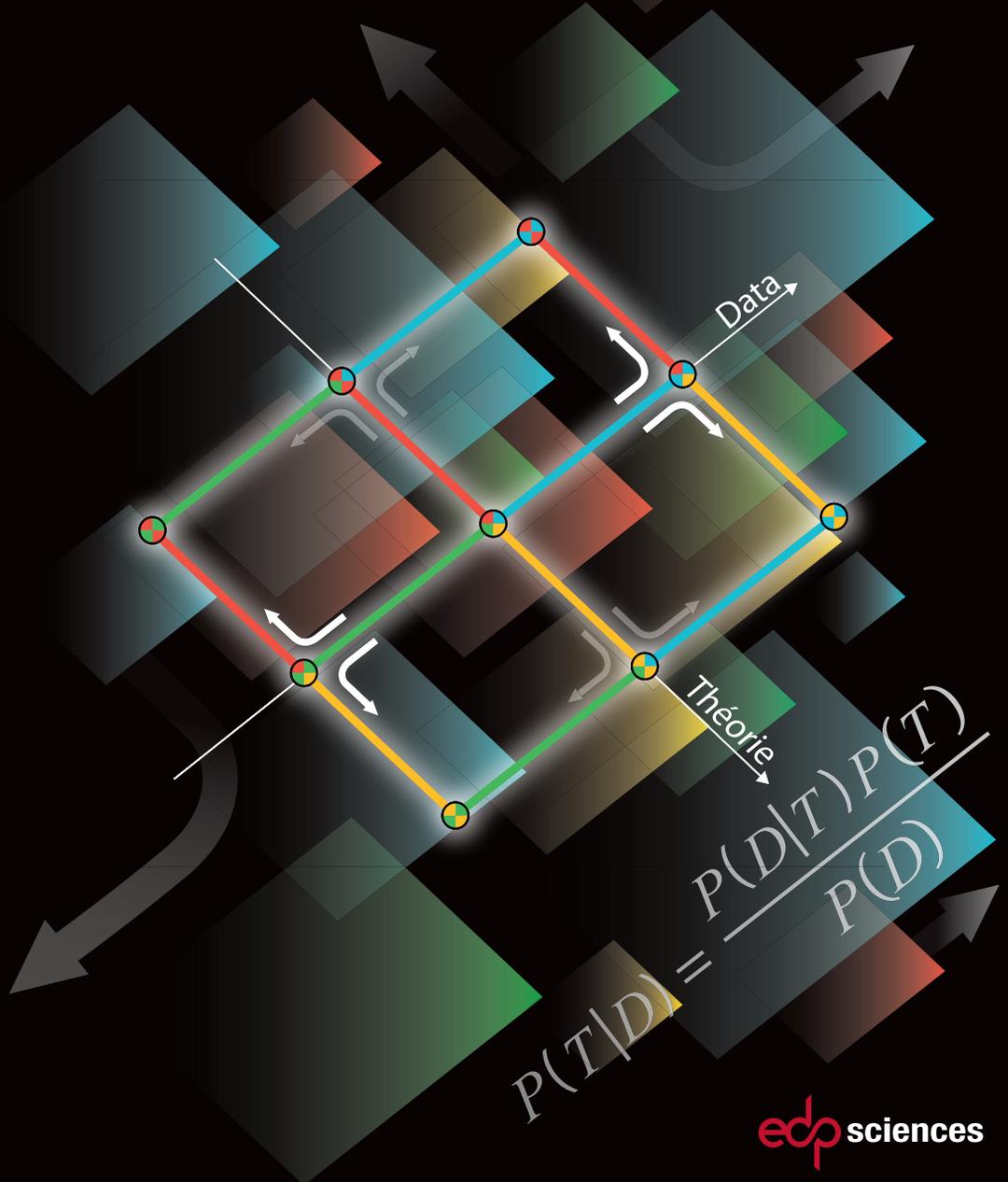


Lê Nguyễn Hoàng

LA FORMULE DU SAVOIR

Une philosophie unifiée du savoir
fondée sur le théorème de Bayes



LA FORMULE DU SAVOIR

Une philosophie unifiée du savoir fondée sur le théorème de Bayes

Lê Nguyễn Hoàng

Ce livre explore et vulgarise une philosophie du savoir appelée *bayésianisme*. En s'appuyant sur les travaux de nombreux philosophes, mathématiciens, statisticiens, informaticiens, neuroscientifiques et chercheurs en intelligence artificielle, le livre défend la thèse selon laquelle le *bayésianisme* est la *bonne* philosophie du savoir — par opposition notamment aux descriptions usuelles de la *méthode scientifique*. En effet, notamment une fois combinée à l'algorithmique, cette épistémologie normative peut se vanter d'être *universelle* et *complète*. De plus, elle est consolidée par un très grand nombre de théorèmes mathématiques et de succès empiriques. S'il contient des passages techniques, la grande majorité de l'ouvrage se veut accessible à un large public. En particulier, aucune connaissance préalable n'est requise.

$$P(T|D) = \frac{P(D|T)P(T)}{P(D)}$$

ISBN : 978-2-7598-2260-7



Lê Nguyễn Hoang

LA FORMULE DU SAVOIR

Une philosophie unifiée du savoir
fondée sur le théorème de Bayes

Lê Nguyễn Hoang

Collaborateur scientifique à l'EPFL, Suisse.
Vidéaste sur la chaîne YouTube Science4All.

Imprimé en France

ISBN (papier) : 978-2-7598-2260-7 – ISBN (ebook) : 978-2-7598-2261-4

Tous droits de traduction, d'adaptation et de reproduction par tous procédés, réservés pour tous pays. La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (alinéa 1^{er} de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles 425 et suivants du code pénal.

© EDP Sciences, 2018

Préface

Arrivant dans une petite ville, avec une lourde valise, vous vous dirigez vers la station de taxi de la gare, où une seule voiture est stationnée. Hélas, le temps que vous vous en approchiez, un voyageur plus rapide l'a déjà empruntée et elle disparaît sous vos yeux. Quelle conclusion pouvez-vous tirer de cette mésaventure ? Qu'il semble y avoir des taxis dans cette ville – vu sa taille, c'était loin d'être assuré – et donc, que si vous attendez patiemment, un autre finira par se présenter ? Ou alors, que l'un des rares taxis de la ville vient de vous échapper et que, vu sa taille, une telle aubaine ne se représentera de sitôt ? Ces deux interprétations sont correctes, mais l'une et l'autre dépendent de ce que vous saviez – ou croyiez – avant de descendre du train.

Ce voyageur qui arrive dans une ville inconnue, fait des hypothèses sur le nombre de taxis et les révisé en fonction de ses observations, n'est pas très différent d'un bébé qui arrive dans un monde inconnu, ou du chercheur qui, s'étonnant de ce que les autres tiennent pour acquis, se demande pourquoi le soleil se lève tous les matins. Les uns comme les autres explorent le monde, font des hypothèses et les révisent en fonction de leurs observations.

Quel enseignement pouvons-nous tirer de nos expériences ? Que pouvons-nous connaître du monde ? Ce sont ces questions que le magnifique livre de Lê Nguyễn Hoang nous invite à nous poser.

Sur ces questions, un point cristallise les controverses depuis plus d'un siècle : est-il possible d'associer, à une hypothèse, une valeur numérique qui mesure sa vraisemblance ? Pour certains, tel Hans Reichenbach, c'est là le but même de la théorie des probabilités. En particulier, toute observation qui confirme une hypothèse, augmente sa probabilité d'être vraie : chaque observation d'un corbeau noir augmente la probabilité que l'hypothèse selon laquelle tous les corbeaux sont noirs soit vraie. Pour d'autres, tel Karl Popper, l'attribution d'une valeur numérique à une telle hypothèse n'est qu'illusion. En observant un corbeau noir, nous ne pouvons que conclure, que notre hypothèse selon laquelle tous les corbeaux sont noirs reste cohérente avec nos observations.

Au cœur de cette controverse, se place une formule d'une simplicité déconcertante, la formule de Bayes, « la formule du savoir », qui donne son titre à ce livre, et qui permet justement de calculer la probabilité que nous devons attribuer à une hypothèse après avoir fait une observation – et donne donc ainsi raison à Reichenbach – mais uniquement à condition que nous ayons su attribuer une

probabilité à cette hypothèse, avant d'effectuer cette observation – et donne donc ainsi raison à Popper.

Si cette question semblait tranchée – en faveur de Popper – au XX^e siècle, l'évolution des techniques de collecte des données, la renouvelle aujourd'hui. Quand nous croyions, au XX^e siècle, qu'il existait des corbeaux blancs, nous pouvons interpréter le fait que les trois corbeaux observés fussent noirs, comme une coïncidence. Quand nous observons, aujourd'hui, mille, un million, un milliard... de corbeaux, et qu'ils sont tous noirs, il faut avoir un certain courage – voire une certaine obstination – pour prétendre que non, tous les corbeaux ne sont pas noirs, et la concordance de nos observations n'est que coïncidence. Au moins sommes-nous contraints de concéder qu'il doit y avoir, parmi les corbeaux, une grande proportion de corbeaux noirs, et même, sans doute, que les corbeaux blancs relèvent de l'exception. Cette objection à la thèse de Reichenbach que constituait le problème des hypothèses a priori mis en évidence par la formule de Bayes, se trouve aujourd'hui relativisée par le déluge des données. D'autres problèmes, en revanche, apparaissent : comment ces données ont-elles été collectées ? La méthode de collecte n'introduit-elle pas un biais, voire une discrimination, à l'égard des corbeaux blancs ? Une fois de plus, nous constatons à quel point l'évolution des techniques, notamment des techniques d'instrumentation scientifique, change la manière dont se posent les questions en philosophie des sciences.

C'est cela qui rend le livre de Lê Nguyễn Hoang passionnant. Il aura été écrit à l'époque d'un basculement. À une époque où l'évolution des techniques changeait le regard que nous portons sur la formule de Bayes et sur sa place dans l'édifice de la connaissance.

Il aura aussi été écrit à une époque où les techniques de communication changeaient notre manière de parler des sciences. Formé à la dure école des vidéos en ligne, Lê Nguyễn Hoang, a su trouver un ton nouveau pour parler des sciences, un ton à la fois rigoureux et narratif, où les exemples illuminent les questions les plus abstraites.

Gilles Dowek,
Chercheur à l'Inria,
Professeur à l'École normale supérieure de Paris-Saclay.

Table des matières

1	Mon voyage initiatique	13
	Collé par un étudiant	13
	Sur les traces du bayésianisme	14
	Une philosophie unifiée du savoir	16
	Une alternative à la méthode scientifique	18
	Le mythe de l'objectivité	20
	Les objectifs du livre	23
2	Le théorème de Bayes	27
	L'énigme des enfants	27
	Le problème de Monty Hall	28
	Le procès de Sally Clark	30
	Le bayésianisme jugé illégal	31
	Le théorème de Bayes	32
	Les composants de la formule de Bayes	34
	Bayes au secours du diagnostic	35
	Bayes au secours de Sally Clark	37
	L'énigme des enfants enfin résolue !	38
	Quelques mots d'encouragement	39
3	Logiquement. . .	41
	Deux modes de raisonnement	41
	Les règles de la logique	43
	Les dames sont-elles toutes bleues ?	45
	Quantificateurs et prédicats	46
	Le syllogisme d'Aristote réinterprété	47
	L'axiomatisation	48
	Platoniciens <i>versus</i> intuitionnistes	48
	La logique bayésienne*	50
	Au-delà du vrai ou faux	51
	Vers une cohabitation de théories incompatibles	53
4	Il faut (bien) généraliser !	57
	Le mouton noir d'Écosse	57
	Une brève histoire de l'épistémologie	58
	Une brève histoire de la planétologie	59

Les sciences contre Popper ?	60
Le fréquentisme*	61
Les statisticiens contre la <i>p-value</i>	63
Le <i>p-hacking</i>	64
Ce qu'en dit un cours de statistique	66
La formule du savoir	67
L'apprentissage cumulatif	68
Revenons-en à Einstein	69
5 Gloire aux préjugés	73
Le problème de Linda	73
Les préjugés au secours de Linda*	74
Les préjugés sont indispensables	76
Le soleil d'xkcd	77
Les préjugés au secours d'xkcd	77
Les préjugés au secours de Sally Clark	78
Les préjugés pour lutter contre les pseudo-sciences	79
Les préjugés au secours des sciences	81
Le bayésien a un préjugé sur <i>tout</i>	83
Les préjugés erronés	86
Les préjugés et la morale	88
6 Les prophètes du bayésianisme	91
Une histoire mouvementée	91
Les origines de la théorie des probabilités	92
Le mystérieux Thomas Bayes	93
Laplace, le père du bayésianisme	94
La loi de succession de Laplace	96
Le grand hiver du bayésianisme	99
Bayes au secours des alliés	100
Des îlots bayésiens dans un océan fréquentiste	102
Bayes secouru par les praticiens	103
Le triomphe de Bayes, enfin !	105
Bayes est partout	106
7 Le démon de Solomonoff	109
Ni homme ni machine	109
Les fondements de l'algorithmique	110
Qu'est-ce qu'un <i>pattern</i> ?	112
La complexité de Solomonoff	113
Le mariage de l'algorithmique et des probabilités	115
Le préjugé de Solomonoff*	118
Bayes au secours du démon de Solomonoff*	119
La complétude de Solomonoff	120
L'incalculabilité de Solomonoff	121
L'incomplétude de Solomonoff	122

En quête de pragmatisme	123
8 Garder le secret	127
Classé confidentiel	127
La cryptographie d'aujourd'hui	128
Bayes à l'assaut des codes cryptés	129
Le sondage randomisé	131
La confidentialité du sondage randomisé	133
La définition de la confidentialité différentielle*	134
Le mécanisme laplacien	136
Robustesse à la composition	137
L'additivité des pertes de confidentialité	138
En pratique, ça ne va pas !	139
Le chiffrement homomorphe	140
9 Les jeux sont faits	143
La magouilleuse	143
Split or Steal	145
La persuasion bayésienne	146
Les points de Schelling	148
L'équilibre mixte	149
Les jeux bayésiens	152
La conception de mécanismes bayésiens*	153
L'enchère de Myerson	154
Les conséquences sociétales du bayésianisme	155
10 Darwin et Bayes font affaire	159
Le biais du survivant	159
Les lézards colorés de Californie	160
La dynamique de Lotka-Volterra*	161
Les algorithmes génétiques	163
Se faire son propre avis ?	163
Un scientifique n'est pas crédible	165
L'argument d'autorité	167
Le consensus scientifique	169
Le putaclic	170
La puissance prédictive des marchés	171
Les bulles financières	174
11 Exponentiellement contre-intuitif	177
Les nombres archi-méga-super géants	177
Le plafond de verre des calculs	178
L'explosion exponentielle	180
La magie des chiffres indo-arabes	183
La loi de Benford	184
L'échelle logarithmique	185

Le logarithme	187
Bayes raffle un prix Gödel	188
Bayes part en vacances	190
La singularité	191
12 Tranchons avec le rasoir d'Ockham	195
Jeudi dernier.	195
Dans le football, rien n'est écrit d'avance	197
Le fléau de la sur-interprétation	198
La complexe quête de simplicité	201
Tout n'est pas simple	202
La validation croisée	204
La régularisation de Tibschirani	206
L'optimisation robuste	207
Bayes au secours de l' <i>overfitting</i> *	208
Seules les inférences bayésiennes sont admissibles*	209
Le rasoir d'Ockham déduit du bayésianisme !	210
13 Les faits sont trompeurs	213
Hôpital ou clinique ?	213
Corrélation n'est pas causalité	215
Cherchez les facteurs de confusion	217
La régression à la moyenne	219
Le paradoxe de Stein	220
L'échec de la stratification endogène	221
Randomisons !	223
Le retour du mouton noir d'Écosse	225
Qu'est-ce qu'un chat ?	226
Le naturalisme poétique	228
14 Vite et (assez) bien	231
Le mystère des nombres premiers	231
Le théorème des nombres premiers	233
Les approximations de τ	234
Les développements limités	235
Les contraintes du pragmatisme	236
Les <i>learning machines</i> de Turing	236
Le bayésianisme pragmatique	239
Les algorithmes sous-linéaires	241
Plusieurs modes de réflexion	243
Devenez post-rigoureux !	244
Les approximations de Bayes	245
15 La faute à pas de chance	249
FiveThirtyEight et l'élection présidentielle de 2016	249
La mécanique quantique est-elle probabiliste ?	250

La théorie du chaos	253
Les automates déterministes imprévisibles	254
La thermodynamique	255
L'entropie de Shannon	256
La compression optimale de Shannon	258
La redondance de Shannon	259
La divergence de Kullback-Leibler	260
La métrique de Wasserstein	262
Les <i>Generative Adversarial Networks</i> (GANs)	263
16 Trou de mémoire	267
La valeur des données	267
Le déluge de données	268
Le problème des toilettes	269
Traiter rapidement un déluge de données	270
Le filtre de Kalman	272
Nos cerveaux confrontés au <i>Big Data</i>	273
Effacer les souvenirs traumatiques	274
Les faux souvenirs	276
Bayes au secours de la mémoire	278
Des mémoires à plus ou moins long terme	279
Les réseaux de neurones récurrents	280
Que faut-il apprendre et enseigner ?	282
17 La nuit porte conseil	285
D'où viennent les idées ?	285
L'art créatif des intelligences artificielles	286
L'allocation de Dirichlet latente (LDA)	287
Le restaurant chinois au secours de LDA	289
Les simulations de Monte-Carlo	290
La descente de gradient stochastique (SGD)	292
Les nombres pseudo-aléatoires	293
L'échantillonnage préférentiel	293
L'échantillonnage préférentiel au secours de LDA	294
Le modèle d'Ising*	296
La machine de Boltzmann	297
MCMC et Google PageRank	298
L'échantillonnage de Metropolis-Hasting	300
L'échantillonnage de Gibbs	301
MCMC et les biais cognitifs	302
La divergence contrastive et les rêves	304
18 La déraisonnable efficacité de l'abstraction	307
Le <i>deep learning</i> , ça marche !	307
L'apprentissage des <i>features</i>	309
La représentation vectorielle des mots	310

L'expressivité exponentielle*	311
L'émergence de la complexité	313
La sophistication de Kolmogorov*	314
La sophistication est un MAP de Solomonoff !*	315
La profondeur logique de Bennett	317
La profondeur des mathématiques	318
La concision des mathématiques	320
La modularité des mathématiques	321
19 Le cerveau bayésien	325
Le cerveau est incroyable	325
Montagne ou vallée ?	326
Les illusions optiques	328
La perception du mouvement	329
Échantillonnage bayésien	330
Le scandale de l'induction	332
Apprendre à apprendre	333
La bénédiction de l'abstraction	334
Le bébé est un génie	336
Le langage	336
Apprendre à compter	338
La théorie de l'esprit	339
Innée ou acquis ?	340
20 Tout est fiction	343
La caverne de Platon	343
L'antiréalisme	344
La vie existe-t-elle ?	345
L'argent existe-t-il ?	346
La téléologie, impasse scientifique ?	349
Ce que la thèse de Church-Turing dit de la réalité	353
L'antiréalisme (instrumentaliste) est-il utile ?	354
Y a-t-il un monde extérieur au cerveau ?	356
Y a-t-il un chat dans un code binaire ?	357
L'antiréalisme du démon de Solomonoff	358
21 Aux origines des croyances	361
Le scandale des séries divergentes	361
Mais c'est faux, non ?	363
Élève officier	364
Mon périple en Asie	365
Tous des monstres en puissance ?	367
Les histoires ont plus d'effet que les chiffres	368
Les superstitions	370
L'évolution darwinienne des idéologies	371
Croire les superstitions est utile	373

La magie de YouTube	375
Le périple continue	376
22 Au-delà du bayésianisme	379
Le bayésien n'a pas de morale	379
La morale (sélectionnée par la sélection) naturelle	380
Inconscients de nos morales	382
Des bâtons et des carottes	385
La morale du plus grand nombre ?	386
La morale déontologique	388
Le savoir est-il une fin raisonnable ?	390
L'utilitarisme	392
La <i>conséquentialiste bayésienne</i>	394
Le mot de la fin	397

La théorie des probabilités n'est, au fond, que le bon sens réduit au calcul ; elle fait apprécier avec exactitude ce que les esprits justes sentent par une sorte d'instinct, sans qu'ils puissent souvent s'en rendre compte.

Pierre-Simon Laplace (1749-1827)

1

Mon voyage initiatique

Collé par un étudiant

À l'issue d'un cours de probabilités et statistiques que je donnais à l'École Polytechnique de Montréal, un étudiant *troll* vint me coller avec une énigme d'apparence triviale. Un homme a deux enfants. Au moins l'un d'eux est un garçon. Quelle est la probabilité que l'autre soit également un garçon ?

Après quelques secondes de réflexion, je réussis à trouver la bonne réponse à cette question — qui, comme on va le voir plus loin, n'est pas $1/2$. L'étudiant acquiesça, et enchaîna avec une seconde énigme. Supposez que vous appreniez maintenant que l'homme a au moins un garçon né un mardi. Que devient la probabilité que l'autre enfant soit également un garçon ?

Cette fois-ci, ma réponse fut une mauvaise réponse. L'étudiant m'avait collé.

Le réflexe usuel est certainement de penser que ces deux énigmes ne sont que des jeux mathématiques. Il y a en effet une bonne réponse, mais il ne s'agit d'une bonne réponse que dans un cadre mathématique rigide et restreint. On retrouve ce genre de problèmes en exercice ou à l'examen à l'école. Mais il ne s'agit *que* de mathématiques.

Pourtant, l'énigme de l'étudiant *troll* n'est qu'une version ultra-simplifiée de nombreuses réflexions qui encombrant nos quotidiens. Faut-il croire un diagnostic médical ? La présomption d'innocence est-elle justifiée ? Les juges font-ils de la discrimination raciale ? Faut-il craindre le terrorisme ? Peut-on

généraliser à partir d'un exemple ? À partir de mille exemples ? Un million ? L'argument d'autorité a-t-il une quelconque valeur ? Faut-il faire confiance aux marchés financiers ? Les OGM sont-ils nocifs ? En quoi la science aurait-elle plus « raison » que les « pseudo-sciences » ? Les robots vont-ils conquérir le monde ? Faut-il condamner le capitalisme ? Faut-il croire en l'existence de Dieu ? Qu'est-ce que le bien et le mal ?

Pour beaucoup, ces questions n'ont absolument rien à voir avec les mathématiques. Et en effet, les mathématiques seules sont impuissantes face à de telles questions. Vous ne résoudrez pas le problème de la faim dans le monde uniquement en prouvant des théorèmes. Mais il y a fort à parier que les mathématiques peuvent contribuer à mieux structurer notre réflexion, à mieux comprendre les tenants et les aboutissants et à fournir des solutions inattendues. Ce n'est sans doute pas un hasard si de nombreuses disciplines se voient de plus en plus mathématisées — y compris l'aide humanitaire¹.

Malgré le pullulement de nombreux modèles mathématiques, il semble que la plupart d'entre nous persiste à vouloir distinguer le « monde réel » des cours académiques que l'École nous force à suivre. En particulier, le monde réel, pense-t-on souvent, dépasse de loin le cadre des mathématiques, si bien que les théorèmes des mathématiques ne semblent jamais devoir, ni même pouvoir, s'appliquer au monde réel. À quel point faut-il être idiot pour penser que les mathématiques ont un quelconque mot à dire sur l'égalité devant la loi² ?

Cette défiance de l'utilité des mathématiques ne se réduit d'ailleurs pas à un réflexe de mauvais élève. Moi-même, pendant les années qui suivirent mon échec face à mon étudiant *troll*, je ne me suis pas rendu compte que cet échec mathématique révélait mon incapacité à raisonner correctement sur le monde réel. Moi-même, je ne compris pas qu'une meilleure compréhension de cette énigme m'aiderait à mieux écouter les conseils de mes amis baroudeurs pour mieux choisir ma prochaine destination de vacances — on y reviendra.

Sur les traces du bayésianisme

Certes, c'est le soir même que je résolus l'énigme de l'étudiant *troll* — au prix de mystérieux calculs obscurs. Mais ce n'est que trois ans plus tard, début 2016, au moment où je me suis intéressé de plus près au débat qui oppose les statisticiens fréquentistes aux statisticiens bayésiens³, que je commençai à vraiment méditer l'énigme de l'étudiant *troll* et, surtout, à la sortir d'un cadre purement mathématique.

¹  [A Set-Partitioning Formulation for Community Healthcare Network Design in Underserved Areas](#) | M. Cherklesly, M.E. Rancourt et K. Smilowitz (2017)

²  [Partager un gâteau, c'est pas du gâteau !](#) Démocratie 22 | Science4All | L.N. Hoang (2017)

³  [Les statistiques à l'heure du Big Data](#) | CESP Villejuif | L.N. Hoang (2016)

En particulier, pendant les deux ans qui suivirent, presque une fois par jour, je me mis à réfléchir à la formule magique qui résolvait cette énigme. Pour mon plus grand bonheur, petit à petit, cette formule mystérieuse commença à me révéler ses secrets. Lentement mais sûrement, cette formule lumineuse me séduisit, au point de changer la manière dont je concevais le monde, les sciences et la connaissance. Au fil des mois, je finis par être submergé par la sublime élégance de cette formule indomptable. Ce fut trop. Il me fallait dédier tout un livre à son sujet. C'est ainsi que fin 2016, je me lançai dans l'écriture du livre que vous venez d'entamer.

La formule indomptable dont je parle, c'est celle que j'aime appeler pompeusement la formule du savoir. Mais les mathématiciens, statisticiens et informaticiens la connaissent davantage sous le nom de formule de Bayes.

La formule de Bayes est un théorème mathématique d'une simplicité remarquable. Il s'agit d'une équation compacte, que l'on enseigne dès le lycée. Sa preuve ne fait qu'une seule ligne, et ne repose que sur la connaissance de la multiplication, de la division et de la notion de probabilité. En particulier, elle paraît beaucoup plus facile à apprendre que de nombreux autres concepts de mathématiques que l'on demande aux lycéens et aux étudiants du supérieur de maîtriser.

Et pourtant. J'ose prétendre que même les meilleurs mathématiciens ne comprennent pas cette formule de Bayes — et il y a même des mathématiciens qui expliquent notre incapacité à saisir cette formule ! Sans aller jusque-là, il ne fait aucun doute à mes yeux que *je* ne comprends toujours pas la formule de Bayes. D'ailleurs, si j'avais compris la formule de Bayes au moment où j'enseignais le cours de probabilités et statistiques, j'aurais immédiatement vu le lien entre le fait qu'un garçon est né un mardi et le sexe de son frangin. J'aurais répondu du tac au tac à mon étudiant *troll*. Il ne m'aurait pas collé.

Depuis près de deux ans, je torture mon esprit pour ne plus jamais me faire coller ainsi. Je veux connaître, comprendre, sentir la formule de Bayes. J'ai déjà beaucoup appris, et je continue à apprendre. Je médite la formule de Bayes presque quotidiennement, comme s'il s'agissait d'une divinité à qui je me devais de consacrer une partie de mes journées. Et quel bonheur que de méditer cette formule ! Loin d'être un effort répétitif, ces méditations semblent avoir continuellement alimenté ma curiosité, me chuchotant, secrètement et au compte-gouttes, ses implications inattendues.

Au bout de longs mois de réflexions, j'ai fini par me laisser convaincre que peu d'idées étaient aussi profondes que la formule de Bayes. Je suis tombé amoureux de cette formule. Au point que je prétends volontiers aujourd'hui que la « rationalité » se réduit essentiellement à l'application de la formule de Bayes — auquel cas personne n'est rationnel ! Tel est, en tout cas, le fondement de ce que l'on pourrait appeler la philosophie de Bayes, ou *bayésianisme*.

Une philosophie unifiée du savoir

N'ayant pas encore eu le temps de vous présenter la formule de Bayes, je suis contraint de rester volontairement flou sur ce qu'est le bayésianisme pour l'instant. Mais s'il fallait le résumer en trois phrases maladroitement, je donnerais la définition suivante. Le bayésianisme consiste à supposer que tout modèle, toute théorie ou toute conception de la « réalité » n'est que croyance, fiction ou poésie ; en particulier, « tous les modèles sont faux ». Les données empiriques doivent ensuite nous forcer à ajuster l'importance ou la *crédence* que l'on assigne aux différents modèles. De façon cruciale, la manière dont ces crédences sont ajustées doit obéir aussi rigoureusement que possible à la formule de Bayes.

J'ai longtemps rejeté la pertinence de cette philosophie du savoir. Elle semble discréditer toutes notions de réalité ou de vérité, pourtant si chères à tant de scientifiques. Cependant, elle semble aussi parfaitement coller avec ce que disait le physicien et prix Nobel Richard Feynman⁴ : « je peux vivre avec le doute et l'incertitude. Je peux vivre sans savoir. Je pense qu'il est beaucoup plus intéressant de vivre sans savoir que d'avoir des réponses qui pourraient être fausses. J'ai des réponses approximatives, j'ai des croyances plausibles avec différents degrés de certitude sur différents sujets. Mais je ne suis absolument sûr de rien. Et il y a beaucoup de choses dont je ne sais absolument rien. Mais je ne suis pas obligé d'avoir une réponse. Je ne me sens pas effrayé par le fait de ne pas savoir. »

Vous pourriez trouver ce point de vue exaltant. Ou vous avez peut-être envie de rejeter en bloc cette approche de la connaissance. Cependant, avant un éventuel rejet ou une potentielle adhésion au bayésianisme, je ne peux que vous encourager à d'abord prendre le temps de longuement méditer la formule de Bayes et ses conséquences.

Dans ce livre, malheureusement, le guide principal que je serai n'a qu'une compréhension très incomplète de cette équation. Pour nous aider dans nos réflexions, je vais invoquer un personnage fictif (féminin), la *pure bayésienne*, et on cherchera à imaginer comment cette *pure bayésienne* réagirait à divers contextes. Plutôt que moi-même, c'est cette *pure bayésienne* que je vous invite à mettre à l'épreuve. C'est d'ailleurs ce que l'on ne cessera de faire dans ce livre. On va multiplier les expériences de pensée, qui seront autant de défis que la *pure bayésienne* devra relever. Et on prendra le soin de guetter, juger et critiquer les diverses réactions de cette *pure bayésienne* — même si ces critiques se transformeront souvent bien vite en celles de notre intuition et de notre inlassable excès de confiance.

Si le premier bayésien de l'histoire digne de ce nom, le géant Pierre-Simon Laplace, n'avait qu'une description partielle de cette *pure bayésienne*, cela fait déjà un demi-siècle que tous les calculs, réflexions et prédictions de la *pure bayésienne* ont été rigoureusement décrits par le génialissime Ray Solomonoff.

⁴  [The Feynman Series - Beauty](#) | Reid Gower (2011)

Malheureusement, et comme on le verra en détails, la *pure bayésienne* que décrit Solomonoff semble nécessairement devoir enfreindre les lois de la physique — en particulier la thèse de Church-Turing dont on reparlera.

Voilà qui nous contraint à un bayésianisme inéluctablement approché, que je qualifierai de *pragmatique*. Ce *bayésianisme pragmatique*, qui se distingue fortement du bayésianisme pur par ses exigences de calculabilité (rapide), sera incarné par un autre personnage fictif (masculin), que j'appellerai *bayésien pragmatique*. Malheureusement (ou pas !), ma description du *bayésien pragmatique* sera très incomplète, puisque le bayésianisme pragmatique est encore un énorme champ de recherche très ouvert — et il n'est pas dit qu'il puisse un jour être entièrement résolu.

Comme vous commencez sans doute à le deviner, comprendre la *pure bayésienne* et le *bayésien pragmatique* ne sera pas chose aisée. Pour ce faire, il va nous falloir aborder de nombreux concepts fondamentaux des mathématiques, de la logique, des statistiques, de l'informatique, de l'intelligence artificielle, voire des notions de physique, de biologie, de neurosciences, de psychologie et d'économie. On devra parler de logarithme, de contraposition, de *p-value*, de complexité de Solomonoff et de réseaux de neurones, mais aussi d'entropie, de l'évolution darwinienne, des faux souvenirs, des biais cognitifs et des bulles financières. Et puis, on va aussi utiliser de nombreux exemples de l'histoire des sciences pour mettre nos deux héros fictifs au défi.

Oui, je sais, ça fait beaucoup à savoir pour comprendre la formule de Bayes...

Mais ça tombe bien, puisque j'aime expliquer les sciences modernes à mes heures perdues — à tel point que j'ai lancé ma chaîne YouTube appelée Science4All ! Du coup, plutôt que de lire ce livre comme un livre de philosophie, je vous invite à le lire (aussi) comme un livre qui promeut les sciences et les mathématiques. D'ailleurs, sur le chemin du bayésianisme, je n'hésiterai pas à faire quelques détours à travers les méandres des sciences, avec l'objectif secret de vous donner envie d'aller plus loin dans votre apprentissage des théories scientifiques !

Mais revenons-en à la philosophie pour l'instant. Comme vous l'aurez compris, j'ai fini par succomber aux sirènes bayésiennes. Après de longs mois de réflexions, sans l'avoir anticipé, le bayésianisme m'a séduit au point que je me sente obligé de vous en parler. J'ai sans cesse trouvé la *pure bayésienne* bougrement intelligente, si bien que j'aspire de plus en plus à lui ressembler... Et même bien après le début de l'écriture du livre, je n'ai cessé de découvrir, encore et encore, d'innombrables merveilles époustouflantes au sujet de ce qui est depuis devenu mon équation mathématique préférée.

Quand j'ai commencé l'écriture du livre, j'étais un bayésien enthousiaste. Désormais, je suis un bayésien convaincu. J'irai même jusqu'à me décrire bayésien extrémiste, notamment par opposition à d'autres qui se disent aussi bayésiens. Mais surtout, j'aimerais un jour devenir un bayésien compétent. Je rêve d'être capable d'appliquer la formule de Bayes, car je suis convaincu que ce n'est qu'ainsi que je pourrai enfin devenir un être rationnel !

Ce qui est amusant, et peut paraître embêtant, c'est que l'élan émotionnel que cette formule de Bayes a insufflé en moi a des airs de délire irrationnel. Je ne peux pas le nier. Je suis même sûr d'être victime d'un énorme biais cognitif qui a conduit à ma sacralisation de la formule de Bayes. Après tout, il m'est impossible d'être indifférent au fait d'avoir découvert par moi-même nombre de secrets de cette formule — même si bien d'autres les ont découverts un demi-siècle avant moi.

Ceci étant dit, conscient de ce biais, je vous promets avoir lutté — et je lutte encore — contre la *pure bayésienne*. J'ai sans cesse essayé de la mettre en défaut ; j'ai sans cesse essayé de gagner un débat contre elle. En vain.

Une alternative à la méthode scientifique

En mathématiques, dès qu'une conjecture semble tenir la route, on s'empresse de la prouver pour l'ériger au rang de théorème. Ce n'est pas loin d'être le cas du bayésianisme !

On verra ainsi que le théorème de Jaynes-Cox prouve que le bayésianisme est la seule généralisation de la logique aristotélicienne capable de traiter la notion de plausibilité de manière cohérente. Le théorème de complétude de Solomonoff, lui, prouve que la *pure bayésienne* finira par toujours repérer toutes les régularités dans un jeu de données, si ces régularités existent. Quant au théorème des gains espérés nécessaires des informations additionnelles, lui montre que la *pure bayésienne* ne perd jamais à acquérir plus de données. Enfin, la théorie statistique de la décision montre que les inférences bayésiennes sont essentiellement les seules méthodes d'apprentissage admissibles, dans le sens où une méthode n'est jamais dominée par une autre, si et seulement si, elle revient à appliquer la formule de Bayes⁵.

À ces théorèmes s'ajoutent de nombreux théorèmes dont on ne parlera malheureusement pas dans ce livre. Il y a par exemple le théorème de Teller⁶-Skyrms⁷, qui montre que seul un bayésien ne sera jamais déficitaire dans un problème de « pari hollandais⁸ ». Mieux encore, le théorème de Joyce⁹ montre que l'on gagne à rendre nos croyances conformes aux lois des probabilités, comme l'impose le bayésianisme. Ces différents résultats sont ainsi parfaitement illustrés par le paradoxe des deux enveloppes¹⁰.

⁵Ces théorèmes sont abordés respectivement dans les chapitres 3, 7, 9 et 12.

⁶ *Conditionalisation and observation* | Synthèse | P. Teller (1973)

⁷ *Dynamic Coherence and Probability Kinematics* | Philosophy of Science | B. Skyrms (1987)

⁸ *Argent, risques et paradoxes* | Démocratie 12 | Science4All | L.N. Hoang (2017)

⁹ *A Nonpragmatic Vindication of Probabilism* | Philosophy of Science | J. Joyce (1998)

¹⁰ *Inégalité bayésienne* | Axiome 9 | T. Giraud et L.N. Hoang (2018)

- ▶ *Rationnellement irrationnels* | Démocratie 11 | Science4All | L.N. Hoang (2017)
- ▶ *Le paradoxe de la morale* | Démocratie 25 | Science4All | S. Debove et L.N. Hoang (2017)
- ▶ *La grande Histoire des petites histoires* | Démocratie 26 | Science4All | S. Mombo et L.N. Hoang (2017)
- ▶ *L'instinct tribal* | Démocratie 27 | Science4All | L.N. Hoang (2017)
- ▶ *L'utilitarisme* | Démocratie 28 | Science4All | T. Giroud et L.N. Hoang (2017)
- ▶ *7 arguments CONTRE la démocratie* | Démocratie 30 | Science4All | L.N. Hoang (2017)
- ▶ *Wikipedia et l'épistocratie* | Démocratie 31 | Science4All | L.N. Hoang (2017)
- ▶ *Fat Tony et Dr John (biais-variance)* | IA 12 | Science4All | G. Mitteau et L.N. Hoang (2018)

- 🎧 *Informatique et éthique* | Podcast Science 266 | G. Dowek et S. Abiteboul (2016)
- 🎧 *Utilitarisme artificiel* | Axiome 3 | T. Giraud et L.N. Hoang (2017)

Références en anglais

- 📖 *The Myth of the Rational Voter: Why Democracy Always Chooses Bad Policy* | Princeton University Press | B. Caplan (2007)
- 📖 *Thinking Fast and Slow* | SpringerFarrar, Straus and Giroux | D. Kahneman (2013)
- 📖 *The Righteous Mind: Why Good People are Divided by Politics and Religion* | Vintage | J. Haidt (2013)
- 📖 *Superintelligence: Paths, Dangers, Strategies* | Oxford University Press | N. Bostrom (2014)
- 📖 *Against Democracy* | Princeton University Press | J. Brennan (2016)
- 📖 *The Big Picture: On the Origin of Life, Meaning and the Universe Itself* | Dutton | S. Carroll (2016)

- 📖 *Motivated Numeracy and Enlightened Self-Government* | Behavioural Public Policy | Dan Kahan, Ellen Peters, Erica Cantrell Dawson, Paul Slovic (2017)
- 📖 *Human-level control through deep reinforcement learning* | Nature | V. Mnih, K. Kavukcuoglu, D. Silver, A. Rusu, J. Veness, M. Bellemare, A. Graves, M. Riedmiller, A. Fidjeland, G. Ostrovski, S. Peterson, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg et D. Hassabis (2015)

- ▶ *New Coke - A Complete Disaster?* Company Man (2017)
- ▶ *Why you think you're right — even when you're wrong* | TED | J. Galef (2016)

- ▶ *Politics and Numbers* | Numberphile | J. Grime (2013)
- ▶ *Why Asimov's Laws of Robotics Don't Work* | Computerphile | R. Miles (2015)
- ▶ *Can You Solve This?* Veritasium | D. Muller (2014)
- ▶ *The Illusion of Truth* | Veritasium | D. Muller (2016)