

Jean-Paul Guillaud

Comprendre les probabilités et la statistique



*Ce livre est dédié à toutes celles et ceux qui
m'ont aidé et encouragé, en particulier à ma
femme Alice et à nos enfants Maëlle et Yann.*

Avant-propos

Sans avoir la prétention de prédire l'avenir – je laisse cette responsabilité aux voyants de toutes les couleurs –, le **calcul des probabilités** sert à déterminer la proportion des cas dans lesquels **un événement lié au hasard** peut se produire.

Le concept de hasard n'est apparu dans les Mathématiques qu'au XVII^e siècle. Il fut alors étudié par Blaise Pascal et Pierre de Fermat (1654) puis par Christian Huygens (1656), John Graunt (1662) et Edmond Halley (1693). Les premiers travaux portèrent sur l'observation du jeu de dés. Les suivants eurent des motivations moins ludiques (Graunt étudia la mortalité humaine, Halley les assurances-vie). La première formulation mathématique est due à Huygens avec un traité complet. À la suite de ces pionniers, Lagrange, Laplace et, surtout, Johann Carl Friedrich Gauss formalisèrent la prolongation naturelle du calcul des probabilités (la *statistique mathématique*) en étudiant la **théorie des erreurs**.

Dans les années 1970, je donnais un cours de D.E.A. (à Grenoble) sur « Probabilités et statistique ». Dans la fougue de ma toute première expérience pédagogique, je m'étais assigné comme but que ce cours soit irréprochable du point de vue mathématique... Hélas, quelques années plus tard, mes étudiants m'ont avoué qu'il était d'une aridité décourageante !

Après ce cuisant échec, cuisant surtout pour mes étudiants, je me suis plongé dans l'abondante littérature sur le sujet pour voir comment mes « pairs » faisaient passer le message et quels étaient les bonnes formules et les bons exemples. Je prie donc les auteurs des nombreux ouvrages cités en référence de ne pas se formaliser de mes emprunts mais plutôt de les considérer comme un hommage à leur pédagogie.

Je me suis remis à la tâche avec, cette fois, pour but de présenter à un étudiant, pourvu *d'un minimum de culture mathématique*, un ouvrage sous la forme d'un cours plus compréhensible et directement utilisable sur les probabilités et la Statistique.

J'ai choisi, pour ne pas alourdir ce cours, et au risque d'être incomplet, d'ignorer la première étape habituelle concernant la Statistique descriptive

(dont le propos est de rassembler et d'analyser les données numériques concernant un groupe d'objets ou d'individus), donc de décrire mais pas d'expliquer. Ce livre débute donc par le **Calcul des Probabilités**, (l'étude des mécanismes aléatoires et des propriétés des objets qui découlent de ces mécanismes), puis les **Dénombrements** (c'est-à-dire le nombre de possibilités de classer (de combiner) divers objets entre eux.)

La description des **lois du hasard** complète l'information nécessaire pour entrer plus en avant dans la Statistique Mathématique.

Celle-ci concerne l'étude de *l'induction statistique*. C'est-à-dire l'analyse de l'information obtenue à travers un mécanisme aléatoire. Elle s'appuie, évidemment, sur le calcul des probabilités, mais ses préoccupations sont essentiellement concrètes.

Alors que le calcul des probabilités fournit les propriétés générales des observations que peut produire un mécanisme aléatoire supposé connu (raisonnement « *aval* »), la statistique mathématique cherche à préciser le mécanisme aléatoire particulier qui a produit les informations dont on dispose moyennant quelques hypothèses sur l'appartenance de ce mécanisme à une classe connue de mécanismes (raisonnement « *amont* »). Une fois ce mécanisme calculé, il est important de savoir dans quelle mesure le résultat de ce calcul est raisonnable et décrit la réalité observée : il faut donc tester ce résultat (cette hypothèse).

Pour alléger le texte, les démonstrations longues et fastidieuses sont rejetées en appendice.

Note : Pour faciliter la compréhension des ouvrages anglo-saxons ce qui figure en *italique souligné* est le terme équivalent anglais. La collection de ces spécificités anglo-saxonnes est dans l'appendice V.

Ce qui est à retenir et est important est étiqueté avec « **➡➡➡** » ou « **⬅⬅⬅** ». (La réunion de ces (47) concepts et formules se trouve en fin d'ouvrage.)

Ce qui est présent par souci d'exhaustivité et qui peut être sauté sans conséquence pénalisante est étiqueté avec « **??? ➡➡** ».

Des exercices (dont la solution figure en fin de l'ouvrage) sont donnés, à titre d'exemples, pour aider l'assimilation (et la manipulation) des notions nouvelles.

Les figures ont été faites avec PAW. Ce logiciel [Brun 1989] du CERN, extrêmement puissant, est incontournable dans l'analyse des données et la publication des résultats scientifiques.

Je remercie vivement Monsieur Couet (du CERN) pour avoir répondu avec gentillesse à toutes mes questions sur l'emploi de ce logiciel.

Vos questions, suggestions et critiques sont les bienvenues (jean-paul.guillaud@cern.ch ou jeanpaulguillaud@gmail.com).

Première partie

Probabilités

I

Probabilités sur un ensemble fini

I – 1) Définitions

Événement aléatoire

Événement à plusieurs issues, dont on ne peut pas déterminer l'issue mais seulement lui associer une probabilité. Par exemple le lancer d'une pièce de monnaie dont les issues possibles sont les côtés PILE ou FACE. Il est possible d'associer à l'issue PILE (comme à l'issue FACE) une probabilité.

Ensemble fondamental (Ω) ou référentiel ou univers (sample space)

L'ensemble fondamental est l'ensemble des différents résultats possibles d'une expérience stochastique.

Expérience stochastique (ε) ou épreuve

Une expérience stochastique est une expérience qui, réalisée dans les mêmes conditions, donne des résultats différents (par exemple, le lancer d'une pièce de monnaie). Son résultat varie *au hasard*. En principe, elle peut être réalisée indéfiniment.

Événement élémentaire

Un événement élémentaire c'est un seul événement. Par exemple a est élémentaire si et seulement si, a se réalise seul. Il n'est donc pas possible pour un événement élémentaire d'utiliser les relations « et », « ou ». Un événement élémentaire est, donc, *un sous-ensemble de Ω constitué d'un seul élément* (un « singleton »).

I – 2) Rappels sur les ensembles

Dans ce qui suit les majuscules ($A, B, \dots$) désignent des ensembles et les minuscules ($a, b, \dots$) des éléments d'ensemble.

Un *ensemble* est une réunion d'objets ayant, en général, une caractéristique commune (les enfants d'une famille par exemple). Ces objets sont les *éléments* de l'ensemble (dans cet exemple, un enfant est un élément).

$a \in A$ signifie que a est un **élément** de l'ensemble A .

Card(A) le **cardinal** de A , désigne le nombre d'éléments de l'ensemble A .

Exemple de cardinal : Si « A » est l'ensemble des jets d'un dé, « a » est le résultat d'un jet, $Card(A) = 6$.

$A \cup B$ désigne la **réunion** des ensembles A et B . $A \cup B$ est l'ensemble qui contient *tous* les éléments de A et *tous ceux* de B et *seulement ceux-là*.

$A \cap B$ désigne l'**intersection** de A et de B . $A \cap B$ est l'ensemble qui contient les éléments qui sont *à la fois* dans A et dans B .

$A \subset B$ désigne l'**inclusion** de A dans B . $A \subset B$ signifie que *tous* les éléments de A sont *aussi* éléments de B (A est sous-ensemble de B). L'inclusion est *transitive* : Si on a $A \subset C$ et $C \subset B$, alors $A \subset B$. Ce qui s'écrit : $(A \subset C \text{ et } C \subset B \Rightarrow A \subset B)$.

$A \times B$ désigne le **produit cartésien** de A et de B . $A \times B$ est l'ensemble qui contient tous les *couples* d'éléments dont le premier membre est un élément de A et le second membre est un élément de B .

C_B^A désigne le **complémentaire** de A dans B (tous les éléments de B n'étant pas dans A). Si A est sous-ensemble de B , C_B^A désigne l'ensemble des éléments de B qui *ne sont pas dans* A . Le complémentaire de A dans l'ensemble fondamental (C_Ω^A) s'écrit $\bar{A}$ et se dit « non- A ».

Les 2 paragraphes suivants détaillent les lois d'Auguste de Morgan (lois de Morgan).

Complémentaire d'une réunion

Le complémentaire d'une réunion est l'intersection des complémentaires. C'est-à-dire $\overline{(A \cup B)} = \bar{A} \cap \bar{B}$. ◀◀◀

Cette relation peut se démontrer [Lazar 1967] de la façon suivante : La condition (nécessaire et suffisante) pour qu'un élément appartienne à $\bar{A} \cap \bar{B}$ est qu'il appartienne à la fois à $\bar{A}$ et $\bar{B}$ c'est-à-dire qu'il ne soit *ni dans* A *ni dans* B , c'est-à-dire qu'il ne soit pas dans $A \cup B$. Cette dernière assertion s'écrit $\overline{(A \cup B)}$.

Complémentaire d'une intersection

Le complémentaire d'une intersection est la réunion des complémentaires.

C'est-à-dire $\overline{(A \cap B)} = \bar{A} \cup \bar{B}$. ◀◀◀

La démonstration de cette relation se fait de la même façon que ci-dessus (complémentaire d'une réunion).

⇒⇒⇒ Il est facile de mémoriser la règle des complémentaires¹ en systématisant $\overline{\cap} \rightarrow \cup, \overline{\cup} \rightarrow \cap$.

Le calcul des probabilités traite d'événements entre lesquels il existe des relations de nature logique (« et », « ou »). Ces relations se font dans le cadre de l'algèbre de George Boole (ou algèbre d'événements).

I – 3) Rappels sur l'algèbre de Boole

L'algèbre de Boole qui traite de la logique des propositions connaît deux états (VRAI, FAUX). Dans ce qui suit, a et b désignent des propositions.

Loi « et »

La relation « et » est symbolisée mathématiquement par « $\wedge$ ».

a **$\wedge$** b est VRAI si et seulement si a est VRAI *et* si b est VRAI

Exercice 1 : Soit l'ensemble des entiers de 1 à 10 ; a=impair ; b>4 : a **$\wedge$** b= ?

Exercice 2 : même question que pour l'exercice 1 mais pour un jet de dé.

Loi « ou »

La relation « ou » est symbolisée mathématiquement par « $\vee$ ».

a **$\vee$** b est VRAI si et seulement si a est VRAI *ou* si b est VRAI

Exercice 3 : Soit l'ensemble des entiers de 1 à 10 ; a impair, b>4 a **$\vee$** b= ?

Exercice 4 : même question que pour l'exercice 3 mais pour un jet de dé.

Propriétés des relations « et », « ou »

- Associativité de la loi $\wedge$: $a \wedge (b \wedge c) = (a \wedge b) \wedge c = a \wedge b \wedge c$. (Les parenthèses sont inutiles.)
- Commutativité de la loi $\wedge$: $a \wedge b = b \wedge a$. (L'ordre n'a pas d'importance.)
- Distributivité de la loi $\wedge$ par rapport à la loi $\vee$: $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$. (Les parenthèses sont importantes, ces 3 propriétés ressemblent à celles de « + » par rapport à « $\times$ » : $\wedge \leftrightarrow \times, \vee \leftrightarrow +$.). Le sens n'a

¹ Cette règle facilite la programmation FORTRAN (ou tout autre langage informatique) dans l'écriture du contraire de « if(a **et** b) » en donnant « if(non-a **ou** non-b) ». Par exemple, le contraire de « if(i<4 **et** j<5) » est « if(i>=4 **ou** j>=5) ». De même le contraire de « if(a **ou** b) » est if (non-a **et** non-b) ». Par exemple, le contraire de « if(i>4 **ou** j>5) » est « if (i<=4 **et** j<=5) » : Le contraire de « et » est « ou » et réciproquement.

pas d'importance : elle est distributive à gauche comme à droite :
 $(b \vee c) \wedge a = (b \wedge a) \vee (c \wedge a)$.

Relation de « et », « ou » avec l'algèbre d'ensembles

« et » correspond à $\cap$. $\leftarrow \leftarrow \leftarrow$

« ou » correspond à $\cup$. $\leftarrow \leftarrow \leftarrow$

Comme il est précisé ci-dessus (page 9), il n'est pas possible d'utiliser les relations « et » et « ou » avec un événement élémentaire.

I – 4) Probabilité

La probabilité est une notion délicate à définir. Pour des raisons de clarté, elle est introduite en deux étapes [Calot 1967] : D'abord une définition intuitive dans un cas extrêmement simple puis dans un cas général.

Définition intuitive

Lors du lancer, au hasard, d'une pièce de monnaie *symétrique*, on ne peut pas savoir sur quelle face elle va tomber. Il y a seulement deux possibilités : obtenir PILE ou FACE. Comme la pièce est symétrique et qu'elle est lancée au hasard, il n'y a pas de raison qu'elle tombe sur PILE plutôt que sur FACE. C'est-à-dire que les deux possibilités sont également probables ou **équiprobables**. Dans ce cas, à chaque possibilité est attribuée² une probabilité $1/2$: il y a une chance sur deux que la possibilité PILE (ou la possibilité FACE) apparaisse. Dans ce cas, l'ensemble fondamental est l'ensemble de toutes les possibilités, il a un cardinal³ 2. (Il y a seulement 2 possibilités.)

$\Rightarrow \Rightarrow \Rightarrow$ D'une façon générale, lorsque les possibilités *sont équiprobables* et en nombre fini, l'ensemble fondamental Ω a un cardinal égal au nombre de possibilités et la probabilité d'apparition *d'une* possibilité est égale à $1/\text{card}(\Omega)$.

Exercice 5 : Dans un jeu de 52 cartes battues, quel est le cardinal des possibilités « tir d'une carte donnée » ?

Exercice 6 : Dans un jeu de 52 cartes battues, quelle est la probabilité de tirer l'as de carreau P(as de carreau) ?

Par extension, on peut définir la probabilité de la possibilité A en divisant le nombre de possibilités équiprobables définissant la possibilité A *par le nombre total de possibilités équiprobables*.

En ce qui concerne l'exercice 6, le nombre de possibilités équiprobables

² Pour que la somme des deux possibilités fasse « 1 ».

³ Pour la définition de cardinal, voir I-2) page 10.

définissant le fait de tirer l'as de carreau est « 1 ». Le nombre de possibilités équiprobables est « 52 » (il y a 52 cartes) et la probabilité de tirer l'as de carreau est $1/52$.

Dans ce qui suit, pour alléger l'écriture, le terme « possibilité » sera remplacé par « événement » (Mais les puristes estimeront que c'est un abus de langage⁴ !).

Exemple de la différence qu'il y a entre une probabilité *avant* une épreuve et la probabilité *après* l'épreuve :

Dans un jeu de 52 cartes, on tire une carte : *on sait que c'est une figure*. Quelle est la probabilité que la carte tirée soit un roi ?

– Avant l'événement, la probabilité était le rapport du nombre de rois au nombre total de cas possibles. C'est-à-dire $4/52=1/13$.

– Après l'événement, la carte tirée est une figure, le nombre de cas favorables est toujours 4, MAIS le nombre total de cas possibles a diminué : c'est le nombre de *figures* (12). La probabilité est $4/12=1/3$. Elle est différente (elle a augmenté).

Définition générale

La probabilité peut se définir de deux manières :

1) $P = \text{nombre de cas favorables} / \text{nombre de cas possibles}$.

Pour utiliser cette formule, il faut faire très attention au fait que les cas soient *également probables*. Cette définition utilisant des cas *équiprobables* comporte, malheureusement, un large degré d'arbitraire.

Une autre précaution indispensable est de formuler le problème d'une façon la plus rigoureuse possible sinon on risque de se trouver avec des résultats erronés. Un exemple célèbre est donné en Appendice II avec « *Le paradoxe des deux enfants* ».

L'exemple des 3 coffrets, donné par Henri Poincaré [Poincaré 1912], illustre l'importance d'une bonne prise en compte des cas équiprobables :

J'ai 3 coffrets (*A*, *B* et *C*) ayant chacun 2 tiroirs *h* (haut) et *b* (bas) identiques. Le contenu (une pièce d'or ou d'argent) de ces tiroirs est le suivant :

Coffret	<i>A</i>	<i>B</i>	<i>C</i>
Tiroir <i>h</i>	Or	Argent	Or
Tiroir <i>b</i>	Or	Argent	Argent

⁴ En effet, il y a une différence subtile entre l'*événement* (résultat de l'épreuve) et la *possibilité* (avant la réalisation de l'épreuve). Pour les « puristes », ces deux termes ne sont pas synonymes. Par esprit de simplification, dans cet ouvrage, comme la probabilité d'un événement est la probabilité d'obtention d'une possibilité, ces deux termes sont, *à tort*, utilisés comme des synonymes. (Puisque la probabilité d'un événement est la possibilité que cet événement se produise.)

1) J'ouvre *au hasard* un des tiroirs. Quelle est la probabilité que je trouve une pièce d'or ? Le fait d'ouvrir les tiroirs *au hasard* entraîne qu'ouvrir un tiroir est un cas équiprobable.

Il y a 6 cas également probables : *Ah, Ab, Bh, Bb, Ch, Cb*. Parmi ces 6 cas, il y a 3 cas favorables : *Ah, Ab, Ch*. Donc la probabilité de trouver une pièce d'or est $3/6=1/2$.

2) À la suite de l'événement précédent, j'ai trouvé de l'or. Quelle est la probabilité que le second tiroir contienne de l'argent ?

Un raisonnement **faux** serait : Ou bien je suis tombé sur le coffret *A* ou bien sur le coffret *C*. La probabilité serait, alors $1/2$. Ce raisonnement est **faux** parce que les deux cas envisagés sont inégalement probables : Le cas *A* est deux fois plus probable que le cas *C* (deux tiroirs contiennent de l'or pour le cas *A* alors que pour le cas *C*, il n'y en a qu'un seul).

Un raisonnement **correct** serait : Je peux trouver de l'or dans trois cas également probables : *Ah, Ab, Ch*. Parmi ces trois cas, un seul est favorable pour que je trouve de l'argent au second tiroir *Ch*. La bonne réponse est, donc, $1/3$.

L'autre définition est :

2) $P = \text{nombre de fois où l'événement arrive} / \text{nombre d'expériences}$.

Pour un nombre d'expériences tendant vers l'infini (c'est-à-dire un nombre très grand⁵).

Propriétés des probabilités

Dans ce qui suit, la probabilité de l'événement *A* est notée $P(A)$.

1. – Une probabilité est un pourcentage. Nécessairement **entre 0 et 1** :

$$0 \leq P(A) \leq 1$$

2. – $P(\Omega)=1$ (Ω , ensemble fondamental, est l'ensemble de tous les événements, cet événement arrive à coup sûr).

3. – $P(\emptyset)=0$ ($\emptyset$, ensemble vide, est un événement qui n'arrive jamais).

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

b. – $P(\bar{A}) = 1 - P(A)$

La relation $P(A \cup B)$ est évidente avec l'analogie [Lazar 1967] de la probabilité comme étant une masse : Pour mesurer la masse de $A \cup B$, il faut additionner la masse de *A* avec celle de *B* puis enlever la masse de la partie $A \cap B$ (commune à *A* et *B*). En d'autres termes, cette formule signifie que la probabilité que se réalise l'un ou l'autre de deux événements quelconques est égale à la somme de leur probabilité *diminuée de la probabilité qu'ils se réalisent ensemble*.

C'est-à-dire : $P(A \text{ ou } B) = P(A) + P(B) - P(A \text{ et } B)$



⁵ En pratique de l'ordre de « 30 ».

Donc

– Si (A et B) n'arrive jamais, $P(A \text{ ou } B) = P(A) + P(B)$ ◀◀◀

– $\bar{A}$ étant le complémentaire de A,

$P(A \cap \bar{A}) = P(\emptyset) = 0$ et $P(A \cup \bar{A}) = P(\Omega) = 1$.

On obtient : $P(\bar{A}) = 1 - P(A)$ ◀◀◀

Probabilité conditionnelle, probabilités composées

La probabilité d'un événement A se produisant si l'événement B est réalisé est notée selon $P(A|B)$. (C'est la probabilité de A « sachant B »). C'est-à-dire, $P(A \text{ si } B)$, la Probabilité de l'événement A sous la condition B.

La règle des **probabilités composées** s'écrit :

$P(A | B) = P(A \cap B) / P(B)$ ◀◀◀

Cette équation est illustrée dans la figure 1 [Brandt 1976].

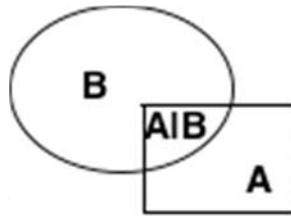


Figure 1 : Illustration de la probabilité conditionnelle $P(A | B)$: La surface de chaque région est proportionnelle à la probabilité de l'événement correspondant. La zone $A | B$ est égale à 1 si B recouvre A et est égale à 0 si les deux zones n'ont pas de recouvrement.

On déduit facilement que : $P(A | B) = P(B | A) \times P(A) / P(B)$.

(C'est le théorème de Bayes développé ci-dessous).

Exercice 7 : démontrez la relation ci-dessus (de Bayes).

Exercice 8 [Escoubès 1998] : Un dé est « pipé » pour donner un 6 avec une probabilité 1/2. On mélange 50 dés normaux avec 50 dés pipés de cette façon. Un dé est tiré au hasard et jeté : on obtient un 6. *Quelle est la probabilité que le dé jeté soit « pipé » ?* (Utiliser la formule ci-dessus).

Indépendance

Deux événements A et B sont dits indépendants si l'événement B n'influe pas l'événement A et *reciproquement*. Dans ce cas :

$P(A \cap B) = P(A) \times P(B)$ ◀◀◀

➡➡➡ Alors la relation des probabilités composées donne $P(A | B) = P(A)$.

(« A si B » est indifférent au fait qu'il y ait B. donc le résultat est seulement lié à A.)

En outre, deux événements A et B sont indépendants si les probabilités conditionnelles satisfont la relation

$$P(B | A) = P(B | \bar{A}) \quad \leftarrow \leftarrow \leftarrow$$

(La probabilité de B est insensible à l'arrivée de A et à l'arrivée de son contraire.)

Il est possible de montrer⁶ que la valeur $P(B|A)$ est égale à $P(B)$. (L'événement B est le même que si il y a A ou pas.)

Enfin, deux variables aléatoires sont dites indépendantes si et seulement si leur distribution conjointe (page 24) est factorisable :

$f(x,y)=g(x)\times h(y)$ (Avec, respectivement, $g(x)$, $h(y)$ les distributions (page 19, de x et de y .)

Théorème de Bayes

Les probabilités de Thomas Bayes, appelées « probabilités des causes », découlent d'un *schéma probabiliste précis* comportant des événements *distincts* (par exemple un tirage de boules d'une urne qui contient des boules blanches et noires)

Le théorème de Bayes⁷ permet de relier les probabilités conditionnelles $P(A | B)$ et $P(B | A)$. Il s'écrit :

$$P(A | B) = P(B | A) \times \frac{P(A)}{P(B)} \quad \leftarrow \leftarrow \leftarrow$$

Cette formule se résume à :

La probabilité de A si B est égale à la probabilité de B si A multipliée par la probabilité de A et divisée par la probabilité de B.

Par exemple, la probabilité qu'il ait plu si mon toit est mouillé est égale à la probabilité que mon toit soit mouillé s'il a plu multipliée par la probabilité qu'il ait plu et divisée par la probabilité que mon toit soit mouillé.

Ou, pour reprendre le cas de l'exercice 8, la probabilité si on a tiré un 6 que le dé soit pipé est égale à la probabilité de tirer un 6 alors que le dé est pipé multipliée par la probabilité de tirer un dé pipé et divisée par la probabilité de tirer un 6.

En mettant en relation deux événements, cette formule mesure la probabilité pour un des phénomènes de provoquer l'autre. On remonte ainsi de la *conséquence* (le phénomène observé) à la *cause* de ce phénomène.

Ce théorème, qui permet de *remonter* aux probabilités avant les épreuves à partir des *résultats* des épreuves, a fait couler beaucoup d'encre. Il a causé la séparation des statisticiens en « bayésiens » et « antibayésiens ».

Pour ceux que cette « controverse » intéresse, l'Appendice II donne un

⁶ Pour le montrer, il faut utiliser la relation d'indépendance « $P(A \cap B) = P(A) \times P(B)$ ».

⁷ Ce théorème dû à Bayes, en 1748, a été repris et *étendu* par Pierre Simon de Laplace en 1774 (qui n'était, peut-être, pas au courant des travaux de Bayes). On doit à ce dernier la forme définitive de ce théorème.

exemple des arguments avancés de part et d'autre y compris un exemple dû à Poincaré [Poincaré 1912] pour montrer une utilisation aberrante du théorème de Bayes.

Extension de Bayes aux causes multiples.

Quoi qu'il en soit, dans le cas de causes multiples, à condition d'avoir des événements *discrets* (c'est-à-dire des événements bien séparés), il est possible d'étendre le théorème de Bayes à *plusieurs* probabilités ***mutuellement exclusives et exhaustives***⁸:

Soit une famille d'événements mutuellement exclusifs $A_1, A_2, \dots, A_n$ et un événement quelconque B, le théorème de Bayes s'écrit :

$$P(A_i | B) = \frac{P(A_i) \times P(B | A_i)}{\sum_{i=1}^n P(B | A_i) \times P(A_i)} \quad \leftarrow \leftarrow \leftarrow$$

Les probabilités $P(A_i)$ et $P(B|A_i)$ doivent satisfaire ces relations :

$$\sum_{i=1}^n P(A_i) = 1$$

$$\sum_{i=1}^n P(B | A_i) = 1$$

Sous cette forme, la formule permet de « décrypter » un phénomène dû à de multiples causes en connaissant la probabilité de *chacune* de ces causes. Elle permet de prendre en compte *tous* les phénomènes possibles (toutes les causes possibles à condition de connaître la probabilité de chacune) afin de trouver *lequel* est la cause du phénomène étudié,

Les applications de ce théorème sont énormes. (Le nombre de publications qui l'utilisent a explosé depuis 2000...) Elles concernent pratiquement toutes les disciplines scientifiques et certaines non-scientifiques.

Par exemple : astronomie, informatique, météorologie, sismologie, physique des particules, chimie, et... science du comportement, diagnostic médical, génétique, production industrielle, etc... L'exercice suivant devrait vous convaincre de sa puissance et ceci dans un domaine non-scientifique :

Exercice 9 [Escoubès 1998] : Un taxi heurte une autre voiture et s'enfuit. Un témoin assure qu'il était vert.

Il fait sombre et ce témoin, est pris, au hasard, dans une population constituée de 20 % de menteurs [les autres disent la vérité]. En plus, dans cette ville, 85 % des taxis sont bleus (les autres taxis sont verts) ; *Quelle est la probabilité pour que le taxi, en réalité, ait été bleu ?*

⁸ C'est-à-dire qu'un événement X_i doit appartenir à A_i et seulement A_i .

II

Variables aléatoires

II – 1) Généralités

Définition

Une variable aléatoire x est un nombre (qui peut prendre des valeurs discrètes ou continues) associé à un événement aléatoire⁹ X . Par exemple, à l'événement « obtenir un 6 avec un dé » est associée la variable aléatoire discrète « $x=6$ ». Mais on peut aussi, à l'événement aléatoire « jet de 2 dés » associer la variable aléatoire « x =somme des 2 chiffres obtenus ».

Note importante : **si x est une variable aléatoire, $f(x)$, fonction de x , l'est aussi.**

Dans ce qui suit, la notation suivante sera utilisée : L'événement aléatoire E est associé à une variable aléatoire (X) qui peut prendre la valeur $X=x$.

Densité de probabilité (ou distribution¹⁰) d'une variable aléatoire

Variable discrète : À chaque valeur possible $x_1, x_2, \dots$ de X , il est possible d'associer une probabilité $P(X=x_1), P(X=x_2), \dots$. Ces probabilités forment la densité de probabilité ou la distribution de probabilités. Par exemple, une urne opaque contient une proportion q de boules noires et une proportion p de boules blanches. À l'événement aléatoire E « tirer une boule de l'urne » est associée la variable aléatoire X qui prend la valeur 0 si la boule tirée est blanche et 1 si elle est noire. Dans ces deux cas, $P(X=1)=q$ et $P(X=0)=p$ (avec $q+p=1$). La densité de probabilité « $f(x)$ » de cet événement aléatoire est montrée dans la figure 2 :

⁹ Pour la définition d'un événement aléatoire voir I-1) page 9.

¹⁰ Dans cet ouvrage, une « distribution » est une loi (densité) de probabilité. **À ne pas confondre avec** le mot anglo-saxon (et de la littérature) « *distribution* » (qui signifie ce qui est défini, ici, comme « fonction de répartition ».) Dans la mesure du possible, ce mot sera utilisé parce qu'il est plus court et qu'il n'y a pas de risque de confusion (puisque la « distribution » anglo-saxonne est appelée, dans cet ouvrage, systématiquement « fonction de répartition ».)

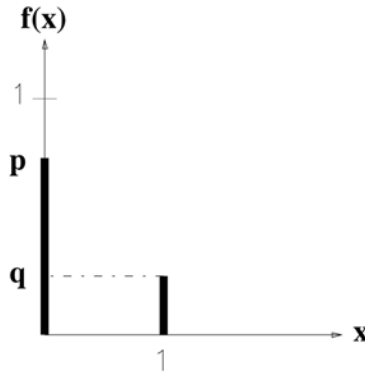


Figure 2 : Densité de probabilité de la variable aléatoire X associée à l'événement aléatoire E « tirer une boule d'une urne opaque » avec des boules ($q\%$ noires et $p\%$ blanches, avec la relation $q+p=1$) $X=0$ Si la boule tirée est blanche, $X=1$ Si elle est noire.

Autre exemple : Dans le cas d'un dé normal (c'est-à-dire « non-pipé »), chaque face a une probabilité $1/6$ de « sortir » après un lancer. La densité de probabilité « $f(x)$ » de la variable aléatoire (correspondant à ce jet) « x = valeur de la face qui sort » associée au lancer est montrée dans la figure 3. Cette densité de probabilité est appelée « diagramme en bâtons ».

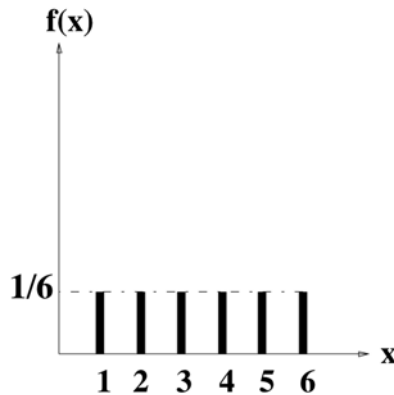


Figure 3 : Densité de probabilité de la variable aléatoire X associée à l'événement aléatoire E « lancer de dé normal ».

Variable continue : La distribution de X est la probabilité de l'événement $x \leq X \leq x+dx$ quand dx tend vers zéro. Elle est notée $f(x)dx$. Alors la probabilité que X prenne une valeur dans l'intervalle $a \leq X \leq b$ est $P(a \leq X < b) = \int_a^b f(x)dx$.

Par exemple [Brandt 1976], une montre est lue au hasard et l'on note la