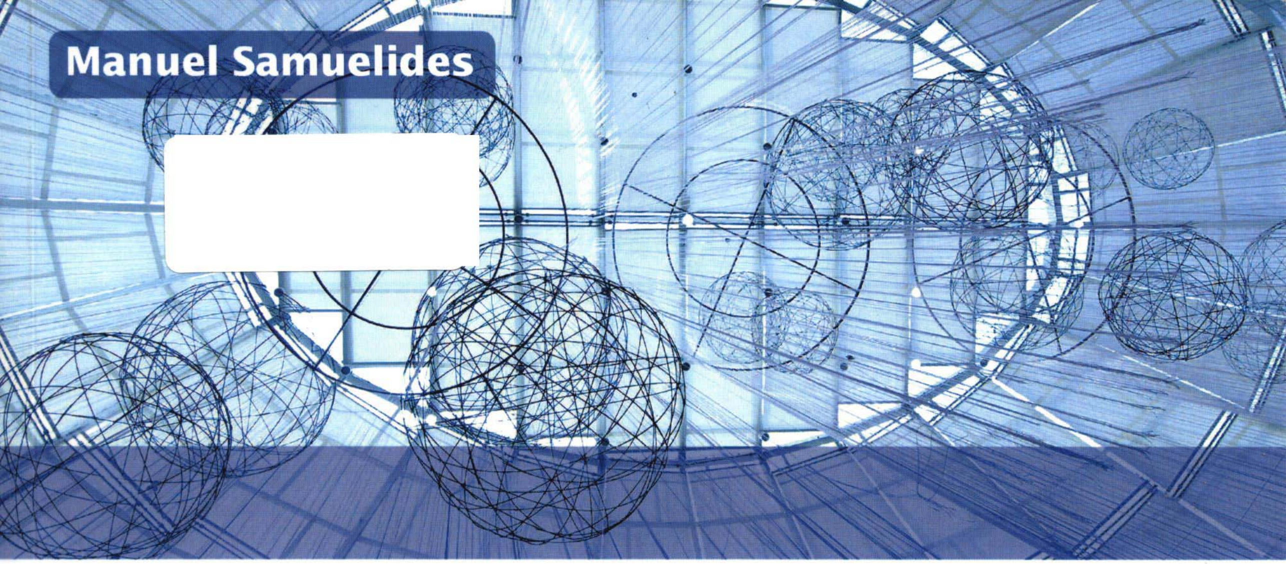
Manuel Samuelides



Probabilités pour les sciences de l'ingénieur

Cours et exercices corrigés

- Écoles d'ingénieurs
- Licence 2 et 3
- Master 1

RESSOURCES NUMÉRIQUES

DUNOD

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

Probabilités pour les sciences de l'ingénieur

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.

TABLE DES MATIÈRES

MODÈLES ALÉATOIRES DISCRETS

1

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.