

MATHIEU AGELOU

MAÎTRISER L'INCERTAIN

STATISTIQUES ET PROBABILITÉS



CNRS EDITIONS

Mathieu Agelou

Maîtriser l'incertain

Statistiques et probabilités

CNRS ÉDITIONS

15, rue Malebranche – 75005 Paris

*« Il y a trois sortes de mensonges : les mensonges,
les sacrés mensonges et les statistiques. »*
Mark Twain

Introduction

Savez-vous qu'aux États-Unis le nombre de morts par noyade en piscine serait lié à la quantité d'électricité d'origine nucléaire produite ? En effet, la corrélation est de plus de 90 %¹ ! Voilà un nouvel argument en faveur d'un arrêt rapide des centrales nucléaires !

Que voulait dire Mark Twain quand il déclarait « *Il y a trois sortes de mensonges : les mensonges, les sacrés mensonges et les statistiques* » ? Que les statistiques, cette branche alors plutôt récente des mathématiques, ne sont pas réellement ce qu'elles prétendent être ? Que leurs conclusions sont fausses ? Qu'elles peuvent faire dire ce que l'on souhaite aux données ? Ou plutôt que leurs résultats, pourtant justement obtenus, vont généralement à l'encontre de nos jugements instinctifs ?

Quelle que fût l'intention de Twain, on comprend qu'il puisse paraître légitime de placer les statistiques dans le camp des mensonges. Notre cerveau a d'une part beaucoup de mal à « bien » raisonner lorsqu'il s'agit de statistiques, qui, d'autre part, ne sont pas toujours simples à manipuler correctement, même pour qui elles sont familières. Il faut cependant y regarder de plus près. La théorie des probabilités et la théorie statistique sont des systèmes mathématiques désormais très bien connus, axiomatisés, qu'aucun mathématicien actuel ne songerait à remettre

1. Site *Spurious correlation* : <https://www.tylervigen.com/spurious-correlations>

en question². En revanche, des débats ont lieu ailleurs : l'interprétation que l'on en tire en termes épistémiques reste une question encore largement débattue (notamment avec l'opposition entre fréquentistes et bayésiens, dont nous parlerons) ; et surtout leur utilisation, parfois même par des chercheurs, n'est pas toujours correcte et souvent mal reçue par le public. Une analogie peut être faite avec un autre domaine de la science qui présente ces mêmes problèmes : la physique quantique. D'un point de vue théorique, elle est parfaitement cohérente et se classe parmi les théories physiques les mieux vérifiées. D'un point de vue épistémique, elle est sujette à plusieurs interprétations sur ce qu'elle nous révèle de la nature profonde de la matière : est-elle une image fidèle de ce qu'est la réalité en soi ? Ou simplement un outil mathématique qui se contente de bien décrire les phénomènes ? Enfin, elle est souvent incomprise par le grand public car elle contrevient à l'intuition commune, et est parfois même honteusement détournée (songeons par exemple à la « médecine quantique » qui n'a de quantique, et de médecine, que le nom !).

Le but de cet ouvrage n'est pas de dire que les chiffres, parce qu'ils sont des chiffres, ont toujours raison, mais plutôt de soutenir qu'une analyse statistique bien pensée et bien menée, même si elle mène à des résultats contre-intuitifs, possède une force démonstrative qu'il convient à tout le moins de ne pas négliger, que ce soit d'un point de vue individuel si l'on se veut rationnel et cohérent, mais, surtout, d'un point de vue collectif. À l'inverse, il conviendra de toujours garder à l'esprit que, mal conçus ou mal interprétés, des calculs statistiques peuvent conduire à des conclusions fausses et par conséquent, à des problèmes potentiellement graves.

2. Cela ne signifie pas que ce domaine est achevé, que tout a été prouvé, et qu'aucune recherche n'est plus nécessaire. Au contraire, c'est un champ de recherche très fécond et de nombreux groupes de par le monde travaillent encore à son développement.

Si nous avons choisi de consacrer cet ouvrage à la statistique (et nécessairement à sa sœur siamoise, la probabilité), c'est qu'elle apparaît partout, dans toutes les branches de la science, bien qu'à des degrés divers et pour des raisons différentes, mais toujours dans le même but : maîtriser l'aléa, prendre en considération l'incertitude et tenter de la réduire. Car, malgré le préjugé populaire selon lequel les faits scientifiques dans les laboratoires sont absolument déterminés et univoques, il faut avoir à l'esprit que ce n'est pas le cas, que les causes des phénomènes sont rarement uniques, et que la grande majorité des données acquises lors d'observations ou d'expériences sont entachées d'erreurs et doivent d'abord subir un traitement statistique, duquel émergera, ou non, une conclusion qui sera alors publiable.

Au-delà de la simple présentation de l'histoire et de l'utilisation des statistiques dans des domaines aussi divers que la physique, la médecine ou l'économie, c'est finalement la rationalité de la méthode scientifique que nous souhaitons mettre en avant. Or de nombreux exemples, dans la presse notamment, nous prouvent tous les jours qu'au contraire, les résultats de la science ne sont pas compris et sont, selon le « camp idéologique » du locuteur, érigés en dogmes indépassables ou ravalés au rang de simple opinion.

Ce livre est organisé de la manière suivante. Le premier chapitre présente une histoire du concept de probabilité et de l'émergence de la statistique, intimement liées à l'idée de risque et de prise de décision en présence d'un environnement incertain. Ce n'est qu'au xvii^e siècle que les premiers calculs sont réellement menés, au xviii^e et xix^e siècles qu'on les appliquera à plusieurs domaines de la connaissance, et il faudra attendre la fin du premier tiers du xx^e siècle pour que les bases mathématiques de la probabilité soient rigoureusement établies.

Le deuxième chapitre s'intéresse justement aux aspects fondamentaux des probabilités et des statistiques du point de vue théorique. Il pourra sembler complexe mais des exemples

illustrent les notions clés et explicitent notamment les tests d'hypothèse, qui sont un outil majeur dans les études scientifiques aujourd'hui.

Le chapitre trois tente de nous dédouaner (un peu) de notre compréhension (généralement) erronée des statistiques. Nous verrons en effet que nous sommes naturellement de forts mauvais statisticiens à cause de nombreux biais cognitifs auquel nous sommes tous soumis à des degrés divers, statisticiens de métier compris.

Enfin, le dernier chapitre expose plusieurs domaines dans lesquels les statistiques interviennent et jouent un grand rôle, dans la recherche scientifique bien sûr mais également dans les décisions qui découlent de l'interprétation faite de ses résultats. Une attention particulière sera portée à l'épidémiologie et la recherche clinique dont les résultats fondent les politiques publiques en matière de santé alors même que les retranscriptions qui en sont faites dans la presse généraliste prennent parfois des raccourcis abrupts, au prix de la vérité.

Chapitre premier

Histoire de la probabilité : de l'aléatoire dans la vie

L'histoire de la probabilité est une riche et longue aventure que la plupart des historiens des sciences font débiter au milieu du xvii^e siècle. C'est à tout le moins le moment où des mathématiciens et philosophes (qui, à cette époque, cumulaient souvent les deux fonctions, ce qui est devenu d'une extrême rareté aujourd'hui) se sont réellement penchés sur la formalisation et la résolution de problèmes liés à des événements aléatoires.

Probabilité : naissance d'une nouvelle forme de rationalité

Avant que ce concept n'acquît une existence réellement autonome, les êtres humains étaient bien entendu déjà confrontés à ces phénomènes aléatoires, à l'incertitude de l'avenir, que l'on appelait les « futurs contingents ». La question s'est donc toujours posée de savoir quelle était la meilleure attitude à adopter dans ces situations où l'on ignore ce qui va advenir. Si le calcul des probabilités nous apparaît aujourd'hui un élément indispensable pour régler notre conduite de manière rationnelle (du moins au niveau de la collectivité), on peut s'étonner qu'il ait fallu attendre le xvii^e siècle pour le voir se développer.

Les débats sur les conditions et les raisons de cette émergence tardive sont nombreux. Plusieurs arguments sont généralement

avancés par les historiens des sciences, sans qu'aucun ne fasse réellement consensus. Ian Hacking (1936-) par exemple, historien et philosophe des sciences, dans son ouvrage d'une rare érudition, *L'Émergence de la probabilité*³, la situe précisément dans la décennie encadrant 1660, avec au moins deux parents directs principaux, Christian Huygens (1629-1695) et Blaise Pascal (1623-1662). Pour Hacking, les conditions d'apparition n'existaient tout simplement pas jusque-là, sans que les raisons invoquées soient toutes déterminantes. Il cite tour à tour l'imprégnation du déterminisme qui accompagnait la révolution scientifique récente et excluait donc l'effet du hasard ; une forme de piété qui empêchait que l'on cherchât à percer les intentions divines derrière les phénomènes contingents ; l'argument économique externaliste qui prône que les avancées en science sont le faits d'éléments externes à la science elle-même : ce serait dans ce cas le développement des assurances et des rentes viagères qui aurait accéléré les réflexions ; l'absence jusqu'alors d'une arithmétique et d'un système de numération performants ; le développement à cette époque des jeux de hasard, formidable terrain de jeux... mathématiques, etc. Aucune de ces raisons ou circonstances, prises isolément, ne semble l'emporter et c'est vraisemblablement une combinaison de toutes qui explique l'avènement, à ce moment particulier de l'histoire de la pensée, du concept et de la mathématisation de la probabilité. Pour Lorraine Daston (1951-), autre historienne et philosophe des sciences, c'est principalement les questions d'ordre juridique qui ont modelé le concept de probabilité parce qu'elles obligent à prendre des décisions « justes » alors que tous les éléments ne sont pas connus – en situation d'incertitude donc⁴. C'était le domaine

3. Ian Hacking, *L'Émergence de la probabilité*, 2002, Le Seuil (traduction de l'édition originale *The Emergence of Probability*, Cambridge University Press, 1975).

4. Lorraine Daston, *Classical Probability in the Enlightenment*, 1988, Princeton University Press.

de ce que l'on a appelé alors les « contrats aléatoires », c'est-à-dire toutes les transactions, les échanges, pour lesquels les avantages ou les pertes ne sont pas connus à la signature du contrat, et dépendent d'un événement futur et incertain. Elle inclut également les travaux mathématiques sur l'analyse combinatoire dans les jeux de hasard, car c'est principalement la recherche de l'équité, de la « justice », dans ces jeux, qui était visée.

Les nombreuses occurrences de l'incertain dans nos vies se trouvent illustrées par les tout aussi nombreux termes utilisés pour en capturer une facette particulière. Aléa, aventure, chance, hasard, péril, fortune, sort, risque. Chacun de ces mots, dont l'étymologie nous éclaire, exprime une nuance ayant trait à ce qui peut arriver et que nous ne maîtrisons pas. Ainsi, l'aléa, le hasard et la chance proviendraient des jeux de dés : *alea* signifiant en latin directement « dé » et par extension « jeu de dé » ou « jeu de hasard », et « chance » dérivant de l'ancien français *chaance*, du latin *cadentia*, participe présent du verbe *cadere*, qui signifie « tomber ». Au départ, la chance est donc le résultat de la chute des dés. « Hasard », lui, dérive de l'arabe *az-zahr*, le dé à jouer, ou *yasara*, la forme verbale « jouer aux dés ». Les autres sont en revanche davantage liés aux événements qui nous arrivent. « Péril » vient du latin *periculum* désignant un essai, une tentative, et son sens s'est peu à peu porté sur l'expérience risquée, l'essai dangereux. « Fortune », du latin *Fortuna*, la déesse présidant au destin des hommes et des peuples qui pouvait faire advenir bonheurs ou malheurs selon sa volonté. Le mot a pris ensuite un aspect généralement positif. L'« aventure », du latin *advenire*, est ce qui advient, et désigne d'abord la destinée (ce qui doit arriver et que sont censées nous dévoiler les diseuses de bonne aventure), puis ce qui arrive par hasard, inopinément (ce sont alors les aventures chevaleresques du Moyen Âge). Le « sort » enfin, coquin ou mauvais, du latin *sors*, désignait au départ divers procédés de tirage au sort dont pouvaient

se servir les oracles pour consulter les dieux. Il a pris ensuite un sens plus proche de la destinée aveugle, liée au hasard⁵.

On le voit à travers ces termes, l'incertitude dans laquelle nous plonge les futurs contingents a longtemps enfermé les hommes dans une attitude passive vis-à-vis de leur destin. Celui-ci était écrit, par un hasard aveugle ou un décret divin, et ne pouvait dès lors pas faire l'objet d'un traitement rationnel et, moins encore, mathématique. « *On tire au sort avec un dé, mais le Seigneur décide de tout* », dit le livre des Proverbes de l'Ancien Testament (16:33). Chercher des lois au hasard revenait en quelque sorte à vouloir connaître les desseins de Dieu. C'est pourtant Pascal, janséniste et en cela défenseur de la théorie de la grâce divine et de la prédestination, qui sera, au XVII^e siècle, à l'origine d'un calcul rationnel pour un pari, celui de la croyance, ou non, en Dieu. Jusqu'à la Renaissance, la connaissance s'entendait principalement sur le mode de la certitude absolue, qui ne pouvait être atteinte que par des raisonnements purement déductifs, comme le syllogisme aristotélicien. La période classique qui lui succède est marquée par une nouvelle attitude, active cette fois, vis-à-vis de l'incertitude intrinsèque de nos vies, et par l'émergence d'une nouvelle forme de rationalité. Elle remet en cause les autorités scolastiques du Moyen Âge et, reconnaissant que la certitude absolue n'était possible qu'en de rares cas, s'en remettait à la recherche d'un raisonnement qui, même s'il ne menait qu'à des degrés plus ou moins élevés de confiance, permettait néanmoins de conduire son action de manière à la fois raisonnable et juste.

Sans rentrer dans ces débats de spécialistes, nous décrirons brièvement deux sources principales qui ont, sinon présidé à la naissance du concept de probabilité, du moins accompagné son émergence : les jeux de hasards et les contrats aléatoires.

5. On pourra retrouver ces étymologies, plus détaillées, dans le Trésor de la langue française, notamment dans sa version informatisée <http://atilf.atilf.fr/>.

Probabilité : les mathématiques pour les jeux de hasard et la notion d'espérance

« Un problème relatif aux jeux de hasard, proposé à un austère janséniste par un homme du monde, a été l'origine du calcul des probabilités » : c'est ainsi que Siméon Denis Poisson (1781-1840), mathématicien et physicien français, résume la naissance du traitement mathématique des probabilités. L'austère janséniste est bien sûr Pascal, et l'homme du monde Antoine Gombaud, dit le « chevalier de Méré » (1607-1684). Ce dernier est un libertin féru de mathématiques et amateur de jeux de hasards. Il fréquente les salons et notamment celui du duc de Roannez où il rencontre Pascal et lui soumet un problème, resté célèbre sous le nom de *problème des partis*.

Le problème, tel que le décrit Pascal dans sa correspondance⁶ avec le mathématicien Pierre de Fermat (1607-1665), est le suivant : deux joueurs misent chacun 32 pistoles dans un jeu qui se gagne en trois manches, mais décident de se séparer en cours de partie. Selon le nombre de manches déjà jouées, quelle doit être la juste répartition des 64 pistoles en jeu⁷ ?

On retrouve ici l'aspect « moral », de justice, dont parlait Lorraine Daston : comment répartir correctement la somme, sachant qu'un certain nombre de parties ont déjà été jouées, et ignorant le résultat des parties restantes, dont le hasard aurait décidé ? Pascal a pour cela mobilisé et développé l'analyse combinatoire, qui existait en partie. Dans sa lettre à Fermat, il part de l'exemple du jeu en trois manches et passe en revue les différents cas (2 manches à 1, 2 manches à 0, 1 manche à 0), avant de généraliser son calcul à un nombre quelconque de parties. Il publiera sa solution l'année suivante dans son *Traité du triangle arithmétique*, dans un chapitre consacré à « l'utilisation

6. Lettre de Pascal à Fermat du mercredi 29 juillet 1654, *Œuvres de Fermat*, tome 2, Gauthier-Villars, 1894, p. 289-296.

7. On comprend ici pourquoi cette question s'appelle problème des *partis* et non *parties*. Nous dirions aujourd'hui problème du partage.

du triangle arithmétique pour déterminer les partis qu'on doit faire entre deux joueurs, qui jouent en plusieurs parties ».

On considère communément qu'il s'agit du point de départ du calcul des probabilités. Mais c'est un peu simplificateur, car ce genre de problèmes, et le parfum de la probabilité (car elle n'est jamais réellement définie), existaient depuis au moins 150 ans et la *Summa de aritmetica* publiée en 1494 par Lucas Pacioli (1445-1517), un mathématicien toscan. Il pensait que les mathématiques étaient un langage universel permettant de régler tout type de problèmes dans de nombreuses disciplines, de l'architecture à la théologie en passant par la médecine. Et c'est au détour d'un chapitre consacré aux militaires qu'il pose le problème suivant : « *Une brigade joue à la paume ; il faut 60 pour gagner, et chaque coup vaut 10 ; l'enjeu est de dix ducats. Un incident survient, qui force les soldats à interrompre la partie commencée, alors que le premier camp a gagné 50 et le second 20. On demande quelle part de l'enjeu revient à chaque camp*⁸. » Il s'agit précisément du problème du chevalier de Méré, auquel il faut ajouter l'hypothèse (cruciale) que les deux équipes ont exactement le même niveau et que c'est donc bien le hasard qui préside aux chances de gagner. Pacioli proposa une solution, qui se révéla fautive. À sa suite, d'autres mathématiciens italiens, comme Tartaglia (1499-1557) et Jérôme Cardan (1501-1576), proposèrent des solutions, fautives elles aussi. Mais Cardan, joueur invétéré, fut parmi les premiers qui firent le lien entre probabilité et analyse combinatoire, en résolvant un autre problème, plus simple, celui des dés. Dans son *Livre sur les jeux de hasard (Liber de ludo aleae)*, il calcula correctement les chances pour un joueur d'obtenir une certaine somme en lançant deux ou trois dés. Il y parvint en dénombrant les combinaisons possibles pour chaque résultat. Il introduisit également la notion

8. Lucas Pacioli cité par Pierre Boutroux, *Les Origines du calcul des probabilités*, Revue du mois, 1908, p 641-654.

de chance comme étant le rapport du nombre de combinaisons favorables sur le nombre total de combinaisons, ce qui est une définition classique de la probabilité. Mais son livre ne fut publié que longtemps après sa mort, en 1663, donc après que Pascal eut lui-même publié sur le sujet.

C'est également grâce aux travaux sur les jeux de hasard que la notion d'espérance mathématique est née. Même si l'idée avait été esquissée par Cardan, puis discutée par Pascal et Fermat, c'est à Huygens que l'on doit sa véritable naissance. Mathématicien et physicien hollandais, il s'intéressait lui aussi aux problèmes que posaient les jeux de hasard depuis un séjour de quelques mois à Paris en 1655, pendant lequel on lui présenta les problèmes des partis. Mais, hasard de l'histoire, croyant Pascal définitivement retiré du monde et des mathématiques à Port-Royal, et Fermat habitant Toulouse, Huygens n'a alors rencontré ni l'un ni l'autre. C'est donc seul qu'il arriva à la même solution et publia en 1657 le fruit de ses recherches dans son traité *Du calcul dans les jeux de hasard (De ratiociniis in ludo aleae)*. La question générale qui se posait alors était la suivante : dans un jeu de hasard où l'on parie de l'argent, quelle mise est-elle juste ? Quel est finalement le juste prix du risque que l'on prend en jouant ? La réponse nous paraît assez évidente : elle dépend du montant du gain et des chances que l'on a de le remporter. Cette évidence n'en était alors pas une. Le premier principe posé par Huygens, qui ne va en fait pas de soi, c'est que, même dans un jeu soumis au hasard, une telle mise existe : « *Quoi que dans les jeux de hasards purs les résultats soient incertains, la chance qu'un joueur a de gagner ou de perdre a cependant une valeur déterminée.* » La résolution du problème des partis menait à cette idée, notamment par la notion de justice qu'elle introduisait. L'archétype d'un jeu où l'on connaît le juste prix à payer est celui de la loterie dite « équitable » : elle est par exemple définie par le fait que, si N billets sont vendus, qu'une somme S est en jeu, et que chaque billet a exactement

la même chance d'être tiré, alors le juste prix du billet est S/N . Une loterie équitable est une loterie pour laquelle l'espérance de gain (sur le long terme) est nulle : si l'on vend 10 tickets pour 100 euros de gain, chaque ticket doit valoir 10 € pour qu'elle soit équitable. Après un grand nombre de parties, ni le joueur, ni le vendeur de billets, n'a gagné ou perdu de l'argent. Huygens définit alors l'espérance dans un pari quelconque comme le juste prix que l'on paierait dans une loterie équitable équivalente. Il se trouve que, même si cette conséquence ne se trouve pas directement sous la plume de Huygens, cette valeur est le gain moyen que l'on peut attendre dans un jeu donné si l'on joue un grand nombre de parties. Si l'on paye davantage, alors, en moyenne, on perd de l'argent, et si l'on paye moins, alors, en moyenne on en gagne. Il choisit judicieusement le terme d'*espérance* (*expectatio* en latin, qu'il a finalement préféré à *espoir*, du latin *spes*) qui exprime finalement ce que l'on peut espérer gagner à long terme, si l'on joue. Et le calcul de cette espérance se fait en multipliant chaque gain par le nombre de chances correspondantes, en sommant ces produits, et en divisant le tout par le nombre total des tirages possibles. C'est la proposition III de son traité : « *Avoir p chances d'obtenir A et q chances d'obtenir B me vaut $\frac{pA + qB}{p + q}$* ». Nous reconnaissons la moyenne pondérée, que nous appelons aujourd'hui « l'espérance mathématique », dans le cas d'une variable aléatoire (voir le chapitre suivant pour sa définition mathématique actuelle). Il est intéressant de noter que, finalement, le concept d'espérance est antérieur à celui de probabilité, définie bien plus tard par Pierre-Simon de Laplace (1749-1827) comme le rapport du nombre de cas favorables à un résultat donné sur le nombre total de cas.

Cette nouvelle notion d'espérance, dont la naissance est liée aux jeux de hasard (et d'argent ! – l'argument externaliste possède finalement une certaine force ici), sera par la suite très

féconde, et mise en pratique notamment dans le cas des calculs de rente.

Probabilité : assurances et rentes, la maîtrise du risque

Bien avant que les concepts de probabilité ou d'espérance n'aient fait leur apparition, une notion connexe vit le jour, celle de risque, c'est-à-dire la prise en compte, pour la conduite des actions humaines, à la fois du caractère incertain des événements futurs et de leurs conséquences éventuelles. Elle apparut notamment avec le développement des échanges commerciaux maritimes au XII^e siècle autour de la Méditerranée et se développa dans ce que l'on a appelé les « contrats aléatoires ».

Un des domaines importants dans lesquels les premiers travaux sur les probabilités ont été rapidement utilisés est celui de l'actuariat, autrement dit de l'assurance et des rentes. La raison en est simple : si vous souhaitez assurer des personnes contre un risque particulier sans en être lésé, il convient de bien connaître statistiquement le risque en question afin de demander la somme adéquate à chaque assuré, somme qui permettra à la fois de payer les conséquences de l'événement risqué lorsque celui-ci aura finalement lieu pour l'un ou l'autre des assurés, et éventuellement de vous procurer une marge supplémentaire, qui sera votre revenu.

Si le mécanisme assurantiel semble actuellement aller de soi tant la nature des entités assurées est vaste (depuis une habitation ou une voiture, en passant par des vacances au ski, les mains des pianistes ou les jambes des footballeurs, un emploi, ou même... une vie⁹), il a lui-même une histoire. Si l'on trouve

9. L'assurance-vie représente le premier moyen d'épargne en France avec un encours de 1750 milliards d'euros (<https://www.ffa-assurance.fr/etudes-et-chiffres-cles/assurance-vie-collecte-nette-positive-en-juin-2019> consulté le 3 août 2019).

trace de contrats de prêt et de transfert de risque chez les Babyloniens dans le code de Hammurabi (1750 avant J.-C.), c'est surtout à partir du XII^e siècle, sur le pourtour méditerranéen, et en Italie en particulier, que la technique de l'assurance s'est développée, dans le domaine du commerce. Un des premiers risques assurés était l'affrètement des bateaux commerciaux. Dans l'Antiquité, on pratiquait déjà le « prêt à la grosse aventure » (parfois raccourci en « prêt à la grosse »), qui consistait, pour un marchand, à se faire prêter de l'argent pour affréter un bateau et transporter ses produits. Le taux du prêt pouvait être très important (jusqu'à 30 %) mais le prêteur jouait gros : si la cargaison n'arrivait pas à bon port (à cause d'un naufrage, ou d'attaques en mer, pas si rares à cette époque), il n'était pas remboursé. Ce type de contrat mêlait donc à la fois le prêt et une forme d'assurance.

À partir du XII^e siècle, le principe de l'assurance telle que nous la connaissons se met en place. C'est à cette époque que le concept même de risque voit le jour, comme le montre Sylvain Piron dans un article¹⁰ consacré à l'apparition, dans les contrats de l'époque, du terme *resicum*, qui a donné notre *risque*. Le débat se concentrait alors sur deux points essentiels : la notion de l'action en situation d'incertitude bien sûr, mais également la licéité, morale et religieuse, d'une rémunération spécifiquement liée à cette incertitude. En effet, l'Église catholique considérait comme un péché de gagner de l'argent avec de l'argent. Thomas d'Aquin, et d'autres théologiens avec lui, reprennent la critique d'Aristote à propos de ce qu'il nommait la « chrématistique », c'est-à-dire la doctrine économique qui vise principalement à accumuler des capitaux pour eux-mêmes, et non pour une utilisation « naturelle » qui serait celle d'achats

10. Sylvain Piron, « L'apparition du *resicum* en Méditerranée occidentale, XII^e-XIII^e siècles. Pour une histoire culturelle du risque », in *Genèse, évolution, actualité du concept dans les sociétés occidentales*, sous la direction de E. Collas-Heddeland, M. Coudry, O. Kammerer, A. J. Lemaître, B. Martin, Éditions Histoire et Anthropologie, p. 59-76, 2004.

Deux grandes lois des statistiques	61
<i>Loi des grands nombres</i>	61
<i>Théorème Central-limite</i>	63
<i>Statistiques inférentielles</i>	68
Échantillonnage	68
Estimation de paramètres	69
Test d'hypothèses	73
<i>Le test z et la loi normale</i>	79
<i>La test t et la loi de Student</i>	80
<i>La loi et le test du khi2</i>	83
La valeur p ou p-value	87
Bayésiens contre fréquentistes	90
Chapitre 3 – Pourquoi sommes-nous de piètres statisticiens ?	97
Perception du hasard : jamais deux sans trois ?	98
Erreurs et paradoxes dus à notre conception du hasard	102
<i>Bombardement de Londres</i>	102
<i>Paradoxe des anniversaires</i>	104
<i>Le paralogisme du joueur et le mythe de la main chaude</i>	105
<i>Coïncidences et loi des séries</i>	108
Erreurs dues à notre perception du risque	111
<i>La loi des petits nombres</i>	112
<i>Négligence du taux de base</i>	114
<i>Sophismes du procureur et de l'avocat de la défense</i>	116
<i>Homo œconomicus et théorie des perspectives</i>	119
Les chausse-trappes statistiques	128
<i>Le paradoxe de Simpson</i>	128
<i>Problème des comparaisons multiples</i>	133
<i>Cause et corrélation</i>	136

Chapitre 4 – Applications des statistiques	141
Espérance de vie	141
Sondages d’opinions	145
Épidémiologie : les statistiques au service de la santé	155
<i>Naissance de la discipline</i>	156
<i>Outils statistiques modernes de l’épidémiologie et la médecine fondée sur les preuves</i>	158
Risque relatif et rapport de cote	160
Études cas-témoins	165
Études de cohortes	167
Études randomisées et contrôlées en double aveugle	172
<i>D’une saine interprétation des études épidémiologiques</i>	180
<i>Les clusters de maladies</i>	191
<i>Les méta-analyses – le cas de l’homéopathie</i>	197
<i>Conclusion sur l’épidémiologie</i>	208
p-value : le graal de la recherche contemporaine et crise de la reproductibilité	209
Conclusion	223
Remerciements.....	227