Probabilités pour les sciences de l'ingénieur

Manuel Samuelides

Probabilités pour les sciences de l'ingénieur

Cours et exercices corrigés

DUNOD

Illustration de couverture : © Federo-istock.com

ISBN 978-2-10-059615-7

Table des matières

Introduction

Cet ouvrage est destiné à permettre aux ingénieurs en formation d'acquérir les compétences nécessaires en probabilités. Il s'agit de la compréhension des méthodes et de la maîtrise des outils de calcul, indispensables à la conduite de processus, au contrôle des mesures et à l'évaluation des risques.

Actuellement, le métier d'ingénieur se développe vite en fonction des besoins et du développement économique et technique de nos sociétés et il se diversifie. Il y a dans le monde et en France plusieurs niveaux de diplômes d'« ingénierie ». Le niveau visé dans cet ouvrage est celui de « l'ingénieur généraliste », capable d'analyser une situation, de proposer des innovations et d'effectuer des transferts de technologie.

La formation à ce métier suppose un « socle commun scientifique » dont la composition peut évoluer mais dont le contenu mathématique est relativement stabilisé. En revanche, les étapes d'acquisition de ce socle sont très variées et ont tendance à se multiplier : classes préparatoires aux grandes écoles scientifiques (CPGE), formations d'ingénieur en 5 ans, cycles préparatoires universitaires, formations de techniciens supérieurs... L'originalité de cet ouvrage est d'avoir été conçu pour s'adapter à tous ces itinéraires. Cette spécificité explique son plan et son articulation en plusieurs parties qui sera détaillée dans la suite de cette introduction. Chacune des parties correspond *grosso modo* à une année de formation de L2 à M1 ou de L3 à M2 selon le « niveau mathématique » de cette formation. Jusqu'à maintenant, les élèves de CPGE n'avaient aucune notion de probabilité dans leur programme L1-L2 de mathématiques. Cette situation va changer à partir de 2014.

Les traits spécifiques des probabilités que nous allons exposer maintenant justifient à notre avis qu'un ouvrage spécifique soit consacré à ce domaine des mathématiques appliquées aux sciences de l'ingénieur qui soit distinct d'un simple ouvrage de « mathématiques de l'ingénieur » ou de « théorie et calcul des probabilités » qui conviendrait à des formations plus applicatives ou au contraire plus théoriques et en tout cas plus spécialisées.

Les probabilités dans la formation de l'ingénieur

Les probabilités, une science relevant des mathématiques appliquées

La théorie des probabilités est une branche relativement récente des mathématiques. Des éléments d'histoire en sont donnés tout au long de cet ouvrage. Les diverses branches des mathématiques se sont développées depuis l'Antiquité en combinant :

- un objet de modélisation (mouvement des corps célestes dans l'antiquité grecque, mécanique du solide indéformable aux XVII^e-XVIII^e siècles, mécanique des vibrations, diffusion de la chaleur aux XVIII^e puis XIX^e siècles...) ;
- une axiomatique définissant le cadre de développement abstrait et ses règles (géométrie de la droite, du cercle et des coniques, calcul différentiel et intégral, transformée de Fourier...) ;
- des calculs développés rigoureusement dans le cadre de cette axiomatique et permettant de prédire et d'expliquer les observations effectuées sur l'objet modélisé.

L'intérêt des mathématiques appliquées dans la formation des ingénieurs réside à notre avis dans l'articulation de ces trois thèmes dans un domaine ou un autre, la modélisation étant l'axe central permettant le transfert d'un domaine d'application à un autre et donc l'innovation. Cet axe est plus important que celui du calcul, plus ou moins mécanisable, ou que celui de la rigueur déductive qui relève plus des chercheurs scientifiques.

Or, les probabilités présentent à cet égard plusieurs difficultés intéressantes :

- Les objets de modélisation des probabilités, dès leur émergence relativement tardive et tout au long de leur histoire, ont été extrêmement variés : assurances maritimes, jeux de hasard, arithmétique politique et sociale, mesures physiques, biologie et écologie, changement d'échelle... Les sciences de l'ingénieur occupent dans cette panoplie une place relativement réduite.
- Le cadre théorique de leur développement a beaucoup évolué de l'analyse combinatoire (XVII^e siècle) au calcul différentiel (XVIII^e siècle) et son unification axiomatique date de moins d'un siècle (1933). Ce cadre qui est celui de la **théorie de la mesure** est particulièrement abstrait relativement aux cadres géométriques, numériques ou fonctionnels des autres domaines des mathématiques appliquées (algèbre, analyse...).
- Les mesures permettant de valider les prédictions probabilistes nécessitent le développement de protocoles particuliers. Leur conception relève d'une science distincte appelée la **statistique**. La statistique n'est pas l'objectif d'étude de cet ouvrage. Son évocation est bien sûr nécessaire, étant donné ses liens étroits avec les probabilités (voir chapitres 6 et 7).

Les probabilités, un outil de modélisation et de calcul du risque

Les probabilités n'ont pas toujours été considérées comme faisant partie du socle commun scientifique de tout ingénieur. Les équations régissant la mécanique du solide et les engrenages, ainsi que la thermique étaient considérées comme suffisamment précises pour permettre de concevoir des dispositifs industriels sûrs et fiables.

Le développement du génie électrique, son utilisation dans la transmission du signal et les télécommunications nécessitent au contraire la prise en compte de perturbations dont l'estimation statistique permet le contrôle. Plus généralement, les probabilités interviennent directement dans les domaines physiques de l'ingénieur partout où il est nécessaire d'analyser les lois du changement d'échelle (passage du microscopique au macroscopique). Ce fut le cas en transmission électromagnétique mais aussi en mécanique des fluides où l'écoulement turbulent fait l'objet de modélisations probabilistes fondamentales et dans un nombre croissant de domaines au fur et à mesure des progrès de la modélisation physique.

À l'époque où l'ingénieur obtenait les résultats numériques dont il avait besoin par la consultation de tables de recensement de ces résultats (logarithmes, fonctions spéciales), l'utilisation des modèles probabilistes fut facilitée par la prééminence dans les domaines de la physique des **lois gaussiennes ou « normales »**. Celles-ci interviennent naturellement dans de nombreux domaines de la physique en raison du théorème central-limite (chapitre 6) et leur intervention dans les dispositifs industriels est pérenne dans la mesure où ceux-ci peuvent être considérés comme linéarisables autour de leur point de fonctionnement. Enfin, le calcul des lois gaussiennes se ramène aux calculs de moyennes et de variance qui relèvent de l'algèbre linéaire, très présente dans le socle commun mathématique de l'ingénieur. Le calcul analytique des lois de probabilité classiques associées à la loi gaussienne (lois gamma, loi du khi 2, loi beta...) complète ce dispositif et permet l'application des méthodes statistiques qui se développent au XIX^{e} siècle et dans la première moitié du XX^{e} siècle.

Le développement de la grande industrie, le déploiement en parallèle et en série de chaînes de machines, le transport d'un site de production à un autre rendent nécessaire le calcul de risques pour concevoir des systèmes de production fiables, évitant les défaillances individuellement imprévisibles par des mesures préventives dont le coût est calculable. Ce nouveau besoin émergent au XX^{e} siècle fait appel à des méthodes probabilistes plus élaborées utilisant un développement nouveau, l'étude des **processus stochastiques**, c'est-à-dire des phénomènes aléatoires dépendant du temps.

L'impact de la révolution numérique

Beaucoup plus récemment, le développement de la « révolution numérique » a eu deux conséquences sur l'utilisation des probabilités par l'ingénieur :

Méthodes de Monte-Carlo. Le développement du calcul numérique a permis de considérer la totalité des lois de probabilité avec une approximation plus ou moins précise. Surtout, la génération de nombres « pseudo-aléatoires » a permis la construction d'échantillons de données de grande taille et le calcul direct approché des grandeurs probabilistes à partir de ces échantillons. Paradoxalement, ces méthodes ont été introduites avec succès dans les calculs déterministes susceptibles d’interprétations probabilistes (calcul d’intégrale).

Collecte d’informations de taille considérable (big data). Plus récemment, les progrès de la collecte et de la transmission d’informations liés à Internet ont mis à disposition une quantité d’information considérable, pouvant être utilisée dans des buts très variés, commerciaux aussi bien qu’industriels, dont le volume nécessite un traitement statistique adapté à l’utilisation projetée. Cette spécialité faisant appel à l’informatique et aux statistiques a reçu le nom de **fouille de données**, (« *data mining* »). Si sa mise en œuvre relève d’ingénieurs spécialisés et à ce titre, sera seulement évoquée dans la conclusion de cet ouvrage, l’ingénieur généraliste se doit de comprendre le principe de mise en œuvre de ces techniques afin de les intégrer dans un processus de conception et de maintenance. En effet, l’importance de ces techniques entraîne un bouleversement général des processus de conception qui incluent maintenant la prise en considération de l’objet produit pendant toute sa durée d’utilisation.

Plan de l’ouvrage

Comme nous l’avons indiqué, l’ouvrage se décompose en trois parties suivies d’une conclusion qui correspondent à des stades différents dans l’acquisition des compétences de l’ingénieur en formation.

Acquérir une intuition probabiliste et comprendre la nature du modèle

Il s’agit de prolonger les acquis de l’enseignement secondaire et de maîtriser les différentes composantes du modèle probabiliste : **espace d’état, variables aléatoires, lois de probabilité, espérance et variance** en utilisant des outils mathématiques élémentaires (théorie des séries et calcul différentiel et intégral à une ou deux variables). Cette partie pourra être traitée, selon les formations, entre les années L2 et L3.

Le chapitre 1 a pour objet les modèles discrets, il prolonge les acquis de l’enseignement secondaire aux espaces d’état dénombrables en utilisant les séries et passe en revue les différentes lois de probabilité sur les entiers.

Le chapitre 2 a pour objet les modèles à densité dans des espaces vectoriels à une ou deux dimensions. Il peut donc se nourrir de l’intuition géométrique élémentaire et s’appuyer sur de nombreuses applications physiques (droite des moindres carrés). Il comporte une introduction dans les limites de la dimension 2 au calcul de lois.

L’indépendance et le conditionnement sont abordés dans ces chapitres sous un aspect élémentaire.

S'initier aux outils mathématiques et maîtriser les techniques de calcul

Cette deuxième partie constitue le corps de l'ouvrage avec les chapitres 3, 4, 5, 6 et 7. Le chapitre 3 expose le modèle général de l'espace de probabilité. Les bases mathématiques sont énoncées dans un appendice pour rendre ce chapitre-clé accessible. Le lecteur passe ainsi aux espaces de dimension quelconque et surtout devient capable de traiter dans le même cadre conceptuel les lois discrètes et les lois à densité, ce qui est à notre avis l'apport indispensable de ce modèle mathématique aux probabilités élémentaires.

Le chapitre 4 est un chapitre de méthodes de calcul qui expose les bases du calcul de lois, donne les exemples fondamentaux de familles de lois stables par somme de variables indépendantes et cite une première application statistique classique.

Le chapitre 5 a pour ambition de traiter le délicat problème du conditionnement de façon complète (probabilité conditonnelle et espérance conditionnelle) en s'appuyant sur des applications incontournables : fusion bayésienne d'information, filtre de Kalman.

Le chapitre 6 expose les théorèmes limites (lois des grands nombres et théorème central-limite) qui sont exploités dans une introduction à l'estimation statistique qui fait l'objet du chapitre 7.

Acquérir les bases de deux domaines appliqués de la dynamique aléatoire : files d'attente et signal aléatoire

On a vu que les principaux domaines d'application des probabilités aux sciences de l'ingénieur (génie industriel et traitement du signal) exploitaient des systèmes dynamiques aléatoires (processus stochastiques). Le modèle mathématique général de ces systèmes dépasse largement le cadre de cet ouvrage. On a choisi d'exposer le modèle des chaînes de Markov (temps discret) dans le chapitre 8 en insistant sur le cas des espaces d'état discrets. Plusieurs domaines d'application sont abordés (finance, gestion, dynamique des populations, physique statistique). Une place particulière est faite aux techniques de simulation. Les deux chapitres suivants détaillent les deux principales applications que nous avons sélectionnées.

Le chapitre 9 s'appuie sur l'exposé des propriétés du processus de Poisson pour détailler les modèles de files d'attente et les éléments de base de la théorie de la fiabilité. On choisit ici de mettre l'accent sur des méthodes générales capables de traiter numériquement des cas d'application et de n'exposer que le modèle théorique le plus simple.

Le chapitre 10 expose le modèle de processus gaussien permettant d'expliciter les représentations temporelles et fréquentielles du filtrage. On introduit pour cela l'outil de l'intégrale de Wiener, permettant l'écriture de l'action d'un filtre linéaire avec le même formalisme dans le cas d'un signal déterministe et d'un signal aléatoire. On obtient ainsi un cadre mathématique dans lequel l'écriture des formules traditionnelles d'un filtre a un sens et on peut faire le lien avec la représentation fréquentielle plus facile à concevoir.

Comment aller plus loin

Au-delà de l'ensemble des sujets traités dans cet ouvrage, bien d'autres sujets sont actuellement en plein développement et passent de l'état d'objets de recherche scientifique à celui de méthodes génératrices d'innovation. Ainsi, on observe le transfert technologique de méthodes inventées pour résoudre des problèmes en petite dimension désormais applicables à des problèmes de dimension beaucoup plus grande. Des modèles biologiques peuvent être utilisés pour développer l'autonomie de systèmes "intelligents". Ces sujets font l'objet d'ouvrages spécifiques destinés à des ingénieurs plus spécialisés (M2, doctorat). Cependant, il nous a semblé utile d'en évoquer quelques-uns dans notre conclusion.

Comment utiliser cet ouvrage

Cet ouvrage est conçu pour accompagner des méthodes pédagogiques plus vivantes. Il s'agit bien sûr des cours et des travaux dirigés. Mais les sources de connaissance actuelles dépassent bien largement ces méthodes traditionnelles. Un grand nombre de cours et d'exercices corrigés sont désormais disponibles sur Internet et font l'objet d'une part importante de nos références bibliographiques. Dans un proche avenir, ces documents qui font déjà une large place aux cours en ligne seront complétés par des accompagnateurs numériques plus interactifs qui combineront tests de connaissance et exposés de notions nouvelles.

Notre ouvrage a été conçu comme un texte de référence couvrant un domaine assez large, des fondements mathématiques aux applications, et donc pouvant contribuer à la constitution d'un socle général de compétences de l'ingénieur en probabilités de l'ingénieur en formation tout en ouvrant des perspectives aux développements récents et à l'innovation.

En mettant au premier plan les capacités de modélisation du modèle probabiliste, cet ouvrage est aussi conçu pour servir de support à une pédagogie de projet combinant modélisation, résolution, simulation et mise en œuvre applicative. Cette pédagogie déjà utilisée dans les formations d'ingénieurs est appelée à se développer laissant aux ingénieurs en formation, l'investissement personnel en acquisition des connaissances de base. À ce titre, des exercices corrigés et des méthodes de programmation explicitées en SCILAB (logiciel français disponible gratuitement) complètent les exposés didactiques de cet ouvrages dans les deux premières parties.

Remerciements à quelques formateurs

Il reste à remercier tous ceux qui par leur soutien ont contribué à cet ouvrage... mais la liste en serait bien longue. Je dois en tout cas remercier Gilbert Saporta dont les conseils m'ont aidé à sélectionner les sujets de la troisième partie de l'ouvrage. Je voudrais aussi citer Bernard Bru avec qui nous avions élaboré il y a longtemps une pédagogie d'enseignement en premier cycle intégrant l'histoire des probabilités.

SUPAERO (Institut Supérieur de l'Aéronautique et de l'Espace) a été un lieu privilégié pour l'enseignement des mathématiques appliquées. La proximité de l'Université, de l'ONERA, du CNES et d'Airbus a favorisé la collaboration multidisciplinaire qui imprègne, je l'espère, cet ouvrage. Les équipes pédagogiques que j'ai animées ont toujours été constituées en partie d'ingénieurs en activité, notamment dans le domaine spatial. Je remercie tous mes collègues et nommerai à titre d'exemple, ceux qui m'accompagnent ainsi depuis longtemps comme Yves Caumel, Jean-Louis Dunau, Henri Sénateur, Patrice Henry et Jean-François Gajewski.

Les plus nombreux et les plus efficaces de mes formateurs en pédagogie ont été bien sûr les nombreux élèves ingénieurs, auditeurs de mes cours. Combien d'élèves-ingénieurs de SUPAERO ont combiné le goût des modèles mathématiques et l'enthousiasme d'apprendre à créer du nouveau. Je voudrais aussi citer les apprentis en formation d'ingénieurs en alternance au CNAM. Ils m'ont incité à chercher de nouvelles approches plus abordables développant des introductions par l'exemple et ont directement inspiré la forme finale des deux premiers chapitres de cet ouvrage.

1 Modèles aléatoires discrets

PLAN

OBJECTIFS

- Comprendre les différentes notions de base du calcul des probabilités : espace d'états, événement aléatoire, variable aléatoire, espérance, variance dans les modèles discrets simples et explicites.
- Pouvoir construire le modèle discret d'une situation incertaine dans le cas d'un espace d'états fini ou dénombrable.
- Être capable d'identifier les modèles discrets où l'espérance et la variance d'une loi de probabilité discrète peuvent être calculées analytiquement.
- Être capable de simuler numériquement un échantillon indépendant de taille donnée d'un modèle discret et d'en déduire une approximation numérique de l'espérance et de la variance d'une variable aléatoire réelle définie sur ce modèle.
- Comprendre la notion de probabilité conditionnelle et pouvoir l'appliquer dans la construction d'un modèle aléatoire sur un espace d'état structuré en arbre.
- Comprendre la notion d'indépendance dans le cas des modèles discrets et connaître les propriétés des variables aléatoires indépendantes.

ENCART 1.1 Éléments d'histoire : les origines du calcul des probabilités

La théorie des probabilités occupe une place à part dans l'histoire des mathématiques. Elle se développe peu de temps après le calcul différentiel et intégral (à partir du XVII^{e} siècle) et souvent les mêmes auteurs contribuent aux deux secteurs des mathématiques en plein essor et dans le but commun de rendre compte de la réalité observable.

Mais si l'analyse mathématique se réfère à cette époque à la géométrie rigoureuse de l'Antiquité, le « calcul » des probabilités n'a pas cette prétention. Il a des origines plus ludiques et essaie de rendre compte des jeux de hasard qui passionnaient les salons de cette époque. Plus précisément, le « problème des paris » du chevalier de Méré exposé dans ce chapitre fait l'objet d'une correspondance entre Pascal et Fermat en 1654 considérée comme à l'origine de l'approche scientifique des probabilités. Huygens (1657) et Bernoulli (1704) reviennent sur ce problème en généralisant et en systématisant le raisonnement de Pascal [1].

L'importance historique des jeux de hasard dans le développement du calcul des probabilités s'explique par la simplicité relative des outils mathématiques utilisés dans ces applications où l'espace des états est un espace fini et où on a tendance à considérer que tous les états sont équiprobables. Les mêmes méthodes (sommations algébriques) peuvent aussi s'appliquer quand les états qu'il est possible d'observer ne sont plus équiprobables. Comment dans ce cas connaître les probabilités des états ou événements élémentaires ? L'observation des fréquences d'occurrence peut-elle se substituer au simple dénombrement d'occurrences équiprobables ?

Si la nécessité pratique du calcul des probabilités apparaît ainsi progressivement (voir encadrés historiques des chapitres suivants), son caractère scientifique s'impose plus difficilement. Ainsi au Moyen Âge, en Europe chrétienne, la vérité scientifique apparaît comme d'origine divine et difficilement accessible et le mot « probable » désigne une opinion reconnue comme fiable et vraisemblable par la majorité des humains. Quand, dans les Temps modernes, le relativisme théologique est battu en brèche et que les hommes s'approprient la construction des « vérités scientifiques », la place du calcul des probabilités reste problématique et au dix-neuvième siècle, Auguste Comte, le maître du « positivisme scientifique » peut encore qualifier le « prétendu calcul des chances » de « honteuse aberration scientifique directement incompatible avec toute vraie positivité » [22].

Les justifications mathématiques des règles statistiques sont en effet complexes et demandent au moins une bonne maîtrise du calcul intégral. Les variables aléatoires continues seront abordées dans cet ouvrage au chapitre 2. Les bases d'une théorie plus générale des probabilités n'ont été clarifiées qu'au vingtième siècle et seront évoquées au chapitre 3 pour pouvoir étudier les phénomènes aléatoires complexes dépendant du temps qui constituent actuellement l'essentiel des applications du calcul des probabilités aux sciences de l'ingénieur.

1.1 Symétrie du hasard et probabilité uniforme

1.1.1 Notion d'espace d'état

La notion d'**état d'un système** est essentielle en modélisation des systèmes physiques. L'état d'un système est un objet mathématique contenant toutes les informations « utiles » sur ce système. Par exemple, l'état d'un point en dynamique est défini comme le couple position-vitesse. La connaissance de l'état d'un système dans une théorie physique donnée permet de connaître complètement le comportement de ce système.

Généralement, l'état d'un système est déterminé par la connaissance de plusieurs variables appelées **variables d'état**. Ces variables peuvent être des nombres réels. Si elles sont en nombre fini d, l'ensemble des états du système que nous noterons Ω ou parfois $\mathscr{E}$ est un sous-ensemble de $\mathbb{R}^d$. Dans la pratique, l'état d'un système peut aussi comporter des variables qualitatives décrivant par exemple l'état de panne ou de bon fonctionnement des composants du système. L'ensemble des états d'un système est appelé **espace d'état** du système.

Dans le cadre d'une théorie physique déterministe, l'état d'un système à un instant donné est un élément de l'espace d'état du système. Dans une modélisation aléatoire, l'état du système n'est pas défini ; le comportement du système à un instant donné est décrit par une loi de probabilité, c'est-à-dire par des nombres décrivant la « probabilité » qu'a le système de se trouver dans tel ou tel état. Le cas le plus simple à modéliser est celui où tous les états du système sont également probables, c'est le cas de la *symétrie du hasard*.

Prenons des exemples de lancer de dés qui, on l'a vu, ont joué un rôle fondateur dans l'élaboration du calcul des probabilités :

> ***Exemple du lancer d'un dé***
>
> Considérons l'expérience aléatoire consistant à lancer un dé à six faces. Le résultat de cette expérience est un entier appartenant à l'ensemble à 6 éléments $\{1,2,3,4,5,6\}$. Tout événement provenant de cette expérience est représenté par un sous-ensemble de cet ensemble, l'ensemble des résultats correspondant à la réalisation de l'événement. Par exemple, l'événement « le résultat est supérieur à 3 » est représenté par l'ensemble $\{4,5,6\}$. L'événement « le résultat est inférieur à 3 » est représenté par l'ensemble $\{1,2\}$. L'espace d'état attaché au résultat de cette expérience aléatoire est effectivement l'ensemble $\{1,2,3,4,5,6\}$.

> ***Exemple d'un jeu de lancer de dés à deux joueurs***
>
> Considérons maintenant un jeu à deux joueurs Alain et Bernard. Chaque joueur mise une somme unité et lance un dé à six faces. Le joueur ayant obtenu le résultat le plus élevé remporte les mises. Si les deux joueurs obtiennent des résul-

tats de même niveau, chaque joueur dispose à nouveau de sa mise. Ici l'espace d'état Ω est l'ensemble à 36 éléments des couples de nombres à valeurs dans $\{1,2,3,4,5,6\}$. Un état (x,y) représente l'issue où le joueur A a obtenu le résultat x et le joueur B le résultat y. L'événement « Alain gagne » est représenté par le sous-ensemble des 15 états suivants :

$$A = \{(2,1),(3,1),(3,2),(4,1),(4,2),(4,3),(5,1),(5,2),(5,3),(5,4),(6,1),(6,2),(6,3),(6,4),(6,5)\}.$$

L'événement « Bernard gagne » est représenté par le sous-ensemble B des 15 états symétriques. L'événement « chaque joueur reprend sa mise » est représenté par l'ensemble à 6 états :

$$E = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$$

On peut considérer qu'il existe deux variables d'état, le résultat obtenu par Alain d'une part et le résultat obtenu par Bernard d'autre part. Nous reviendrons plus loin sur la notion de variable aléatoire. La représentation de l'expérience aléatoire par un ensemble d'états permet d'effectuer des opérations logiques sur les événements attachés au résultat de l'expérience aléatoire. On va considérer les opérations logiques élémentaires « OU », « ET », « NON ». Dans notre exemple, l'événement « un des deux joueurs gagne » est le résultat de l'opération logique :

« Alain gagne » OU « Bernard gagne »

Cet événement est représenté par l'ensemble d'états $A \cup B$.

De même l'événement « Chaque joueur reprend sa mise » est la négation de l'événement précédent :

« Chaque joueur reprend sa mise » = NON(« Alain gagne » OU « Bernard gagne »)

L'identité logique précédente est représentée dans l'espace d'états par l'égalité ensembliste :

$$E = (A \cup B)^c$$

On peut résumer ces correspondances dans l'énoncé suivant :

Exemple

Considérons une expérience aléatoire dont le résultat est modélisé par un élément ω d'un ensemble d'états noté Ω. Tout événement observable à l'issue de cette expérience aléatoire est modélisé par un sous-ensemble de Ω, l'ensemble des états dans lesquels on observe effectivement l'occurrence de l'événement. Il existe une correspondance entre les événements aléatoires et les opérations logiques d'une part, et les sous-ensembles d'états et les opérations ensemblistes d'autre part, résumée dans le tableau ci-après :

Tableau de la correspondance logique-ensembliste

Logique	Notation ensembliste	Commentaire
état observable	ω	
événement toujours observé	Ω	ensemble des états observables
événement jamais observé	$\emptyset$	ensemble vide
événement aléatoire « A »	$A \subset \Omega$	
opération « A OU B »	$A \cup B$	
opération « A ET B »	$A \cap B$	
opérateur « NON A »	complémentaire A^c	

Remarquons que si nous avons montré comment associer à chaque événement aléatoire un sous-ensemble d'états, il n'est pas évident de faire correspondre à tout sous-ensemble de Ω un événement aléatoire.

Dans le cas envisagé ici où Ω est fini, cette correspondance réciproque peut s'imposer. En effet, tout état $\omega \in \Omega$ représente un événement aléatoire observable. Si on considère un sous-ensemble $A \subset \Omega$, il représente l'événement aléatoire formé par l'assemblage par « OU » de tous les événement élémentaires associés aux état $\omega \in A$. Dans le cas où un ensemble d'états fini Ω est associé à l'expérience aléatoire, on a donc une correspondance bijective entre les événements aléatoires observables à l'issue de l'expérience et l'algèbre de Boole $\mathscr{P}(\Omega)$ de tous les sous-ensembles de l'espace d'états Ω.

Rappelons la définition suivante d'une algèbre de Boole de parties d'un ensemble Ω :

Définition 1.1 Algèbre de Boole de parties d'un ensemble

Soit Ω un ensemble et $\mathscr{A} \subset \mathscr{P}(\Omega)$ un ensemble de parties de Ω. On dit que $\mathscr{A}$ est une **algèbre de Boole** si elle vérifie les propriétés suivantes

a) $\Omega \in \mathscr{A}$

b) $A \in \mathscr{A} \Rightarrow A^c \in \mathscr{A}$

c) $A \in \mathscr{A}$ & $B \in \mathscr{A} \Rightarrow A \cup B \in \mathscr{A}$

En particulier $\mathscr{P}(\Omega)$ est une algèbre de Boole.

1.1.2 Probabilité uniforme

Dans l'analyse des « jeux de hasard », le calcul de la probabilité d'un événement aléatoire se ramène au problème de compter le nombre d'états constituant l'ensemble d'états représentant cet événement et de le diviser par le nombre d'états total.

Ainsi, dans l'exemple précédent du jeu à deux joueurs, on a

$$\text{Probabilité(« Alain gagne »)} = \frac{15}{36}$$

ou

$$\text{Probabilité(« Chaque joueur reprend sa mise »)} = \frac{6}{36}$$

En effet, l'équilibre du jeu impose que tous les états ont la même probabilité. C'est le principe de **« symétrie du hasard »**. On obtient ainsi la définition de la **probabilité uniforme** qui historiquement fut la première loi de probabilité considérée et à laquelle on cherchait toujours à se ramener avant qu'elle n'apparaisse que comme un cas très particulier.

Définition 1.2 Loi de probabilité uniforme sur un ensemble fini Ω

Soit Ω un espace d'états fini associé à une expérience aléatoire. On appelle **loi de probabilité uniforme sur** Ω l'application P de $\mathscr{P}(\Omega)$ dans $[0,1]$ définie par

$$A \in \Omega \to P(A) = \frac{\sharp(A)}{\sharp(\Omega)}$$

où $\sharp(A)$ désigne le nombre d'éléments de A.

Propriétés de la loi de probabilité uniforme sur un ensemble fini

On démontre sans difficulté à partir de la définition précédente les propriétés suivantes :

Théorème 1.1

La loi de probabilité uniforme sur un ensemble fini Ω est une application de $\mathscr{P}(\Omega)$ dans $[0,1]$ qui vérifie les propriétés suivantes :

(a) *masse unité :* $P(\Omega) = 1$

(b) *additivité :* $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

1.2 Loi de probabilité sur un ensemble fini

Exemple des dés distinguables et indistinguables

Reprenons l'exemple de l'expérience aléatoire du lancer de deux dés. Nous avons vu que l'espace d'état de cette expérience aléatoire était $\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$ et comptait 36 états. Supposons maintenant que les dés au lieu d'être de deux couleurs différentes soient de la même couleur et soient lancés en même temps. Le modèle probabiliste n'est pas changé pour autant car rien dans la physique du problème n'a changé. Pourtant les états ne sont plus observables. Les événements aléatoires observables sont des réunions de $15 + 6 = 21$ événements aléatoires élémentaires, les 15 couples d'états du type $\{(1,2),(2,1)\},\{(1,3),(3,1)\},\ldots,\{(5,6),(6,5)\}$ et les 6 états $(1,1),(2,2),(3,3),(4,4),(5,5)$ et $(6,6)$. Ces événements aléatoires sont toujours composables par les opérations logiques « OU » et « ET » et l'opérateur logique « NON ». Ils forment donc une sous-algèbre de Boole $\mathscr{A}$ de $\mathscr{P}(\Omega)$ et la restriction de la probabilité uniforme à cette sous-algèbre vérifie les propriétés de masse-unité et d'additivité.

Cependant, l'observateur incapable de distinguer les deux dés peut construire un autre modèle : il peut choisir l'espace d'état empirique E à 21 états que nous noterons

$$((1,2)),((1,3)),\ldots,((5,6)),((1,1)),((2,2)),..,((6,6))$$

Cet observateur va remarquer empiriquement que ces états n'ont pas la même occurrence et qu'un état du type $((x,y))$ avec $x \neq y$ est plus souvent observable qu'un état du type $((x,x))$. **Le principe de symétrie du hasard n'est plus observé**.

Deux stratégies sont possibles et ont successivement été mises en œuvre dans l'histoire des probabilités comme on va le voir dans l'exemple du problème posé à Pascal par le chevalier de Méré :

– reconstruire un espace d'état Ω plus complexe que l'espace d'état empirique E et sur lequel le principe de symétrie du hasard s'applique et la probabilité uniforme permet une prédiction des observations,

– munir E d'une probabilité non uniforme pour pouvoir prédire efficacement les observations. Dans l'exemple considéré, on poserait : si $x \neq y$, $P(\{((x,y))\}) = \frac{1}{18}$ et $P(\{((x,x))\}) = \frac{1}{36}$.

L'application de la première stratégie était possible dans l'application aux jeux de hasard et a permis l'invention d'algorithmes de dénombrement efficaces et ingénieux. Au bout de quelques décennies, la seconde stratégie s'est imposée et a permis de généraliser le concept de probabilité et donc d'étendre considérablement son champ d'applications.

Définition 1.3 Loi de probabilité sur un ensemble fini Ω

Soit Ω un ensemble fini, on appelle **(loi de) probabilité sur** Ω une application P de $\mathscr{P}(\Omega)$ dans $[0,1]$ vérifiant les propriétés :

(a) masse unité : $P(\Omega) = 1$

(b) additivité : $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

Généralisation de l'additivité

Considérons n ensembles $A_1,\ldots,A_n$ disjoints deux à deux : $\forall i, j, A_i \cap A_j = \emptyset$.

On montre alors par récurrence sur n que $P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots P(A_n)$.

En effet $A_{n+1} \cap (A_1 \cup \ldots \cup A_n) = (A_{n+1} \cap A_1) \cup \ldots \cup (A_{n+1} \cap A_n) = \emptyset$.

Théorème 1.2

Soit Ω un ensemble fini et P une probabilité sur $\mathscr{P}(\Omega)$. On a

- $\sum_{\omega \in \Omega} P(\{\omega\}) = 1$,
- $\forall A \subset \Omega, P(A) = \sum_{\omega \in A} P(\{\omega\})$

Réciproquement soit $\omega \in \Omega \to p_\omega \in [0,1]$ une application de Ω dans $[0,1]$ telle que $\sum_{\omega \in \Omega} p_\omega = 1$, alors il existe une et une seule probabilité P sur $\mathscr{P}(\Omega)$ telle que $\forall \omega \in \Omega, P(\{\omega\}) = p_\omega$.

La démonstration est une conséquence directe de la définition d'une probabilité sur un ensemble fini.

ENCART 1.2 Éléments d'histoire : équilibre de l'urne d'Ehrenfest

À la fin du XIXe siècle dans leur effort d'unification des sciences, des physiciens théoriciens essaient de montrer que la thermodynamique et notamment l'évolution des fluides peut se justifier dans le cadre de la mécanique classique : il faut pour cela que les quantités accessibles à la mesure sont des quantités macroscopiques résultant d'une dynamique gouvernée par les lois de la mécanique classique à l'échelle microscopique. Cette nouvelle approche est appelée **mécanique statistique**. Elle est due notamment à Maxwell (1831-1879), principal auteur de la théorie cinétique des gaz, et à Boltzmann (1844-1906), auteur du H-théorème qui donne une définition probabiliste de l'entropie (voir [1] et articles du site Wikipédia consacrés à ces auteurs). Nous aurons l'occasion dans la partie de cet ouvrage consacrée aux dynamiques probabilistes et aux processus stochastiques de revenir sur ces questions.

Cette vision probabiliste de la physique était fortement contestée à la fin du XIXe siècle. Pour en montrer la validité, les époux Ehrenfest ont produit en 1907 un modèle jouet de dynamique aléatoire dont nous allons ici simplement citer la probabilité d'équilibre.

On considère N particules (par exemple des molécules de gaz) se répartissant « au hasard » entre deux urnes séparées par une paroi (éventuellement perméable en dynamique) et appelée « urne 0 » et « urne 1 ». Le modèle de symétrie du hasard s'applique et on veut étudier la probabilité de la répartition des molécules entre les deux urnes qui déterminera par exemple les forces de pression sur la paroi séparant les deux urnes.

On choisit comme espace d'états $\Omega = \{0,1\}^N$, ensemble fini à 2^N éléments. La i-ème variable d'état est égale à 0 si la i-ème molécule est dans l'urne 0 et à 1 si elle est dans l'urne 1. Tous les états sont équiprobables et donc la probabilité d'un événement aléatoire A est $P(A) = \frac{J(A)}{2^N}$.

Cependant, pour l'observateur les particules sont indistinguables et les seuls états qu'il peut identifier sont composés par des événements élémentaires du type $a_k = (N-k,k)$, événement aléatoire défini par la présence de k particules dans l'urne 1 et $N-k$ particules dans l'urne 0 pour $0 \leq k \leq N$. Pour calculer la probabilité de a_k, il nous faut dénombrer le nombre de partitions d'une population de N éléments en deux sous-populations de $N - k$ et de k éléments. On sait que ce nombre noté $\binom{n}{k}$ et appelé le **nombre de combinaisons de k éléments parmi** N est donné par la formule du binôme de Newton

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}$$

On a donc $$P(a_k) = \frac{N!}{k!(N-k)!}\frac{1}{2^N}$$

Le principe de symétrie du hasard ne s'applique pas ici non plus à la probabilité restreinte à l'espace des états indistinguables $\mathrm{E} = \{a_k\}_{0\leq k\leq N}$. Les événements élémentaires a_k de la représentation d'état indistinguable ont des probabilités fortement différentes et d'autant plus inégales que N est grand. Nous étudierons sous le nom de « loi binomiale » cette probabilité plus en détail dans la suite de ce chapitre.

Exercice 1.1 Paradoxe du Grand Duc de Toscane

Galilée est surtout connu pour ses travaux en astronomie. Cependant, il étudia aussi les probabilités en précurseur à la demande de son protecteur, le duc de Toscane. Celui-ci lui soumet le problème suivant provenant d'un jeu où on lance trois dés. Les joueurs avaient remarqué que le total de 10 est plus souvent obtenu que le total de 9. Cette observation leur semblait contradictoire avec le fait que 10 comme 9 peut s'obtenir à partir de 3 dés de 6 façons différentes :

$$10 = 2+2+6 = 1+3+6 = 1+4+5 = 2+3+5 = 2+4+4 = 3+3+4$$
$$9 = 1+2+6 = 1+3+5 = 1+4+4 = 2+2+5 = 2+3+4 = 3+3+3$$

L'explication du paradoxe réside dans le fait que les joueurs raisonnaient en probabilité uniforme sur un espace d'état où les dés sont indistinguables et que dans le cas des dés, le principe de symétrie du hasard s'applique aux raisonnements sur les espaces d'état où les dés sont distinguables. Calculer les probabilités d'obtenir 9 et 10.

Corrigé page 34

Exercice 1.2 Formule de Poincaré

Cette formule classique semble avoir été découverte au moins dans un cadre plus restrictif par le mathématicien du XVII[e] siècle De Moivre, plus connu pour les formules de manipulation des nombres complexes. Soit P une probabilité sur une algèbre de Boole $\mathscr{A}$ de parties d'un ensemble.

1) Montrer que $\forall (A,B) \in \mathscr{A} \times \mathscr{A}$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2) Etendre cette formule par récurrence au cas de n ensembles $A_1, \ldots, A_n$.

Corrigé page 35

1.3 Probabilité sur un espace d'état dénombrable

Exemple de la poule à trois joueurs

On considère 3 joueurs A, B et C qui se livrent à une série de duels. Dans un premier temps, on suppose que le nombre de duels est limité à 4 et que les duels sont équilibrés (lancer d'une pièce à pile ou face avec probabilité $\frac{1}{2}$). Les joueurs A et B se livrent le premier duel et le vainqueur affronte le joueur C. Si C perd, alors le vainqueur successif de ses deux adversaires est déclaré le vainqueur de la poule. Au bout de quatre lancers le jeu s'arrête. Quelles sont les probabilités de gain des joueurs A, B et C ? Quelle est la probabilité que le jeu s'arrête sans conclusion ?

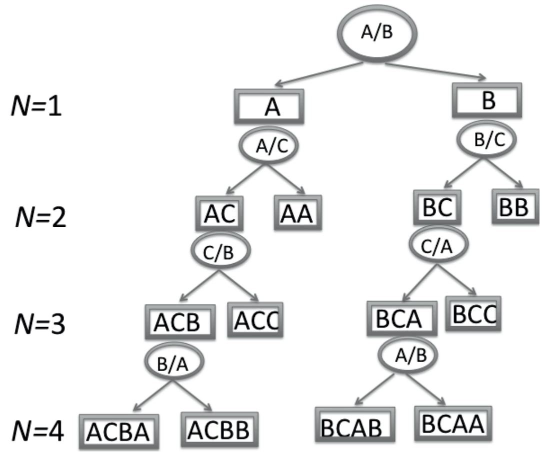

Figure 1.1– **Graphique des déroulements possibles de la poule à 3 joueurs pour $N = 4$**

Si nous construisons un modèle en utilisant une probabilité uniforme, comme $N = 4$, nous utiliserons la probabilité uniforme sur un espace d'états $\Omega = \{0,1\}^4$ à 16 états.

L'arbre de la figure précédente montre que le jeu a en fait dix issues. Dans les deux cas AA et BB, le jeu s'arrête au bout de deux lancers de dés. Dans chacun de ces cas, les lancers de dés suivants sont fictifs et non observables. On a donc $P(AA) = P(BB) = \frac{1}{4}$.

Pour les deux feuilles suivantes de l'arbre ACC et BCC, le jeu s'arrête au bout de trois lancers de dés et on a $P(ACC) = P(BCC) = \frac{1}{8}$.

Enfin, on a $P(ACBB) = P(ACBA) = P(BCAA) = P(BCAB) = \frac{1}{16}$. On est donc conduit à définir un espace d'états à 10 états non équiprobables $\mathrm{E}_4 = \{AA, ACC, ACBB, ACBA, BB, BCC, BCAA, BCAB\}$.

Supposons maintenant que nous voulions modéliser un jeu se déroulant selon les mêmes règles mais où le nombre de lancers n'est pas précisé et où le jeu doit se dérouler jusqu'à ce qu'un vainqueur apparaisse à la suite de deux victoires consécutives. L'espace d'états est maintenant infini et comporte les états :

$$\mathrm{E}_\infty = \{AA, ACC, ACBB, ACBAA, \ldots, ACBACBA\ldots, BB, BCC, BCAA, BCABB, \ldots, BCABCAB\ldots\}$$

On ne peut exclure en effet *a priori* que le jeu ne s'arrête pas. Essayons maintenant de munir $\mathscr{E}_\infty$ d'une probabilité correspondant aux spécifications de l'énoncé. Les états caractérisés par n duels ont la probabilité 2^{-n}. Reste à attribuer une probabilité aux états infinis. Or la probabilité d'un de ces états, par exemple $\{ACBACBA\ldots\}$, est nécessairement inférieure aux probabilités des événements ACB, ACBACB, ACBACBACB... Cette probabilité est donc 0.

Exercice 1.3 *Calcul des probabilités dans le problème de la poule à trois joueurs*

1) Calculer la probabilité pour que le jeu s'achève en N coups en fonction des valeurs de N.

2) Calculer les probabilités de gain de chacun des joueurs A, B et C.

Corrigé page 35

Nous pouvons tirer quelques conclusions de l'étude de cet exemple :

- Une probabilité sur l'algèbre de Boole $\mathscr{P}(\Omega)$ des parties d'un ensemble dénombrable Ω doit pouvoir se définir à partir des probabilités de chaque état en remplaçant les sommes finies par des sommes infinies (limite de la suite des sommes partielles).
- Il est nécessaire de disposer d'une propriété de passage à la limite qui n'est pas utile dans les algèbres de Boole finies.

Définition 1.4 Tribu

Soit Ω un ensemble, on appelle $\sigma-$algèbre de Boole ou **tribu** un ensemble $\mathscr{A}$ de parties de Ω qui soit une algèbre de Boole et qui de plus possède une des propriétés équivalentes suivantes pour une algèbre de Boole :

(a) Soit (A_n) une suite croissante de $\mathscr{A}$ c'est-à-dire telle que $\forall n, A_n \subset A_{n+1}$ alors $\cup_n A_n \in \mathscr{A}$.

(b) Soit (A_n) une suite quelconque de $\mathscr{A}$, alors $\cup_n A_n \in \mathscr{A}$.

Démonstration Pour montrer que (a)⇒(b). Si (A_n) est une suite quelconque de $\mathscr{A}$, on pose $B_1 = A_1, B_2 = A_1 \cup A_2, \ldots, B_n = A_1 \cup \ldots \cup A_n$ et alors B_n est une suite croissante et (a) peut s'appliquer.

Définition 1.5 Probabilité sur une tribu

Soit Ω un ensemble et $\mathscr{A}$ une tribu de parties de cet ensemble. Une probabilité sur $\mathscr{A}$ est une application de $\mathscr{A}$ dans $[0,1]$ possédant les propriétés suivantes :

(a) masse unité : $P(\Omega) = 1$

(b) additivité : $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(c) continuité monotone croissante : Soit A_n une suite croissante de $\mathscr{A}$ (telle que $\forall n, A_n \subset A_{n+1}$), alors la suite croissante $P(A_n)$ a pour limite $P(\cup_n A_n)$

Remarque

Les propriétés (b) et (c) sont équivalentes à la propriété de $\sigma-$additivité :
Soit A_n une suite de $\mathscr{A}$ d'ensembles disjoints deux à deux : $i \neq j \Rightarrow A_i \cap A_j = \emptyset$ alors on a

$$P(\cup_n A_n) = \sum_n P(A_n)$$

Il est clair que la σ-additivité entraîne comme cas particuliers l'additivité (suite finie) et la continuité monotone croissante (suite croissante). Pour prouver la réciproque, on reprendra la construction déjà utilisée

$$B_1 = A_1, B_2 = A_1 \cup A_2, \ldots, B_n = A_1 \cup \ldots \cup A_n$$

En passant aux complémentaires, on montre que la continuité monotone croissante est équivalente à la continuité monotone décroissante qui s'énonce ainsi :

(d) continuité monotone croissante : Soit A_n une suite décroissante de $\mathscr{A}$ (telle que $\forall n, A_{n+1} \subset A_n$), alors la suite décroissante $P(A_n)$ a pour limite $P(\cap_n A_n)$.

Définition 1.6 Espace de probabilité

On appelle **espace de probabilité** la donnée d'un espace d'état Ω, d'une tribu $\mathscr{A}$ de parties de Ω représentant les événements aléatoires considérés dans le modèle et d'une probabilité P sur $\mathscr{A}$.

Les espaces de probabilité généraux seront étudiés par la suite (chapitre 3). Leur étude fait appel à la théorie mathématique de la mesure et de l'intégration. Dans ce chapitre, nous manipulons les espaces de probabilité dits « discrets » construits sur des espaces d'état finis ou dénombrables. Leur manipulation utilise seulement la théorie des ensembles et l'algèbre élémentaire pour les espaces de probabilité finis et la théorie des séries quand l'espace d'état est dénombrable.

Le théorème 1.2 qui montre comment les probabilités sur un ensemble d'états fini se construisent à partir de la donnée de la probabilité sur les états se généralise sans difficulté aux espaces d'état dénombrables en utilisant les propriétés des séries numériques positives On a ainsi :

Théorème 1.3

Soit Ω un espace d'états dénombrable et P une probabilité sur $\mathscr{P}(\Omega)$. On a

- $\sum_{\omega\in\Omega} P(\{\omega\}) = 1$
- $\forall A \subset \Omega, P(A) = \sum_{\omega\in A} P(\{\omega\})$

Réciproquement soit $\omega \in \Omega \to p_\omega \in [0,1]$ une application de Ω dans $[0,1]$ telle que $\sum_{\omega\in\Omega} p_\omega = 1$, alors il existe une et une seule probabilité P sur la tribu $\mathscr{P}(\Omega)$ telle que $\forall \omega \in \Omega, P(\{\omega\}) = p_\omega$.

1.4 Probabilité conditionnelle et indépendance

Notion de probabilité conditionnelle

Considérons un espace de probabilité discret $(\Omega, \mathscr{P}(\Omega), P)$ et supposons que nous bénéficions *a posteriori* d'une information partielle : on sait que l'événement aléatoire B s'est réalisé. Cette information nous amène à changer notre modèle. Ainsi, dans le nouveau modèle, l'événement B devient le nouvel espace d'état de probabilité 1, les états non éléments de B n'ont plus besoin d'être considérés. Si P était la probabilité uniforme sur l'espace d'état Ω : $P(A) = \frac{\sharp(A)}{\sharp(\Omega)}$, la nouvelle probabilité est simplement la probabilité uniforme sur B : les états à l'issue desquels B est réalisé sont équiprobables, les autres états ne sont plus considérés comme réalisables. La nouvelle probabilité d'un événement aléatoire A notée $P(A \mid B)$ et appelée probabilité conditionnelle de A sachant B, en abrégé probabilité de A si B est donnée par :

$$P(A) = \frac{\sharp(A \cap B)}{\sharp(B)}$$

Plus généralement, cette définition s'étend au cas des probabilités non uniformes sur les espaces de probabilité

Définition 1.7 Probabilité conditionnelle

Soit $(\Omega, \mathscr{A}, P)$ un espace de probabilité et $B \in \mathscr{A}$ un événement aléatoire de probabilité non nulle. On appelle **probabilité conditionnelle sachant** B la nouvelle probabilité sur $\mathscr{A}$ définie par

$$A \in \mathscr{A} \rightarrow P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

On vérifie sans peine que la probabilité conditionnelle ainsi définie vérifie les propriétés permettant la définition d'une probabilité (masse 1 et σ-additivité).

Exercice 1.4 *(suite de l'exercice 1.1) Paradoxe du Grand Duc de Toscane*

Dans le lancer de trois dés de l'exercice 1.1 (paradoxe du Grand Duc de Toscane page 9), calculer les probabilités conditionnelles de l'obtention de 9 et de 10 sachant que deux dés au moins ont eu des résultats semblables.

Corrigé page 35

La notion de probabilité conditionnelle s'est développée dès l'origine du calcul des probabilités pour permettre au joueur de faire des paris sur l'occurrence d'événements inconnus (par exemple la main de l'adversaire dans un jeu de cartes) à partir d'observations accessibles (par exemple ses annonces ou ses premiers coups). Elle est aussi très utile pour pouvoir interpréter le résultats de tests à issue incertaine, que ce soit en médecine pour des diagnostics ou en génie industriel pour détecter des dysfonctionnements. On doit au pasteur britannique Thomas Bayes (1702-1761) la formulation claire et générale de cette proposition dans son ouvrage posthume *Essay Towards Solving a Problem in the Doctrine of Chances* (1763).

Théorème 1.4 Théorème de Bayes

Soit $(\Omega, \mathscr{A}, P)$ un espace de probabilité et (B_n) une partition finie ou dénombrable de Ω en événements aléatoires i.e.

(a) $\forall n, B_n \in \mathscr{A}$

(b) $n \neq p \Rightarrow B_n \cap B_p = \emptyset$

(c) $\cup_n B_n = \Omega$

Alors on a

$$\begin{cases} \forall A \in \mathscr{A}, \quad P(A) = \sum_n P(A \mid B_n)P(B_n) \\ P(B_i \mid A) = \dfrac{P(A \mid B_i)P(B_i)}{\sum_n P(A \mid B_n)P(B_n)} \end{cases}$$

Démonstration On a par σ-additivité de P la formule de recomposition de $P(A)$:

$$P(A) = \sum_n P(A \cap B_n) = \sum_n P(A \mid B_n)P(B_n)$$

On obtient la formule d'inversion de Bayes à partir de la formule de recomposition en appliquant simplement la définition de la probabilité conditionnelle :

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A \mid B_i)P(B_i)}{\sum_n P(A \mid B_n)P(B_n)}$$

Ce théorème important permet ainsi de formuler des hypothèses statistiques sur une propriété inconnue d'un système B_i à partir d'une observation incertaine (mesure, test) A et d'une analyse *a priori* du système (connaissance des probabilités de B_n et de A sachant B_n).

L'indépendance de deux événements aléatoires est une notion fondamentale. Sa définition est une conséquence directe de la définition d'une probabilité conditionnelle. Deux événements aléatoires A et B sont dits indépendants si la probabilité de A n'est pas changée par la connaissance de l'occurrence de B, autrement dit si l'information apportée par l'observation de B ne change pas notre prévision de l'occurrence de A.

Définition 1.8 Indépendance de deux événements aléatoires

On dit que deux événements aléatoires A et B d'un espace de probabilité $\Omega, \mathscr{A}, P$ sont **indépendants** si $P(A \cap B) = P(A)P(B)$.

Cette définition montre la symétrie de la relation d'indépendance. On a

Théorème 1.5

A et B sont indépendants si et seulement si A et B^c sont indépendants.

Démonstration Supposons A et B indépendants. On a

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)P(B^c)$$