$C_n^p = \frac{Anp}{p!} = \frac{n!}{p!(n-p)!}$$

Remarque : A la notation ancienne $C^p{}_n$, on préfère parfois la notation moderne (np). Les nombres n et p constituent les coefficients binomiaux.

Exemples:

Dans le cadre de l’exemple de la séquence d’ADN, le nombre de dinucléotides attendus sans tenir compte de l’ordre des bases dans la séquence (hypothèse justifiée dans le cas de l’ ADN non codant) est donc : $C^2{}_4$= 4! / 2!(4−2)! = 4×3 / 2×1 = 6 dinucléotides

Les 6 dinucléotides possibles sous cette hypothèse sont :

AC CA	AG GA	AT TA	CG GC	CT TC	GT TG

Ceci correspond aux 12 arrangements possibles sans répétitions ($A^p_n = \frac{n!}{(n-p)!}$) divisé par le nombre de permutations possibles avec 2 nucléotides ($P_p=p!$).

4.3. Combinaisons avec remise

Le nombre de combinaisons de p objet parmi n avec remise est :

$C_{n+p-1}^{\ p}$ = (n+p−1)!p!(n−1)!Cn+p−1p=(n+p−1)!p!(n−1)!

Voici pourquoi :

Soit la constitution de mots de 3 lettres à partir d'un alphabet à 5 lettres avec remise, on distingue 3 cas possibles :

• C^3_5nombre de mots de 3 lettres différentes et sans ordre

• $C^2_5\times 2$ nombre de mots de 2 lettres différentes et une lettre redondante

• C^1_5nombre de mots de 3 lettres identiques

d'où au total : $C^3_5 + 2\,C^2_5 + C^1_5 = C^3_7$ =35 en utilisant la formule des combinaisons composées.

en effet $C^3_5 + 2\,C^2_5 = C^3_6$et $C^2_5 + C^1_5 = C^2_6$d'où $C^2_5 + C^1_5 = C^2_6$soit C^3_7 = 35 mots possibles de 3 lettres à partir d'un alphabet à 5 lettres.

ainsi $C_7^3 = C^3_{5+3-1} = C^p_{n+p-1}$ avec n=5 et p=3

4.4. Propriétés des combinaisons

4.4.1. La symétrie

Le nombre de combinaisons de pp objets pris parmi nn étant Cnp=n!p!(n−p)!Cpn=n!p!(n−p)!, alors

1. C0n=Cnn=1 car C0n = Cnn = n! / n!

2. si n≥1 C1n=Cn−1n=n car C1n=Cn−1n=n! / (n−1)!

3. si $n \geq 2$ $C2n=Cn-2n=n(n-1) / 2$

avec $C2n=Cn-2n= n! / 2!n-2! = n \times (n-1)(n-2)! / 2!n-2!$

Par récurrence, on déduit des relations précédentes, la propriété de symétrie à savoir :

si $0 \leq p \leq nCn-pp=n!p!(n-p)!$ ainsi $Cpn=Cn-pn$

Il revient au même de donner la combinaison des p objets choisis ou bien celle des (n−p) qui ne le sont pas.

Chapitre 2 — Introduction au calcul des probabilités

1. Introduction
2. Espace fondamental et évènements

2.1 Définitions

2.2. Evènements remarquables

2.3. Opérations sur les évènements

2.4. Système complet d'évènements

2.5. Espace probabilisable

Chapitre 3 — Eléments de calcul des probabilités

Probabilités

1 Définitions

2 Propriétés des probabilités

3 Indépendance statistique

Chapitre 4 — Probabilités conditionnelles

1. Définition
2. Probabilités composées
3. Probabilités totales
4. Diagramme en arbre
5. Le théorème de Bayes

1. Espace fondamental et évènements

La théorie des ensembles qui est succinctement présentée dans ce chapitre constitue un outil puissant dans plusieurs branches des mathématiques, notamment en probabilités.

En face de situations dont **l'issue est incertaine**, on a bien souvent envie d'attribuer à chacune des éventualités possibles une vraisemblance plus ou moins grande. Afin de donner une rigueur mathématique à ce concept, il est nécessaire tout d'abord de donner quelques définitions.

- Une expérience ou une épreuve est qualifiée d'**aléatoire** si on ne peut pas prévoir son résultat et si, répétée dans des conditions identiques, elle peut donner des résultats différents.
- Le résultat d'une expérience noté ω constitue une **éventualité** ou un **événement élémentaire**.
- L'ensemble des évènements élémentaires possibles pour une expérience aléatoire donnée constitue l'**espace fondamental** appelé **univers** ou **univers des possibles** noté Ω.

Exemple:

Lors d'un contrôle sanguin, l'ensemble des résultats possibles si l'on s'intéresse

1. au groupe sanguin et au facteur rhésus d'un individu est Ω={A+, A-, B+, B-, AB+, AB-, O+, O-}

2. au nombre de globules blancs $\Omega=\mathbb{N}^*=\{1,2,....,n,...\}$

3. au taux de glycémie $\Omega=[0;15]$ au-delà de 15, l'individu n'est plus en état de subir une prise de sang.

Ainsi pour une même épreuve, l'univers Ω peut être **fini** (toutes les éventualités sont connues : cas 1) ou **infini** (toutes les éventualités ne sont pas connues : cas 2 et 3). Dans ces deux derniers cas, l'univers peut être **dénombrable** si on peut numéroter les éventualités connues (cas 2) ou bien **continu** comme dans le cas du taux de glycémie (cas 3).

Un <u>événement</u> quelconque A est un ensemble d'évènements élémentaires et constitue **une partie de l'univers** des possibles Ω dont on sait dire à l'issue de l'épreuve s'il est réalisé ou non.

Si $\omega \in A$, alors A **est réalisé**. Mais si $\omega \notin A$ alors A **n'est pas réalisé** et c'est $\bar{A}$, l'**événement contraire** qui est réalisé. Un événement est donc une assertion relative aux résultats d'une expérience.

Il est possible qu'un événement ne soit constitué que **d'un seul** événement élémentaire.

Les évènements sont représentés par des lettres majuscules A,B,C,A1,A2, etc.

Exemple:

Dans l'exemple concernant les groupes sanguins,

- l'événement A « l'individu est de rhésus positif » est représenté par : A = {A+, B+, AB+, O+} avec $A \subset \Omega$ l'événement B « l'individu est donneur universel » est représenté par : B ={O-}un seul événement élémentaire.

Dans le cadre de cet exemple, l'événement A est réalisé si le résultat du typage donne l'un des 4 groupes sanguins A+,B+,AB+,O+.

Remarque : Pour ce même exemple, le résultat « la glycémie vaut √2 » **ne constitue pas un événement** car il est impossible de savoir s'il est réalisé ou non. Toute partie de Ω n'est pas forcément un événement. Ainsi il faut toujours définir après avoir déterminé l'univers Ω, l'ensemble des évènements $\varepsilon(\Omega)$.

Si Ω est fini, chaque partie A de l'univers $\Omega(A \subset \Omega)$ est constituée d'un nombre fini d'éventualités et dans ce cas l'ensemble des évènements est tel que :

$\varepsilon(\Omega)=P(\Omega)$ l'univers des possibles

Dans le cadre de ce cours, nous nous placerons dans le cas où l'ensemble des évènements de l'univers Ω est clairement défini.

1.1. Evènements remarquables

L'**événement impossible**, noté $\emptyset$ est l'événement qui **ne peut être réalisé** quelle que soit l'issue de l'épreuve. Bien que constitué d'aucune éventualité, $\emptyset$ est considéré comme un événement : $\emptyset \in (\Omega)$

L'**événement certain**, noté Ω **est toujours réalisé** quelle que soit l'issue de l'épreuve. Il est constitué de toutes les éventualités et l'on impose que ce soit un événement : $\Omega \in (\Omega)$

L'**événement contraire** ou complémentaire d'un événement A, noté CA ou $\bar{A}$ est l'événement qui est réalisé si et seulement si A ne l'est pas. Il est donc constitué des évènements élémentaires ω qui ne sont pas dans A.

$\omega \in \bar{A} \Leftrightarrow \omega \notin A$

Le complémentaire CA ou $\bar{A}$ correspond à la *négation* logique **non–A**.

Exemple:

Dans l'exemple concernant les groupes sanguins, l'événement contraire de A « l'individu est de rhésus positif » est constitué des évènements élémentaires suivant : $\bar{A}$={A-, B-, AB-, O-}

Par définition, on obtient les relations suivantes :

$\bar{\bar{A}} = A$

$\bar{\emptyset} = \Omega$

$\bar{\Omega} = \emptyset$

1.2. Opérations sur les évènements

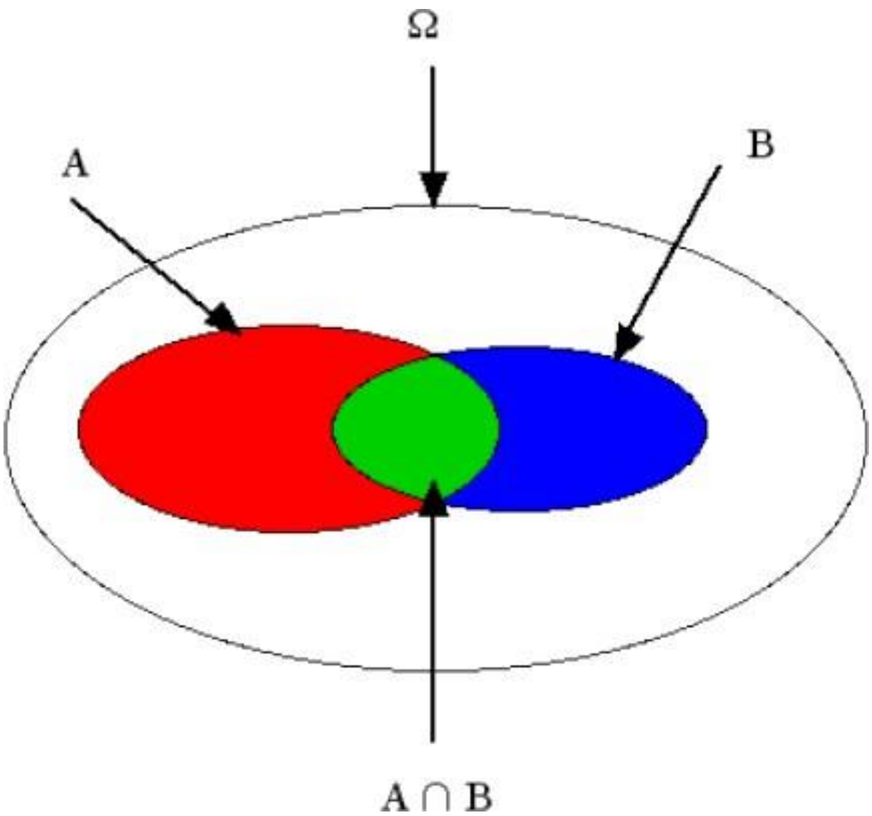
Si l'on considère simultanément la réalisation de deux évènements A et B, il est possible d'effectuer des opérations sur ces ensembles.

2.3.1. L'intersection de deux évènements

On appelle **intersection** de deux évènements A et B, l'événement qui est réalisé si et seulement si A **et** B le sont. Il est donc constitué des éventualités appartenant à la fois à A **et** B. C'est un événement noté $A \cap B$ tel que : $\forall A, B \in (\Omega)$, $A \cap B \in (\Omega)$ avec $\omega \in A \cap B \Leftrightarrow (\omega \in A$ et $\omega \in B)$

L’intersection $A \cap B$ correspond à la *conjonction* logique « A **et** B ».

Remarque : L’univers des possibles Ω n’étant pas limité uniquement aux évènements A (parties **rouge** et **verte**) et B (parties **bleu** et **verte**), l’événement complémentaire $\bar{A}$, est formé des parties **bleu** et **blanche**.



Exemple: Dans l’exemple concernant les groupes sanguins, si à l’événement A « l’individu est de rhésus positif » , on ajoute l’événement B « l’individu possède l’allèle B », l’intersection de ces deux évènements donne : $A \cap B = \{B+, AB+\}$

Deux évènements A et B sont **incompatibles** ou **disjoints**, s’ils ne peuvent être réalisés simultanément. On a alors : $A \cap B = \emptyset$

Quelques propriétés de l’intersection ($\cap$) :

$A \cap \bar{A} = \emptyset$ évènements incompatibles

Ω∩A=A élément neutre (Ω)

∅∩A=∅ élément absorbant (∅)

A∩B=B∩A commutativité

A∩(B∩C)=(A∩B)∩C associativité

A∩(B∩C)=(A∩B)U(A∩C) distributivité avec la **réunion** (U)

2.3.2. La réunion de deux évènements

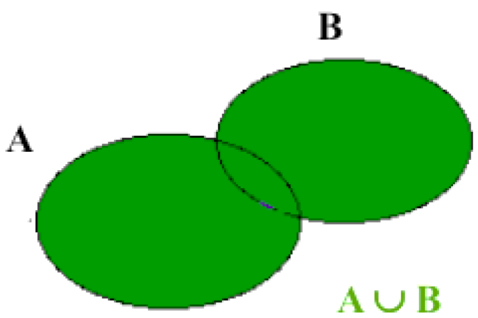
On appelle **réunion** de deux évènements A et B, l'événement qui est réalisé si et seulement si A **ou** B est réalisé. Il est donc constitué des éventualités appartenant à A **ou** B.

C'est un événement noté AUB tel que :

∀A, B∈(Ω), AUB∈(Ω)

avec ω∈AUB⇔(ω∈A ou ω∈B)

La réunion AUB correspond à la *disjonction* logique « A **ou** B ».



Remarque : La **réunion** de deux évènements n'**est pas la somme algébrique des évènements** dans la mesure où la zone de recouvrement n'est pas comptabilisée deux fois.

Exemple : Dans l'exemple concernant les groupes sanguins, si à l'événement AA « l'individu est de rhésus positif », on ajoute l'événement BB « l'individu possède l'allèle B », la réunion de ces deux évènements donne : AUB={A+,B+,B-,AB+,AB-,O+}

Quelques propriétés de la réunion ($\cup$):

$A \cup \bar{A} = \Omega$ évènements complémentaires

$\emptyset \cup A = A$ élément neutre ($\emptyset$)

$\Omega \cup A = \Omega$ élément absorbant (Ω)

$A \cup B = B \cup A$ commutativité

$A \cup (B \cup C) = (A \cup B) \cup C$ associativité

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ distributivité avec la **intersection** ($\cap$)

Selon les lois de **Morgan**, nous avons :

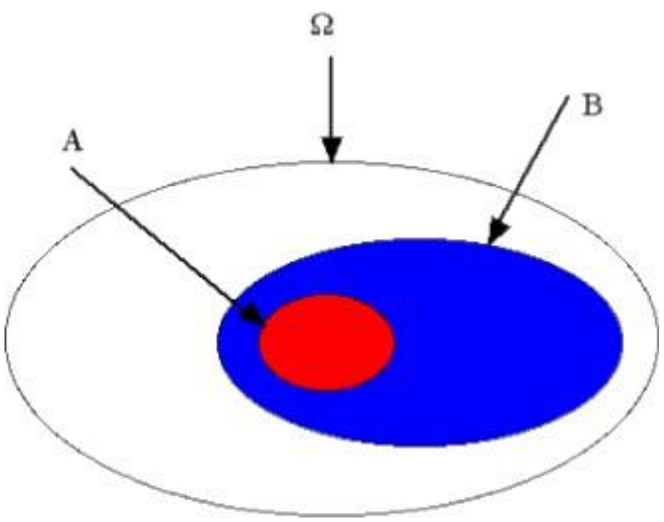
$\overline{A \cap B} = \bar{A} \cup \bar{B}$ ce qui correspond à la partie hachurée sur ce **graphe**.

$\overline{A \cup B} = \bar{A} \cap \bar{B}$ ce qui correspond à l'ensemble vide $\emptyset$ lorsque l'univers des possibles Ω n'est constitué que des évènements A et B (**graphe**).

2.3.3. L'inclusion d'un événement

Un événement A entraîne un événement B si la réalisation de A implique celle de B. On dit que l'événement A est **inclus** dans l'événement B.

$A \subset B$



L'***implication*** logique « $A \Rightarrow B$» se traduit par l'inclusion $A \subset B$.

Exemple : Soit une urne contenant des billes rouges unies et des billes vertes unies et striées. Si l'on note A l'événement « obtention d'une bille striée » et B l'événement « obtention d'une bille verte », la réalisation de A implique la réalisation de B car A est inclus dans B.

2.4. Système complet d'évènements

$A_1, A_2, \ldots, A_n$ forment un système complet d'évènements si les **parties** $A_1, A_2, \ldots, A_n$ de Ω constituent une **partition** de Ω telle que :

$\forall i \; A_i \neq \emptyset$

$\forall i \neq j \; A_i \cap A_j = \emptyset$

$\bigcup_i A_i = \Omega$

Un **système complet d'évènements** est formé de toutes les parties de Ω, c'est-à-dire des familles d'évènements **2 à 2 incompatibles** dont la réunion constitue l'événement certain Ω.

Le nombre de partitions possibles dans un ensemble fini de n évènements est : si $Card(\Omega)=n$ alors $Card\ P(\Omega)=2^n$

3. Probabilités

Le passage d'une description de type ensembliste des phénomènes aléatoires à l'élaboration d'un véritable **modèle mathématique** se fait en introduisant les mesures de probabilité.

3.1 Définitions

3.1.1 Concept mathématique

On appelle **probabilité** P toute **application** de l'ensemble des évènements Ω dans l'intervalle [0,1], tel que :

$P:\varepsilon(\Omega)\rightarrow[0,1]$

$A\mapsto P(A)$

satisfaisant les propriétés (ou axiomes) suivantes

(P1)$\forall(A)\in(\Omega)P(A)\geq 0$

(P2)$P(\Omega)=1$

(P3)$\forall A,B\in(\Omega)$ si $A\cap B=\emptyset$ alors $P(AUB)=P(A)+P(B)$

Remarque : Le concept mathématique de ***probabilité*** modélise les notions intuitives de *proportion* et de ***fréquence***. Si l'on avance que la probabilité d'être immunisé contre la tuberculose est de 0,8, on modélise le fait qu'environ 80 % de la population est immunisé contre la tuberculose.

3.1.2. Probabilités combinatoires

Soit Ω un espace fondamental fini constitué de N **évènements élémentaires** sur lequel on fait l'hypothèse d'**équiprobabilité** de réalisation des N évènements élémentaires. On suppose ainsi que tous les évènements élémentaires ont « la même chance » de se réaliser. Dans ce cas la probabilité pipi d'un événement élémentaire quelconque ωI est telle que :

$pi=1/N$ avec $pi=P(\omega I)$

satisfaisant:

(P1) avec $\forall I\ p_i \geq 0$

(P2) $\sum_i p_i = 1$

Soit A un événement quelconque constitué de **k évènements élémentaires** de, on en déduit :

$P(A)=k/N$ avec $P(A)=\sum_{\omega_i \in A} p_i$

Cette formule s'énonce souvent comme :

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

Cette formule permet de ramener les calculs de probabilités à des décomptes d'évènements élémentaires effectués par des techniques d'**analyse combinatoire** qui ne sont pas des probabilités.

Exemples :

1. En tapant 5 lettres au hasard sur une machine à écrire (possibilité de taper plusieurs fois sur la même touche), la probabilité d'obtenir le mot « lutte » est d'**une chance sur 12 millions**. En effet il y a exactement 11 881 376 mots de 5 lettres possibles (?theme=proba&chap=1#Arrangement avec répétitions) avec répétition).

2. La probabilité d'obtenir un multiple de trois lors du lancé d'un dé à 6 faces, non pipé est : $A=\{3,6\}$ d'où $P(A)=2/6=1/3$ avec $k=2$ et $p_i=1/6$

3.1.3. Loi des grands nombres

Si l'on répète N fois une expérience dans laquelle la probabilité d'apparition d'un événement A est P, **la fréquence** de cet événement au cours des N expériences, kN tend vers P lorsque N tend vers l'infini. $N \to \infty \Rightarrow kN \to P$

Lorsque le nombre d'épreuves augmente indéfiniment, les fréquences observées tendent vers les probabilités et les distributions observées vers **les lois de probabilité**.

Exemple:

Lors d’un croisement entre plantes hétérozygotes Aa pour un caractère à dominance stricte (allèle A, forme sauvage et allèle a, forme mutée), on examine successivement deux **échantillons** de plantes résultant de ce croisement.

	N=40 fleurs		N=1000 fleurs		
	Effectifs	Fréquences	Effectifs	Fréquences	Probabilités attendues
Phénotype sauvage	29	0,725	754	0,754	0,750
Phénotype mutant	11	0,275	246	0,246	0,25

Il est aisé de vérifier que les fréquences pour les deux phénotypes possibles sont plus proches des probabilités attendues sous le modèle de croisement mendélien pour l’échantillon de très grande taille. L’écart entre les fréquences observées et les probabilités attendues peut être tester à l’aide du **test du khi-deux**.

3.1.4. Espace probabilisé

Nous définirons un **espace probabilisé** en utilisant l’axiomatique de **Kolmogorov**,

Définition 1 :

On appelle probabilité sur (Ω,C) une application P de C dans l’intervalle $[0,1]$ telle que :

- $P(\Omega)=1$
- Pour tout ensemble dénombrable d’évènements incompatibles 2 à 2, on a : $P(\bigcup_i A_i)=\sum_{i=1}^{n} P(A_i)$

Définition 2 :

On appelle **espace probabilisé**, le triplet (Ω,C,P)

<>

Ainsi un espace probabilisé désigne un espace fondamental et ses évènements, muni d’une mesure de probabilités.

3.2 Propriétés des probabilités

Des axiomes précédents découlent les propriétés **additives** des probabilités, d'usage permanent.

3.2.1. Additivité

- **Cas d'évènements incompatibles**

Si A1,A2,...,Ai,...,An sont n évènements **incompatibles deux à deux** ($Ai \cap Aj = \emptyset$ si $i \neq j$) alors : $P(A1 \cup A2 \cup ... \cup Ai \cup ... \cup An) = P(A1)+P(A2)+...+P(Ai)+...+P(An)$

La probabilité de la **réunion** d'un ensemble fini ou dénombrable d'évènements

2 à 2 incompatibles est égale à la somme de leur probabilité d'où :

$P(\bigcup iAi)=\sum i=1nP(Ai)P(\bigcup iAi)=\sum i=1nP(Ai)$

voir **Démonstration**

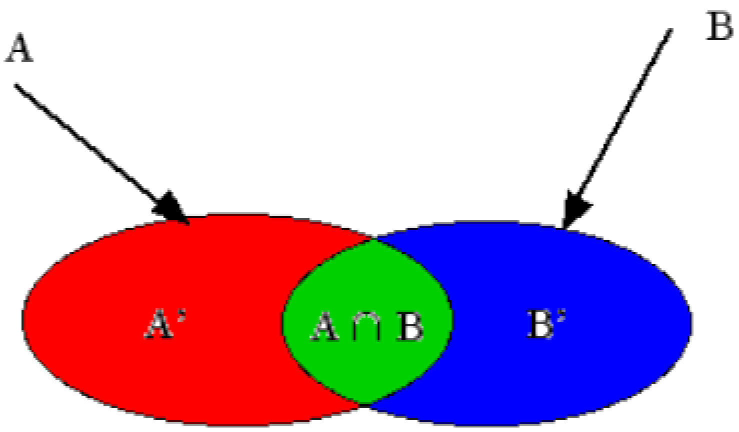
- **Cas de deux évènements quelconques**

Si A et B sont deux **évènements quelconques**, alors :

$P(A \cup B)=P(A)+P(B)-P(A \cap B)$

Voici pourquoi :

A et B étant deux évènements quelconques, $(A \cap B) \neq \emptyset$, ces évènements peuvent se décomposer comme la réunion de deux évènements incompatibles :



Alors :

1. $A=A'\cup(A\cap B)A=A'\cup(A\cap B)$ avec $A'\cap(A\cap B)=\emptyset A'\cap(A\cap B)=\emptyset$

alors $P(A)=P(A')+P(A\cap B)P(A)=P(A')+P(A\cap B)$

et $P(A')=P(A)-P(A\cap B)P(A')=P(A)-P(A\cap B)$

2. $B=B'\cup(A\cap B)B=B'\cup(A\cap B)$ avec $B'\cap(A\cap B)=\emptyset B'\cap(A\cap B)=\emptyset$

d'où $P(B')=P(B)-P(A\cap B)$ $P(B')=P(B)-P(A\cap B)$

3. $P(A\cup B)=P(A')+P(A\cap B)+P(B')P(A\cup B)=P(A')+P(A\cap B)+P(B')$

d'où $P(A\cup B)=P(A)+P(B)-P(A\cap B)$ $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

Exemple :

Dans l'exemple du lancer d'un dé à 6 faces, non pipé, on considère l'événement A « le résultat est pair » et l'événement B « le résultat est un multiple de trois ».

On a alors :

$A=\{2,4,6\}$ et $B=\{3,6\}$ donc $A\cup B=\{2,3,4,6\}$ et $A\cap B=\{6\}$

avec $P(A)=3/6$ $P(B)=2/6$ $P(A\cup B)=4/6$ $P(A\cap B)=1/6$

on vérifie alors que : $P(A\cup B)= P(A) + P(B) - P(A\cap B) = 3/6+2/6-1/6 = 4/6$

3.2.2 Evènement contraire

Si A est un événement quelconque, alors $P(A)^{‾}=1−P(A)$

Voici pourquoi :

Nous avons vu précédemment que

$A∪A^{‾}=Ω$ et $A∩A^{‾}= ∅$ Propriétés de la **réunion** et de **intersection**

$P(A∪A^{‾})= P(A) + P(A^{‾})$ $P(A∪A^{‾}) = P(A) + P(A^{‾})$ Propriétés d'additivité des probabilités

d'où $P(Ω) =1= P(A) + P(A^{‾})$ ainsi $P(A^{‾}) = 1−P(A)$ $P(A^{‾}) = 1−P(A)$

Exemple: La probabilité lors du lancer d'un dé non pipé d'obtenir « plus de 2 » se traduit par A={3,4,5,6} et A={1,2} d'où P(A) = 1–P(A) = 1–2/6 = 4/6 = 2/3

Remarque: L'application de cette propriété est très utile lorsque le nombre d'évènements élémentaires de A , k , est important et que le calcul des probabilités p i est fastidieux.

3.2.3 Evènement impossible

$P(∅)=0$

Voici pourquoi:

Nous avons vu précédemment que $∅∪Ω=Ω$ **élément neutre**

$P(∅∪Ω)=P(∅)+P(Ω)$ Propriétés d'additivité des probabilités

d'où $P(Ω)=P(∅)+P(Ω)$ ainsi $P(∅)=0$

3.2.4 Inclusion

Si $A⊂BA⊂B$ alors $P(A)≤P(B)$

Voici pourquoi :

si $B = B' \cup A$ avec $B' \cap A = \emptyset$

alors $P(B) = P(B' \cup A) = P(B') + P(A)$

d'où $P(A) \leq P(B)$

avec $P(A) = P(B)$ lorsque $P(B') = 0$

3.3 Indépendance statistique

3.3.1 Définition

L'hypothèse d'**indépendance** entre évènements et plus généralement entre épreuves successives est un préalable lors de l'établissement des **lois de probabilités**

On dit que deux évènements A et B sont **indépendants** si l'on a :

$P(A \cap B) = P(A)P(B)$

Ainsi si A et B sont deux évènements **statistiquement indépendants**, la probabilité de la réalisation conjointe de ces deux évènements est le produit de leur probabilité respective.

Remarque :

Il ne faut pas confondre évènements **indépendants** et évènements **incompatibles**.

Supposons A et B à la fois indépendants et incompatibles. On a alors :

$P(A \cap B) = P(A)\ P(B)$ indépendants

$P(A \cap B) = P(\emptyset) = 0$ incompatibles

d'où nécessairement $P(A) = 0$ ou $P(B) = 0$

Exemples :

1. Dans l'exemple du lancer d'un dé à 6 faces, non pipé, les deux évènements : A « le résultat est pair » et B « le résultat est un multiple de trois » sont **statistiquement indépendants**.

En effet, soit $A=\{2,4,6\}$ $B=\{3,6\}$ $A\cap B=\{6\}$ ainsi $P(A)=3/6$ $P(B)=2/6$ $P(A\cap B)=P(A)=3/6$ $P(B)=2/6$ $P(A\cap B)= 1/6$ on vérifie alors que : $P(A\cap B)=P(A)P(B) = (3/6)\times(2/6) = 6/36 = P(A\cap B)=P(A)P(B)=(3/6)\times(2/6)=6/36= 1/6$

2. Si l'on considère une famille de deux enfants, les deux évènements : A « enfants de sexe différent » et B « au plus une fille » **ne sont pas statistiquement indépendants**.

En effet, l'espace probabilisé Ω, contient 4 évènements élémentaires (si l'on considère une famille ordonnée),

$\Omega=A\cup B=\{GG,GF,FG,FF\}$

avec $A=\{GF,FG\}$, $B=\{GG,GF,FG\}$ et $A\cap B=\{GF,FG\}$

d'où sous l'hypothèse d'équiprobabilité : $P(A) =1/2$, $P(B) =3/4$ et $P(A\cap B)= 1/2$

On vérifie alors que : $P(A\cap B) \neq P(A)P(B) = (1/2)\times(3/4) = P(A\cap B) \neq P(A)P(B) = (1/2)\times(3/4) = 3/8 = 1/2$

3.3.2 Propriétés

Les propriétés associées à l'indépendance sont :

1. si A est un évènement **quelconque**,

A et Ω sont indépendants : $A\cap\Omega=A$ **élément neutre**

$P(A\cap\Omega)= P(A)\ P(\Omega) = P(A)\ P(A\cap\Omega) = P(A)\ P(\Omega) = P(A)$ car $P(\Omega)=1$

A et ø sont indépendants : $A\cap\emptyset=\emptyset$ **élément absorbant**

$P(A\cap\emptyset)=P(A)\ P(\emptyset) = P(\emptyset)\ P(A\cap\emptyset) = P(A)\ P(\emptyset) = P(\emptyset)$ car $P(\emptyset)=0$

2. si A et B sont deux évènements **quelconques**,

AA et BB sont indépendants si et seulement si A et $\bar{B}$ ou ($\bar{A}$ et B) sont indépendants (**démonstration**).

A et B sont indépendants si et seulement si $\bar{A}$ et $\bar{B}$ le sont.

3.3.3 Généralisation à n évènements

nn évènements ($n\geq 2$), $A1,A2,...,Ai,..,$ An sont dit **indépendants** dans leur ensemble (ou mutuellement indépendants) si on a : $P(A1\cap A2\cap ...\cap Ai\cap ...\cap An) = P(A1)\times P(A2)\times ...\times P(Ai)\times\ \times P(An)$

$P(\cap iAi)=\prod i=1\ nAi$

Remarque :

n évènements peuvent être indépendants deux à deux,

$[P(Ai\cap Aj)=P(Ai)xP(Aj)]$ avec $i\neq j$ sans être indépendants au sens de la définition ci-dessus.

Exemple :

On jette deux dés non pipés et on considère les évènements suivants :

A1 « le premier dé donne un nombre pair »

A2 « le deuxième dé donne un nombre pair »

A3 « la somme des deux lancers est paire »

Le nombre d'événements élémentaires est :

$card(\Omega)=36$

Les 3 évènements A1, A2et A3 sont 2 à 2 indépendants mais ne sont pas indépendants dans leur ensemble. En effet :

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Les probabilités associées aux 3 évènements sont :

$P(A1)=1/2$; $P(A2)=1/2$; $P(A3)=1/2$

$P(A1\cap A2)=9/36=1/4=P(A1)\ P(A2)$

$P(A1\cap A3)=9/36=1/4=P(A1)P(A3)$

$P(A2\cap A3)=9/36=1/4=P(A2)P(A3)$

$P(A1\cap A2\cap A3)=9/36=1/4\neq P(A1)P(A2)P(A3)=1/8$

Les cases grisées représentent les évènements élémentaires réalisés dans le cadre soit $A1\cap A2$ ou $A1\cap A3$ ou $A2\cap A3$ ou $A1\cap A2\cap A3$.

4. Probabilités conditionnelles

4.1. Définition

Soit deux évènements A et B d'un espace probabilisé Ω avec $P(B)\neq 0$, on appelle probabilité conditionnelle de l'évènement « A si B» (ou « A sachant B »), le quotient

$P(A/B)=P(A\cap B)\ P(B)$ notée $PB(A)$

On définit ainsi une probabilité sur Ω au sens de la **définition** donnée précédemment.

Théorème :

Soit B un évènement de probabilité non nulle, alors :

$PB:\varepsilon(\Omega)\rightarrow[0,1]$

$A\mapsto P(A/B)=P(A\cap B)P(B)$ est une probabilité sur Ω

Voici pourquoi :

(P1) $\forall A\in(\Omega)\ P(A|B)\geq 0$ (quotient de deux réels positifs)

(P2) $P(\Omega|B)=P(\Omega\cap B)P(B)=P(B)P(B)=1$ car $\Omega\cap B=B\Omega\cap B=B$ car Ω **élément neutre**

(P3) si $(A1\cap A2)=\emptyset$ $(A1\cap A2)=\emptyset$, $PB(A1\cup A2) = P[(A1\cup A2)\cap B]\ P(B)=P[(A1\cap B)\cup(A2\cap B)]\ P(B)=P[(A1\cap B)+(A2\cap B)]\ P(B)=PB(A1)+PB(A2)\ PB(A1\cup A2) = P[(A1\cup A2)\cap B]\ P(B)=P[(A1\cap B)\cup(A2\cap B)]$ **additivité indépendants**

Remarque : La probabilité P(A) est appelée la probabilité a priori et P(A|B) ou PB(A) la probabilité a posteriori car sa réalisation dépend de la réalisation de B.

On observe les relations suivantes :

$P(A|A)=1$

Si $B\subset AB\subset A$, alors $A\cap B = B$ et donc $P(B|A)=P(B)\ P(A)\ P(B|A) = P(B)\ P(A)$

Exemple :

Soit un croisement entre hétérozygotes Aa pour un caractère à dominance stricte, quelle est la probabilité d’obtenir à la génération suivante parmi les individus de phénotype A, un individu homozygote ?

L’ensemble des évènements élémentaires est : $\Omega=\{AA,Aa,aA,aa\}$ Si h=h= homozygote et $\bar{h}=\bar{h}$= hétérozygote

$P(h|A) = P(h∩A)\ P(A) = ¼\ ¾ = P(h|A) = P(h∩A)\ P(A) = ¼\ ¾ = 1/3\ 1/3$ **probabilité a *posteriori***

La probabilité ***a priori*** d'obtenir un homozygote est **1/4**.

4.2. Probabilités composées

Théorème :

Soit deux évènements A et B d'un espace probabilisé . Alors,

$P(A∩B) = P(B|A)\ P(A) = P(A|B)\ P(B)$ **Formule des probabilités composées**

Voici pourquoi:

Par définition, $P(A|B) = P(A∩B)\ P(B)\ P(A|B) = P(A∩B)\ P(B)$ d'où $P(A∩B) = P(A|B)\ P(B)\ P(A∩B) = P(A|B)\ P(B)$

Par symétrie, $P(A∩B) = P(A|B)\ P(B) = P(B|A)\ P(A)$

Si A et B sont deux évènements **indépendants** et que $P(B) \neq 0$ alors ceci équivaut à affirmer que $PB(A)=P(A|B)=P(A)$

Lorsque deux évènements sont **indépendants**, le fait que l'un des évènements soit réalisé, n'apporte aucune information sur la réalisation de l'autre. Dans ce cas la probabilité conditionnelle PB(A) (*a posteriori*) est égale à la probabilité P(A) (*a priori*).

Voici pourquoi :

La formule des probabilités composées donne $P(A∩B) = P(A|B)\ P(B)$

L'indépendance statistique entre A et B équivaut à $P(A∩B)=P(A)\ P(B)$

d'où la relation $P(A|B)=P(A)$

Si A et B sont deux évènements **indépendants** alors ceci équivaut à affirmer que $PB(A)=PB^{-}(A)=P(A)\ PB(A) = PB^{-}(A) = P(A)$

Lorsque deux évènements sont **indépendants**, la probabilité conditionnelle de A est la même que ce soit B ou B qui est réalisé (voir **démonstration**).

Dans l'exemple du lancer d'un dé à 6 faces, non pipé, les deux évènements : A « le résultat est pair » et B « le résultat est un multiple de trois » sont **indépendants** (voir **exemple**). Ainsi la probabilité que la face soit paire sachant que c'est un multiple de 3 est :

si $A=\{2,4,6\}$ $B=\{3,6\}$ $A\cap B=\{6\}$

et $P(A)=3/6$ $P(B)=2/6$ $P(A\cap B)=1/6$

$P(A|B)=P(A\cap B)$ $P(B)$ $=1/6$ $1/3$ $=$ $P(A|B)$ $=$ $P(A\cap B)$ $P(B)$ $=$ $1/6$ $1/3= 1/2 = P(A)$ d'où la relation $P(A|B) = P(A)$

4.3. Probabilités totales

Théorème :

Si $\{A1,A2,....,Ai,.....,An\}$ est un système complet d'évènements , quel que soit l'évènement B, alors : $P(B)=P(B|A1)P(A1)+P(B|A2)P(A2)+ ...... +P(B|An)P(An)$

$P(B) = \sum i = 1n\ P(B|Ai)\ P(Ai)$ $P(B) =\sum i=1nP(B|Ai)\ P(Ai)$ **Formule des probabilités totales**

Voici pourquoi :

Si $i\neq j$, alors $(Ai\cap B)\cap(Aj\cap B) = (Ai\cap Aj)\cap B = \emptyset$

Grâce à la distributivité , on a :

$(A1\cap B)\cup(A2\cap B)\cup......\cup(An\cap B) = (A1\cup A2\cup ....\cup An) \cap B = \Omega\cap B = B$

Grâce à l'additivité, on a :

$P(B)=P(A1\cap B)+P(A2\cap B)+.....+P(An\cap B)$

Grâce à la formule des probabilités composées, on a :

$P(A1\cap B)=P(B|A1)P(A1)P(A1\cap B)=P(B|A1)P(A1)$

d'où $P(B)=P(B|A1)P(A1)+P(B|A2)P(A2)+ ....+P(B|An)P(An)$

Exemple :

Une population animale comporte 1/3 de mâles et 2/3 de femelles. L'albinisme frappe 6 % des mâles et 0,36 % des femelles. La probabilité pour qu'un individu pris au hasard (dont on ignore le sexe) soit albinos est :

Si A = {mâle} et $\bar{A}$ = {femelle} constitue un système complet d'évènements

B = {albinos} et $\bar{B}$ = {non albinos}

sachant que $P(B)=P(B|A)P(A)+P(B|\bar{A})P(\bar{A})$

alors $P(B)=(0,06\times1/3)+(0,0036\times2/3)=$ **0,0224**

soit **2,24% d'albinos** dans cette population.

4.4. Le théorème de **Bayes**

Un corollaire au théorème des probabilités totales est connu sous le nom de **formule de Bayes**.

Théorème :

Si $\{A1,A2,....,Ai,.....,An\}$ est un **système complet d'évènements** , et quel que soit l'évènement B tel que $P(B)\neq0$, alors :

$P(Ai|B)=P(B|Ai)P(Ai)P(B|A1)P(A1)+...+P(B|Ai)P(Ai)+...+P(B|An)P(An)$

$P(Ai|B)=P(B|Ai)P(Ai)\sum i=1nP(B|Ai)P(Ai)$

Formule de Bayes

Voici pourquoi :

D'après la formule des **probabilités composées**,

$P(Ai\cap B)=P(Ai|B)P(B)=P(B|Ai)P(Ai)$

D'après la formule des **probabilités totales**,

$P(B)=\sum i=1nP(B|Ai)P(Ai)$

D'après la formule des **probabilités conditionnelles**, $P(Ai|B)=P(Ai\cap B)P(B)$

d'où $P(Ai|B)=P(B|Ai)P(Ai)\sum i=1nP(B|Ai)P(Ai)$

Remarque :

La formule de Bayes est utilisée de façon classique pour calculer des **probabilités de causes** dans des diagnostics (maladies, pannes, etc.). L'application du théorème de Bayes est à la base de toute une branche de la statistique appelée **statistique bayesienne**.

Exemple :

Dans une population pour laquelle 1 habitant sur 100 est atteint d'une maladie génétique A, on a mis au point un test de dépistage. Le résultat du test est soit positif (T) soit négatif ($\bar{T}$). On sait que : $P(T|A)=0,8$ et $P(\bar{T}|\bar{A})=0,9$

On soumet un patient au test. Celui-ci est positif. Quelle est la probabilité que ce patient soit atteint de la maladie A soit PT(A) ou P(A T)?

D'après la formule de Bayes :

$P(A|T)=P(A\cap T)P(T)=P(T|A)P(A)P(T|A)P(A)+P(T|\bar{A})P(\bar{A})$

d'où $P(A/T)=0,01\times 0,80,8\times 0,01+0,1\times 0,99 = 0,075$

Ainsi **avant le test**, la probabilité d'**être malade** était de P(A) = **0 ,01** (probabilité a priori)

et après le test **la probabilité d'être malade** est de P(A|T) = **0,075** (probabilité *a posteriori*).

Ainsi le test apporte un supplément d'information.