



Éléments d'histoire de la statistique

Jean-Claude Oriol

► **To cite this version:**

| Jean-Claude Oriol. Éléments d'histoire de la statistique. Statistix, Statistix, 2010. <inria-00466297>

HAL Id: inria-00466297

<https://hal.inria.fr/inria-00466297>

Submitted on 23 Mar 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Une approche historique de la statistique : des premiers pas à la « géométrie du hasard » et aux « maîtres de l'erreur »¹

Jean-Claude Oriol²

¹ Nous avons ici la première partie (historique) de la thèse intitulée « Formation à la statistique par la pratique d'enquêtes par questionnaires et la simulation : étude didactique d'une expérience d'enseignement dans un département d'IUT » consultable et téléchargeable sur le serveur HAL à l'adresse <http://tel.archives-ouvertes.fr/tel-00191166/fr/>

² Renseignements (CV etc.) disponibles sur <http://sites.google.com/site/jeanclaudeoriol/>

Aborder ce travail par une étude de l'histoire de la statistique peut apparaître au premier abord une approche banale et sans grand intérêt ; banale car le déploiement d'une question dans l'espace et dans le temps est une approche traditionnelle, et sans grand intérêt en raison des ouvrages que l'on peut citer pour circonscrire l'histoire de la statistique.

Nous nous inscrirons en faux avec ce liminaire car loin de nous laisser guider par cette première partie vers une chronogénèse des concepts statistiques nous marquerons les différences entre une approche que l'on peut lier à l'histoire de l'apparition des outils statistiques et une démarche plus proche d'une topogénèse des concepts statistiques.

Ainsi notre travail donnera une situation privilégiée à l'enquête comme lieu contribuant à la fois aux interrogations précédant la mise en place d'outils théoriques mais également comme sujet et occasion de justifier, comme a posteriori, la mise en place de théories novatrices en leurs temps.

Bien entendu étudier l'histoire de la statistique en tant que mode de connaissance du monde ne peut se faire sans approfondir les outils formels développés par le « calcul des probabilités » puis par la « statistique mathématique » qui en constituent à la fois la preuve devant la communauté scientifique mais aussi qui nourrissent un ensemble de techniques, de résultats et de théorèmes. Nous rencontrerons également dans cette étude les fondements matériels et institutionnels qui ont accompagné le développement de la statistique, justifiant son existence dans le cadre des états où elle assure une fonction à la fois utilitaire et symbolique.

Nous aurons donc à parcourir ce champ sous un triple éclairage : celui de l'évolution de la science, celui de l'organisation des États et celui relatif à la cité, à la chose publique, au gouvernement de l'état ou plus simplement à la politique.

1. Des débuts prometteurs

Il semble qu'à ses débuts la statistique soit à la fois le signe de la naissance d'un pouvoir que ce dernier ait été le fait d'une cité ou du moins d'un état centralisé et d'un pouvoir religieux les deux étant souvent confondus.

1.1. Les systèmes artificiels à mémoire (SAM)

A l'origine du stockage de l'information on a trouvé depuis une centaine d'années un certain nombre d'objets gravés, principalement sur des os ou des bois de rènes au Paléolithique supérieur (environ -35 000 ans en Europe et -60 000 ans en Afrique).

Souvent les premières interprétations de l'utilisation de ces objets étaient qu'ils étaient dans le registre du rite, i.e. fréquemment quand on ne comprend pas la signification ou l'utilisation d'un objet on décide que c'est un objet appartenant au registre du religieux. C'est ce que souligne Francesco d'Errico³ lorsqu'il met en évidence la difficulté de l'interprétation de ces objets archéologiques. Ainsi c'est à partir de ces objets que ce chercheur construit des modèles ethnographiques qui l'amènent à répondre à la question de la « modernité de l'homme » et à substituer au modèle du « big bang culturel » une évolution plus régulière entre -200 000 et -20 000 avant JC⁴.

³ Francesco d'Errico, directeur de recherche au CNRS, Institut de Préhistoire et de Géologie du Quaternaire, CNRS UMR 5199 PACEA

⁴ « Un modèle a longtemps été admis pour rendre compte de cette grande transformation. La modernité de l'homme serait associée à une brusque révolution culturelle ayant eu lieu il y a 40 000 ans environ, soit au début du Paléolithique supérieur. Cette mutation culturelle aurait eu lieu en Europe et coïnciderait avec l'arrivée des hommes anatomiquement modernes, des hommes comme nous en somme, sur le vieux continent. Ce changement a longtemps été considéré comme soudain et explosif. Contre le modèle du big bang culturel du Paléolithique supérieur, un autre modèle a été récemment proposé. Selon ce scénario la modernité culturelle aurait débuté en Afrique, le continent où selon la génétique notre espèce a eu son origine il y a environ 200 000 ans, et se serait déployée par étapes entre 200 000 et 20 000 ans BP (au cours de l'époque dite du Middle Stone Age africain). Ce modèle suppose donc une évolution beaucoup plus graduelle et non européenne. En particulier le deuxième modèle, qui est en train de s'imposer comme le paradigme dominant, lié directement à l'origine biologique de notre espèce à l'origine des comportements modernes. L'idée est simple : le processus qui a produit notre espèce en Afrique a du octroyer à cette dernière certains avantages (langage, pensée symbolique, capacités cognitives supérieures) qui ont favorisé sa colonisation de l'Eurasie et le remplacement par celle-ci des populations humaines vivant dans ces régions » in Conférence de Francesco d'Errico au Cercle Genevois d'Archéologie, 4 avril 2006.

1.2. Des actes fondateurs : les listes, les nombres

1.2.1. L'usage de listes

C'est dans les premières listes que l'on reconnaît à la fois les premières manifestations de désignation d'objets et du concept de nombre. Cela ressemble fort à une naissance d'un couple : l'écriture et les nombres avec comme première conséquence un forme de statistiques correspondant au classement et à la présentation des données.

C'est d'ailleurs ce que fait remarquer Jack Goody (Goody J., 1993, p.167) dans un entretien ayant pour sujet « Ecritures et sociétés » :

« Les études menées récemment sur les débuts de l'écriture indiquent en effet qu'elle s'est développée à Sumer surtout pour des raisons de comptabilité : les marchands utilisaient ce moyen pour envoyer des messages à d'autres marchands. Il s'agissait de brèves indications (de mots isolés ou numéros) portées sur des pots ou des tablettes, de listes de mots apparemment destinées à instruire les élèves dans les écoles. Ces premières écritures ne comportent ni phrase ni paragraphe. »

Evidemment nous savons qu'avec le temps l'écriture va remplir des fonctions fort différentes, passant de la liste de couples (description d'un objet, nombre) à des fonctions plus transcendantes citons par exemple la description du cadastre, algorithmes de calcul d'aires, contrats de commerce, etc. et surtout transcription du discours ou du langage parlé. Ainsi Goody (Goody J., 1993, p.167) poursuit :

« Des écrits plus proches de la parole se développèrent bien entendu sous forme de contrats, de codes, et même de documents historiques et littéraires. Mais les types d'écritures non liées au discours jouèrent un rôle particulièrement important, favorisant le développement des mathématiques et de ses applications (comme par exemple le calcul de l'aire d'un champ). C'est dans la période babylonienne du premier millénaire qu'apparaissent des notations liées au cadastre ainsi que des listes concernant le lever du soleil et de la lune, permettant ainsi l'apparition d'une vraie astronomie prévisionnelle. »

On voit donc qu'en quelques pas nous sommes passés de listes d'objets et de nombres, c'est-à-dire un début des statistiques descriptives à une véritable approche de statistiques prévisionnelles.

Ce double aspect de ces listes, d'une part le codage réservé à ceux qui en connaissent les clés, et d'autre part la prévision d'un futur par essence imprévisible, va installer pour longtemps respect et défiance vis-à-vis du statistique. Comment en effet ne pas être empli de

respect pour celui qui sait à l'avance ce que l'on va trouver dans une jarre ou dans une tablette d'argile et ne pas craindre celui qui par la lecture des tables arrive à prévoir les phases de la lune ou l'arrivée de Sirius permettant de prévoir les crues du Nil (même si ces dernières comportent, on s'en doute, quelques imprécisions) ?

Tout au long de notre travail nous rencontrerons cette facette de la statistique comme une entrave à son évolution, mais également dans la partie didactique de ce travail comme un obstacle au développement d'une pensée statistique chez l'individu.

1.2.2. Une invention de comptables

Georges Ifrah (Ifrah, 1981, p. 160) intitule sous forme interrogative « *Les inventeurs de l'écriture : une invention de comptables ?* » le chapitre 10 de son ouvrage « Histoire universelle des chiffres ». Notons par anticipation que l'on retrouvera sous forme de quolibet le nom d'actuares dans la bouche des bourbakistes pour désigner les statisticiens (INSEE, 1977).

Mais revenons à Georges Ifrah qui poursuit l'opinion contenue dans le titre du chapitre :

« L'idée de l'écrit, en Mésopotamie comme en Elam, serait donc bien née de besoins strictement utilitaires dans un contexte essentiellement économique. » (Ifrah G., 1981, p.178).

Et donc l'écriture en tant que version du langage articulé ne s'est imposée que plus tard, c'est ce que l'auteur renforce quelques lignes après :

« Il ressort, en tout cas très nettement de ce qui précède que la comptabilité écrite a très vraisemblablement précédé, en Elam et sans doute aussi en pays de Sumer, la transcription du langage articulé. » (Ifrah G., 1981, p.178).

Et pour résumer nous pouvons dire que les besoins statistiques et comptables ont entraîné la naissance de l'écrit, et que cet écrit a ensuite rempli les fonctions qu'on lui connaît.

1.3. Une origine de la statistique descriptive

Nul besoin n'est ici de rappeler le « miracle grec » à partir du VII^{ème} siècle avant JC qui permit l'instauration de la logique, de la démocratie et de la méthode scientifique.

Le philosophe Aristote⁵ (Stagire, Macédoine, aujourd'hui Stavro, - 384 - Chalcis, Eubée, -322) est un de ceux qui dans un traité de politique intitulé « Constitution d'Athènes »⁶

⁵ en grec Aristotelês, dit le Stagirite

aborde une forme de statistique descriptive, telle qu'on peut le constater dans les quelques lignes ci-dessous tirées du chapitre XXV⁷ :

« En même temps, suivant la politique inaugurée par Aristide, on assura à la multitude largement sa subsistance. Il arriva que par les contributions extraordinaires, par les droits et impôts, par les alliés, plus de vingt mille hommes étaient nourris. Il y avait en effet six mille juges, mille six cents archers et en outre douze cents cavaliers ; le Conseil comptait cinq cents membres, les gardes des arsenaux étaient au nombre de cinq cents, et les gardes en ville au nombre de cinquante ; environ sept cents hommes exerçaient des magistratures dans le pays; environ autant, en dehors du pays. Plus tard, quand Athènes eut entrepris la guerre, il y eut deux mille cinq cents hoplites, vingt vaisseaux croiseurs, d'autres vaisseaux pour la perception des tributs ayant à bord les deux mille hommes désignés par le sort. Ajoutons le prytanée, les orphelins, les géôliers. Tout ce monde tirait sa subsistance des revenus publics. »

Thomas d'Aquin au XIII^{ème} siècle donne aux idées d'Aristote un statut de doctrine officielle de l'église, cette ligne de l'église romaine a pour conséquence immédiate une influence considérable des idées d'Aristote et ses textes ont inspiré les premières études statistiques en Europe en Occident, jusqu'au XVI^{ème} siècle : en effet à cette période, Galilée, Torricelli et Pascal remettent en cause le système aristotélicien à partir de données expérimentales.

1.4. D'autres approches de la statistique

Si l'on cherche ailleurs que dans une filiation directe on trouve plusieurs exemples montrant l'universalité de la démarche statistique

Le cas chinois : l'empereur chinois Yao, ordonne le recensement des productions agricoles en 2238 avant J.-C. in Gerbaud (Gerbaud, 2006, p. 3) ; l'Egypte : dès 1700 avant J.-C. le pharaon Amasis décrète la peine de mort contre celui qui ne se plie pas aux déclarations de son nom, sa profession et ses moyens de subsistance ; l'exemple indien de

⁶ Le texte en question provient d'un papyrus acheté en Egypte en 1879, une querelle est toujours en cours concernant le fait qu'Aristote soit l'auteur ou non de ce texte, cependant il existe de nombreux témoignages de l'antiquité attestant de la réalité d'un recueil d'un grand nombre de constitutions des cités grecques par les élèves d'Aristote.

⁷ Aristote. Constitution d'Athènes, traduite par B. Haussoullier, Paris : E. Bouillon, 1891, Bib. de l'École des hautes-études. Sciences philologiques et historiques. Fascicule 89

Une approche historique de la statistique

l'Arthashastra : au IV^{ème} siècle avant J.-C. Kautilya, ministre du roi Candragupta (Empire indien des Maurya) rédige un traité de science politique et économique qui décrit les techniques de recensement des populations ; et enfin citons les quipus des Incas qui étaient des systèmes fondés sur des cordelettes, des nœuds et des couleurs permettant d'avoir des statistiques sur les récoltes.(Morlat, 2000, p. 3).

Ces divers exemples brièvement abordés montrent bien l'universalité de la démarche statistique. Cet aspect global de l'approche statistique par les états se retrouve dans un autre élément présent dans les débuts de la statistique à savoir la mesure des sols et des propriétés tels qu'ils apparaissent dans la constitution des divers cadastres, c'est ce que nous allons aborder maintenant.

2. Le cadastre

2.1. Liminaire lapidaire

L'établissement du cadastre⁸ est dans un premier temps une des pierres sur laquelle nous avons construit la statistique et ses outils. Sans vouloir en faire une étude historique exhaustive donnons quelques points de repères tels que les présente Jean-Paul Miserez⁹, cette chronologie étant reprise par l'Institut de la gestion publique et du développement économique (IGPDE). Nous distinguerons trois périodes : une première époque couvrant l'Antiquité et le Moyen Âge, puis de la Renaissance à la Révolution française et enfin de cette dernière à la fin du XIX^{ème}.

2.2. Les périodes antique et médiévale

J.P. Miserez nous donne un aperçu des pratiques concernant le cadastre et le cens depuis l'Antiquité jusqu'au Moyen Âge. Nous avons résumé dans le tableau suivant :

⁸ Les données de ce paragraphe proviennent des documents de l'Institut de la gestion publique et du développement économique (IGPDE)

⁹ ingénieur-géomètre, chargé de cours à l'École polytechnique de Lausanne

Période	Lieu	Souverain	Evènement / découverte
Env. 4000 av JC	Chaldée		Tablette donnant le plan coté, la superficie, la description d'un groupe de parcelles de la ville de Dunghi.
3200-2800 av JC	Égypte		Recensement foncier en en vue de la perception de l'impôt, basé sur la surface des propriétés et leur rendement
2238 av JC	Chine	Empereur Yao	Recensement des productions agricoles.
2000 av JC	Égypte	Sésostris	Les terres sont divisées en parcelles qui déterminent l'assiette de l'impôt foncier.
1600-1400 av JC	Italie		Plan gravé sur un rocher plat. Les lignes figurent les ruisseaux, les canaux d'irrigation et les routes.
Env. 1200 av JC.	Israël	Josué	Organisation territoriale et tribale d'Israël. Etablissement d'un cadastre et le partage des sols entre les tribus (Josué, XVIII, 4-9).
Env. 700 av JC	Lucanie		Cadastre sur 10000 hectares divisés en lots rectangulaires de 6 hectares chacun.
Vers 578-535 av JC	Etrurie	Servius Tullius	Cadastre de Rome.
63 av JC - 14 ap. JC	Gaule	Auguste	Etablissement du cadastre.
77 après JC	Gaule		Gravure du cadastre d'Orange.
380 après JC	Rome	Théodose I le Grand	Registre public : contenance, nature et qualité des biens par déclaration des propriétaires.
645-649	Japon	Ere de Taïka	Confection de registres d'état-civil et du cadastre.
Env. 1000	Inde	Raja le Grand	Confection d'un cadastre
1086	Angleterre	Guillaume le Conquérant	Fin de l'établissement du Domesday Book
1303 et 1493	France		. La royauté tente « à plusieurs reprises de faire dresser, çà et là, un "cadastre, compoix ou estime", local d'abord, puis régional et enfin général, qui servît d'assiette équitable des subsides demandés au royaume ¹⁰ . »
1364-1380	France	Charles V	Organisation des « parcellaires »
1368-1398	Chine	Hongwu	Recensement de la population et cadastre général

On le voit dans cette énumération le cadastre, le cens sont les outils qui permettent au pouvoir de l'état de dégager les ressources nécessaires à son fonctionnement.

2.3. De la Renaissance à la Révolution française

La maîtrise et la connaissance des outils scientifiques se propagent durant cette période traversée par de grands changements et on essaie d'installer des réformes

¹⁰ DUPONT-FERRIER (G.). – Études sur les institutions financières de la France. – Paris, *Librerie Firmin-Didot*, 1930-1932, 2 vol., I, p. 10 et II, p. 48, 136, 140, 141 et 246. [BH : 8° 14207.]

Période	Lieu	Souverain	Evènement / découverte
1427	Florence		Cadastre
1467	Tabriz (Iran)	Uzun Hasan	Cadastre agraire
A partir de 1599	Vietnam du Nord	Trinh- Cuong	Elaboration d'un cadastre

Les plans cadastraux se succèdent alors, on a ainsi et parmi d'autres : 1664.Montauban (sous l'impulsion du ministère Colbert), 1670 Lausanne, 1713-1740 Prusse (Frédéric-Guillaume Ier), 1716-1718 Catalogne (Philippe V), 1730 Savoie (Charles-Emmanuel III), 1740 Bâle (prince évêque), 1781-1786 Sicile (le vice-roi Caracciolo n'arrive pas à établir le cadastre).

2.4. De la Révolution française à la fin du XIX^{ème}

Cette période nous est plus proche et nous en trouvons encore les traces dans les mairies des villes et des villages¹¹.

Beaucoup de débats ont eu lieu pendant la révolution française : fallait-il un cadastre général ? Charles Gomel (GOMEL, 1896, p.130-131) rapporte que « (...) *par suite d'illusions qui parfois se rencontrent encore aujourd'hui dans les discussions des Chambres* [il écrit en 1896 sur 1789], *plusieurs cahiers exprimèrent l'espoir qu'un cadastre général pourrait servir de base à la répartition de l'impôt foncier entre les différentes provinces. L'expérience a prouvé depuis que le cadastre parcellaire n'est utile que pour la répartition entre propriétaires d'une même commune.* »

Jean Pigeire rapporte qu'un peu plus tard en nivôse an XI (décembre 1800) Chaptal¹², ministre de l'intérieur, envoie « *les instructions successives pour l'établissement d'un cadastre national, garantie des cultivateurs, base des jugements et des contributions* », (PIGEIRE, 1932).

En 1807, Napoléon décrète une loi sur le cadastre : ce dernier devait être un instrument fiscal et administratif, mais aussi, fait nouveau, juridique.

¹¹ Il existe des mairies où dans certains cas on ressort le « cadastre napoléon »

¹² Ministre de l'intérieur de 6 novembre 1800 au 4 août 1804

Il fallut presque cent ans pour qu'entre 1891 et 1898, les chambres décident le renouvellement du cadastre à la suite de l'exposé du budget pour 1892¹³, afin qu'il serve aussi d'instrument juridique.

Durant ce siècle le cadastre avait été généralisé dans tous les états modernes, de la Prusse à l'Italie et de la Belgique à l'Australie (Act Torrens 1858).

2.5. En guise de conclusion de cet aperçu de l'histoire du cadastre

De cette rapide histoire du cadastre nous tirerons quelques conclusions se rapportant à notre travail :

- Le cadastre a été en premier lieu un instrument fiscal et/ou administratif
- Les perfectionnements scientifiques ont permis une plus grande précision et une fiabilité accrue des cadastres
- Au fil des siècles les sociétés en ont aussi fait un instrument juridique

La statistique fonctionne de manière homéomorphe au cadastre : les ingrédients sont différents mais sa structuration au fil des siècles lui a fait également changer de statut.

En effet elle est passée d'un instrument essentiellement destiné à dénombrer les richesses, les sujets et à prélever des impôts au statut d'instrument prévisionnel, on peut dire qu'elle a outrepassé les statistiques pour atteindre à la statistique.

On peut penser sur la base de certaines pratiques que ces outils, qui appartiennent assurément à la statistique descriptive, ont eu une certaine fonction prévisionnelle et annoncent, de ce fait, une statistique comportant une conjecture sur le futur une espèce de statistique inférentielle « en acte ».

¹³ ROUVIER (Maurice), ministre des Finances. – Discours prononcé par M. Maurice Rouvier, ... séance du 24 octobre 1891. Discussion du projet de loi portant fixation du budget général de l'exercice 1892. – *Paris, impr. des journaux officiels*, 1891. « Le cadastre perpétué à l'aide d'un système permanent de conservation, ne serait pas seulement un instrument fiscal et administratif : il devrait satisfaire à d'autres besoins. Des abornements généraux et une triangulation rigoureuse précéderaient le renouvellement des opérations : le cadastre constituerait la base de la propriété foncière ; il assurerait la sécurité des hypothèques et la régularité des transactions immobilières ; il fournirait enfin à l'agriculture, par le développement des institutions de crédit, les moyens d'action qui lui font défaut aujourd'hui. En un mot, il deviendrait le Grand Livre terrier de la France ; » (§ 26 et 27.) [BnF : 8° Le⁹⁰. 957.]

3. De la théocratie

Il y a une certaine légitimité à supposer que la prise de conscience par l'homme de son humanité soit arrivée très tôt, et qu'elle précède puis coïncide avec des productions symboliques. Parmi toutes les activités symboliques possibles, celles dont on a trouvé les traces les plus anciennes consistent en l'organisation de sépultures. On conviendra qu'un ensemble d'ossements humains fossilisés n'est pas forcément une sépulture, et il semble que le meilleur critère soit « la préservation du squelette en connexion » (VANDERMEERSCH, 1986), ce qui, au passage, fait appel à une preuve d'ordre statistique.

La datation des plus anciennes sépultures découvertes est évaluée à 100 000 ans. Il s'agit du site de Skhül en Israël fouillé depuis 1930 par Dorothy Garrod (1892-1968, professeur à l'Université de Cambridge), puis par d'autres chercheurs. Cette prise de conscience de sa condition par l'homme l'a inmanquablement conduit au religieux et de là, aux religions. Il était alors naturel que s'installent des théocraties dans divers groupes sociaux, certaines étant d'actualité. Leurs positions par rapport à la science et donc à la statistique sont nombreuses, diverses et variées, parfois en totale opposition et souvent ambiguës.

Dans des hypothèses selon lesquelles une entité supérieure domine le futur il est évidemment sacrilège de faire des calculs sur ce futur. Ainsi dans la Bible selon que le recensement est l'ordre de Dieu ou du roi, il connaît des fortunes opposées. Dans le Premier livre des chroniques (Bible, p. 1288) on a une statistique des enfants d'Israël ordonnée par Dieu :

« Tous les Israélites avaient été enregistrés et voilà qu'ils étaient inscrits sur le Livre des rois d'Israël.

Quant aux Judéens, ils avaient été déportés à Babel à cause de leur infidélité.

Les premiers habitants qui réoccupèrent leurs propriétés et leurs villes furent ceux d'Israël : les prêtres, les Lévites, les oblats.

À Jérusalem habitèrent des fils de Juda, des fils de Benjamin, des fils d'Ephraïm et de Manassé. ».

Et il semble que cet inventaire ne pose aucun problème.

Par contre, toujours dans le Premier livre des chroniques chapitre XXI (Bible, p.1325) lorsque le roi David prend l'initiative de faire de même pour son propre compte pourrait-on dire :

« Satan se dressa contre Israël et il incita David à dénombrer Israël. David dit à Joab et aux chefs du peuple : « Allez compter Israël, de Bersabée à Dan, puis revenez m'en faire connaître le chiffre. Joab répondit : « Que Yavé accroisse son peuple de cent fois autant ! Monseigneur le roi, ne sont-ils pas tous les serviteurs de Monseigneur ? Pourquoi Monseigneur fait-il cette enquête ? Pourquoi Israël deviendrait-il coupable ? » Cependant l'ordre du roi s'imposa à Joab. Joab partit et parcourut tout Israël puis rentra à Jérusalem. [...] L'ordre du roi avait tant répugné à Joab qu'il n'avait recensé ni Lévy ni Benjamin. »

La Bible nous raconte que Dieu se fâche et demande à David de choisir entre plusieurs punitions toutes plus abominables les unes que les autres et frappant tout le peuple :

« Dieu vit avec déplaisir cette affaire et il frappa Israël. David dit alors à Dieu : « C'est un grand péché que j'ai commis en cette affaire ! Maintenant veuille pardonner à ton serviteur, car j'ai commis une grande folie » [...] Yavé envoya donc la peste en Israël... À la suite de quoi David érigea le fameux temple à Yavé. »

On le voit à cet exemple, une pratique statistique, même simple, est un domaine où les hommes doivent s'aventurer avec prudence.

Si l'on fait un énorme saut dans l'espace et dans le temps on rapprochera l'épisode précédent à celui de Staline déportant les statisticiens qui avaient fait des sondages, seuls les recensements ayant un statut de preuve à ses yeux. Lors d'un « Café de la Statistique » (toujours à propos de Staline) ayant pour thème « *L'indépendance de la statistique à l'égard du pouvoir politique* » organisé par la SFDS (Société Française de Statistique) le 16 mai 2006 et animé par Jean-Marie Delarue¹⁴, un des participants raconte que « *Staline avait anticipé le résultat du recensement démographique ; l'office statistique a trouvé un chiffre inférieur ; son directeur n'a pas d'emblée été sanctionné, mais on a nommé une commission d'enquête ; comme il advient pour tout recensement, elle a trouvé des omissions (et, sans doute aussi, des doubles comptes) : on a conclu au sabotage et le directeur a été fusillé. Puis, le président de la commission d'enquête a été nommé à sa place ; il a alors mieux compris les conditions de l'opération et en est venu à partager l'avis de ses collaborateurs : il a été fusillé à son tour* ».

¹⁴ ancien vice-président du Conseil national de l'information statistique

Nathalie Moine (MOINE, 2002) a développé dans un article « Le miroir des statistiques, Inégalités et sphère privée au cours du second stalinisme » une étude fort complète des rapports pour le moins complexes entre les statistiques et le pouvoir dans la société soviétique sous Staline.

On pourrait penser qu'une société laïque, débarrassée en partie du poids de la religion, puisse plus facilement accéder à une pensée statistique mais nous allons voir au paragraphe suivant qu'il reste encore quelques obstacles.

4. Calculer le futur se pose comme un obstacle épistémologique

4.1. Une pensée débarrassée des superstitions ?

Dans l'ouvrage *Psychopathologie de la vie quotidienne* (Freud, 1969), l'auteur nous fait rencontrer la superstition comme une confirmation de « *l'existence d'une connaissance inconsciente et refoulée de la motivation des actes manqués et accidentels* ».

A partir de là on peut émettre l'hypothèse que les superstitions seront toujours constituantes de l'appareil psychique de l'individu, tout au plus faudra-t-il apprendre à vivre avec, et l'on ne s'étonnera pas que tel président ait consulté régulièrement une célèbre et médiatique astrologue, que le suivant ait vu, dit on, des marabouts africains, et que les jeux de la « Française des jeux » aient autant de succès.

Et donc, en ce qui concerne les apprentissages statistiques ces derniers ne seront, dans un premier temps, qu'un cosmétique sur les superstitions qu'à des degrés divers chaque individu garde en lui. C'est à force de persévérance et par la multiplication des situations qu'une pensée statistique s'installera à côté de « petites superstitions » fonctionnant comme des addictions marqueurs de névroses plus ou moins handicapantes.

En revanche, pour le même individu, ils viendront compléter l'outillage logique et l'« habitus », (BOURDIEU, 1980, pp. 87), tout en portant un certain nombre de singularités et de contradictions que nous abordons ci après.

4.2. Un paradoxe par rapport à la pensée d'Aristote : prendre et rejeter à la fois

Mais revenons au développement d'une pensée ayant pour sujet le statistique. La méthode scientifique se nourrit au sein de la pensée d'Aristote et c'est également vrai de la statistique.

Nous avons déjà rencontré les rapports d'Aristote et de la pensée officielle chrétienne par l'intermédiaire de Thomas d'Aquin (1225-1274). Ce dernier trouve dans les écrits d'Aristote un système qu'il pense le meilleur pour développer une doctrine de l'église. Dans ce renouveau de la scolastique il est en conflit avec les partisans d'un courant traditionnel, en

particulier les franciscains, qui s'appuient sur la glose d'Augustin (354-430) inspirée du système néo platonicien. Notons qu'Averroès¹⁵ se réclame aussi d'Aristote et qu'il est sans doute un des principaux enseignants de sa pensée à l'Occident.

Mais ce qui a permis le développement de la pensée et de la méthode scientifique devient un obstacle pour développer une pensée statistique, ainsi lorsqu'on lit chez Aristote (Aristote, *Derniers analytiques*) dans le début du chapitre VIII des « *Derniers analytiques* » :

*« Toute conclusion démontrée est éternelle : il n'y a donc pas de démonstration pour les choses périssables, de même qu'il n'y a pour elles que science d'accident. - Les définitions sont éternelles comme les démonstrations, dont elles ne sont qu'une forme. - La démonstration peut s'appliquer à certaines choses passagères, mais dont l'essence est éternelle, par exemple certains phénomènes naturels. »*¹⁶

On retient qu'« il n'y a pas de démonstration pour les choses périssables ».

Et dans le paragraphe 10 de ce chapitre Aristote rajoute :

§ 10. Τῶν δὲ συμβεβηκότων μὴ καθ' αὐτά, ὃν τρόπον διωρίσθη τὰ καθ' αὐτά, οὐκ ἔστιν ἐπιστήμη ἀποδεικτική. οὐ γὰρ ἔστιν ἐξ ἀνάγκης δεῖξαι τὸ συμπέρασμα· τὸ συμβεβηκὸς γὰρ ἐνδέχεται μὴ ὑπάρχειν· περὶ τοῦ τοιοῦτου γὰρ λέγω συμβεβηκότος.

Ce qui peut se traduire par :

*« § 10. Il est impossible de savoir par démonstration les accidents qui ne sont pas essentiels dans le sens même de la définition que nous avons donnée de ce mot : c'est qu'en effet on ne peut jamais pour les accidents démontrer que la conclusion est nécessaire, puisqu'un accident est ce qui peut ne pas être, seule espèce d'accident dont je veuille ici parler. »*¹⁷

Ceux qui voulaient développer une pensée statistique étaient obligés de faire un grand écart : d'une part accepter la partie du système qui leur permettait d'avancer (la logique, etc.), mais d'autre part transgresser le système mis en place par Aristote lui-même.

¹⁵ Averroès (1126-1198) a commenté pendant une grande partie de sa vie les écrits d'Aristote

¹⁶ Traduction de Jules Barthélemy-Saint-Hilaire

¹⁷ [ibidem](#)

4.3. Franchir l'obstacle épistémologique

Les difficultés précédentes montrent bien qu'il existe un obstacle à la construction de la pensée statistique ; en ce sens nous suspectons ici un obstacle de nature épistémologique et nous devons sans doute ne pas le sous-estimer dans le cadre des apprentissages statistiques.

Devant les affirmations contenues dans les « Derniers analytiques », qui plus est appuyées par les thuriféraires d'une doxa thomiste, les contradicteurs ont dû :

- temporiser afin d'attendre que les forces ennemies soient moins puissantes ;
- attendre que certaines théories d'Aristote soient contredites par l'expérience ;
- patienter jusqu'à ce que la doctrine de l'église s'émancipe par rapport aux théories aristotéliennes ;
- trouver un terrain d'exercice, autrement dit un « angle d'attaque », ce sera essentiellement le jeu ;
- construire des outils appropriés.

Et on le constate, les premiers développements vont se dérouler dans le cadre des jeux de hasard (interdits puis tolérés par l'église), ces situations étant investies à l'aide d'outils construits pour la circonstance : la combinatoire et le calcul des probabilités naissant.

5. Aleae Geometria

5.1. Introduction

C'est de ce titre inouï, donné par Pascal (Pascal, 1954, p.75) dans le texte adressé en 1654 à l'académie parisienne des sciences «CELEBERRIMÆ MATHESIOS ACADEMIÆ PARISIENSI» ou encore «A LA TRES ILLUSTRÉ ACADEMIE PARISIENNE DE SCIENCE» faisant référence à ce qui va devenir le calcul des probabilités que nous engagerons cette partie de notre étude. En effet si on ne peut difficilement parler de statistique sans parler de calcul de probabilités, l'inverse devrait être également vrai.

5.2. La géométrie du hasard

Dans ce texte, Pascal transmet plusieurs «travaux» différents et il introduit et décrit ainsi la «Géométrie du hasard» :

« Et puis un traité tout à fait nouveau d'une matière absolument inexplorée jusqu'ici, savoir : la répartition du hasard dans les jeux qui lui sont soumis, ce qu'on appelle en Français faire les partis des jeux »

En fait ce que l'on présente quelquefois comme une découverte est un sujet à la mode à l'époque. Il s'agit de savoir comment répartir les gains en cas d'arrêt non prévu d'un jeu.

D'une manière indépendante Fermat (1601-1665), magistrat à Toulouse qui travaille la théorie des nombres, l'optique, le calcul infinitésimal, la géométrie analytique et le calcul des probabilités a trouvé une solution au «problème des partis» et il pense que quelqu'un pourrait s'occuper à publier l'ensemble de ses travaux et il pense à Carcavi, à Huygens et à Pascal (1623-1662) déjà connu pour ses travaux en mécanique et en mathématiques (en fait aucune des trois personnes ne fit ce travail et c'est le fils aîné de Fermat, Samuel, qui le réalisa après la mort de son père).

Le contact entre Fermat et Pascal a lieu en 1654 par le biais d'une correspondance dont nous avons heureusement une trace au moins partielle.

Privilège nous est donné d'avoir d'une part deux solutions différentes développées l'une par Pascal et l'autre par Fermat, et d'autre part une partie de la correspondance (certaines lettres sont perdues) entre les deux brillants esprits. On trouve une analyse de cette correspondance sous la plume de Nicolas Trotignon (TROTIGNON, 2006). Pascal va publier

une solution dans la « Règle des partis » (PASCAL, 1954, p.115), Fermat se contentera d'essais restés manuscrits de son vivant.

Mais revenons à cette correspondance intéressante à plus d'un titre.

En effet elle nous donne à voir un archétype, celui d'un échange entre deux façons d'envisager le calcul des probabilités. Les solutions avancées par Fermat qui utilise l'analyse combinatoire et le principe des probabilités composées semblent plus magistrales et plus fécondes que la méthode énoncée par Pascal.

Elle peut également alimenter notre réflexion sur la « validation entre pairs »¹⁸, bien que leurs approches soient différentes ; mais plus fondamentalement elle sonne la fin d'une croyance en un futur contingent relevant de la simple supposition, En effet sur ce dernier point Pascal et Fermat donnent au calcul la capacité de pouvoir renseigner sur le futur contingent : l'idée nouvelle est de partir de l'ensemble des dénouements envisageables et par une pensée rétroactive, donner en remontant au présent, in fine, la connaissance d'un futur « calculé ».

Mais si Fermat reste un magistrat de province pour qui les mathématiques sont un passe temps, Pascal essaie de « faire carrière » et pour cela fréquente les salons parisiens où gravitent toutes sortes d'idées et de gens. Parmi ceux-ci, le chevalier de Méré qui se demande à partir de combien de « coups » on peut parier, en jouant avec deux dés, avoir un double six, et avoir une espérance supérieure à $\frac{1}{2}$. Pascal donne une solution et trouve là une certaine notoriété.

Soulignons, in fine, une petite note que Pascal (PASCAL, 1665, p.219) rajoute à une lettre du 29 août 1654 de Fermat à lui-même (publiée après la mort de Pascal en 1665) « *il écrivait dans sa province ce que j'inventais à Paris, heure pour heure, comme nos lettres écrites et reçues en même temps en témoignent.* ». Or la note (de Pascal) est complétée par un renseignement concernant un travail de Pierre Boutroux qui « *a établi [...] dans l'édition des Grands Ecrivains de la France (t. III, p.334) [que] les découvertes de Fermat et de Pascal ne furent pas simultanées. Fermat avait dix-huit ans de priorité.* »

Nul doute que l'insistance de Pascal ne soit, au minimum, le reflet d'un doute sur la simultanéité de leurs découvertes.

¹⁸ idée reprise par un certain nombre de didacticiens des mathématiques tel Nicolas Balacheff écrivant : « La nécessité de prouver est liée à la situation dans laquelle on se trouve; la preuve est un acte social, elle s'adresse à un individu (éventuellement soi-même) qu'il faut convaincre. » in Encyclopaedia Universalis.

5.3. Cardano et Ludo Aleae (1663) ; Galilée et le « Problème du Duc de Toscane » (1620)

Notons au passage que si l'acte « officiel » de naissance du calcul de probabilité est, tel que nous l'avons vu ci-dessus, attribué à Fermat et Pascal, on donne à Cardano et à Galilée les prémisses de cette naissance.

Ainsi Cardano (CARDANO, 1663, p. 273) développe au chapitre XXVIII dans « De ludo Aleae Liber », écrit aux environs de 1564 mais publié en 1663 (HALD, 2003), une série de textes concernant divers jeux.

CAPVT XXVIII.
De longo consilio, & iudicio,
procedentiaque.

MAximum autem in ludo Aleæ præcipuè est prouidentia, & iudicium ad multos iactus, vt in alueare tabulæ sic disponantur, vt cum duplex sit via ad finem, vna, vt celerius progrediamur, altèrâ, vt collusorem retardemus: celerius progredimur occupando meliora loca, ac citius, vt contra retardamus, vel vt deteriora loca habeat, & pauciora, vel vt lentius ingrediatur vltimam fidem. Cum ergo dictæ viæ sint vtraque

Ce qui peut se traduire par :

« Plans de grande envergure, Jugement, et Procédé

DE LA PLUS GRANDE IMPORTANCE dans le jeu sont la prévoyance et le jugement concernant les nombreux jets, car le jeu de plateau est tel qu'il y a deux directions à notre but, l'une qui est d'avancer le plus rapidement, l'autre qui est de retarder l'adversaire ; nous avançons plus rapidement en occupant de meilleures positions, et plus rapidement, si au

contraire, nous retardons notre adversaire de manière à ce qu'il ait de plus mauvaises positions ou moins de choix ou qu'il aille en arrière en repartant de la dernière rangée. »¹⁹

Quant à Galilée il traite d'un problème (GALILEO GALILEI , 1718) connu sous le nom « Problème du Duc de Toscane ». Galilée met en scène un haut personnage lui demandant d'« expliquer » la différence lors d'un lancer de trois dés entre la fréquence obtenue d'une somme égale à 9 et celle d'une somme égale à 10, alors que ces nombres 9 et 10 sont obtenus « d'autant de façons » (sous entendu non ordonnées soit 6 façons). Galilée affirme que c'est une « lunga osservazione » qui aurait permis de constater une différence entre les fréquences mais nous savons que les probabilités ne diffèrent que de 1/108 ce qui rend difficile un tel constat. Galilée fournit une réponse tout à fait satisfaisante à la question initiale dans des pages publiées seulement en 1718.

Remarquons bien que dans les deux cas soit de Galilée et du Duc de Toscane, soit de Pascal et du chevalier de Méré le calcul des probabilités est convoqué en réponse à une question issue de la statistique, mais que les statistiques sont elles même fournies comme attestant les résultats trouvés par la théorie des probabilités.

¹⁹ la traduction de Sydney Henry Gould « Cardano, The Gambling Scholar » de Ore Oystein (1953)

6. Les maîtres de l'erreur

Stéphane Callens (CALLENS, 1997) rapporte le mot (péjoratif) d'Auguste Comte à propos des statisticiens comme étant les « maîtres de l'erreur ». Cette maîtrise de l'erreur a commencé bien avant cet épisode, et c'est dans la ligne des travaux de De Moivre (DE MOIVRE, 1733) et de Bernouilli que l'analyse s'adapte à la mesure des erreurs (essentiellement des erreurs de mesure en astronomie). La fin du XVIII^{ème} siècle et le XIX^{ème} sont riches d'évènements où statistique et probabilités sont mêlés : Lagrange par exemple conçoit la méthode des moindres carrés (paternité contestée par Gauss) et Laplace (LAPLACE, P.S., 1810) obtient les succès que l'on sait. Cet écart important dans le temps entre De Moivre et Laplace, souligné par Bernard Ycart²⁰, recouvre une transformation du rôle de la science entre le XVIII^{ème} siècle et le XIX^{ème}. Ainsi dans la préface à *l'Essai philosophique sur les probabilités*, de Laplace, R. Thom écrit en parlant du texte de Laplace :

« C'est un texte révélateur d'une transition essentielle, la charnière qui relie l'humanisme éclairé du XVIII^{ème} siècle au scientisme autoritaire et dogmatique qui dominera à la fin du XIX^{ème} siècle - et qui domine encore une bonne part de la société contemporaine... Le XVIII^{ème} siècle fut celui des mathématiques de l'intelligibilité, le XIX^{ème} celui des mathématiques du contrôle. »

Mais la statistique se veut également instrument social de Condorcet à plus tard Cournot ; Quételet va établir des premières bases de données, par ailleurs la statistique mathématique anglaise va prolonger les travaux de Darwin concernant la biologie, Galton s'appuyant sur la variabilité et constatant la « régression » des tailles vers la moyenne. Puis les tests, déjà en germe dans les travaux de Laplace vont faire une entrée en force dans le XX^{ème} avec le test du χ^2 en 1900, suivi par les statistiques inférentielles (Fischer, Student, Pearson Karl et Egon, Neymann) en ce qui concerne les tests d'hypothèses et les intervalles de confiance, et complété par les travaux de Wald (théorie des décisions statistiques) et de Shewhart (cartes de contrôle). Malgré les travaux de Borel (1924), Lévy et quelques autres, la

²⁰ B. Ycart http://www-lmc.imag.fr/lmc-sms/Bernard.Ycart/smel/articles/etoiles/cadre_etoiles.html

Une approche historique de la statistique

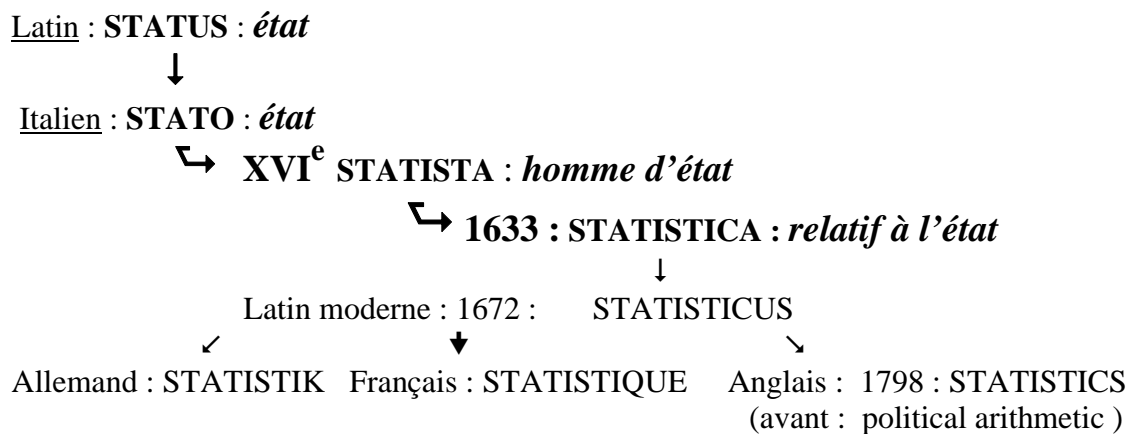
France est en retard sur ce plan, et ce retard va s'accroître avec l'attitude du groupe Bourbaki très anti statisticiens qu'ils traiteront d'actuaux.

C'est l'axiomatique de Kolmogorov qui donne une unité à l'ensemble en 1933, la fin du XX^{ème} voyant fleurir de nouvelles méthodes d'analyse des données : l'analyse en composantes principales, l'analyse factorielle des correspondances qui ne peuvent s'épanouir que grâce aux progrès considérables faits par les instruments de calcul.

7. Conclusion

On le voit nous avons donné dans cette partie un aperçu sur l'histoire de la statistique et du calcul des probabilités. Cette césure historique se retrouve dans l'étymologie des mots que l'on utilise dans le champ étudié.

Ainsi si l'on attribue souvent l'introduction du terme « statistique » à un professeur de Göttingen, G. Achenwall, qui aurait en 1746 créé le mot Statistik, dérivé de la notion Staatskunde, l'étymologie nous donne une histoire un peu plus complexe (Le ROBERT, 1993, p.2016) dont voici un résumé :



Le nom signifie : « étude méthodique par des procédés numériques des faits sociaux qui définissent un état. »

Le nom change de valeur, d'abord en Angleterre, en devenant « statistics » (1798), puis en France en 1832, et il prend le sens de : « ensemble de techniques d'interprétation mathématique appliquées à des phénomènes. »

Le nom désigne ensuite (1862) l'objet des statistiques : « ensemble de données numériques concernant une même catégorie de faits »

Alors que l'origine du mot hasard (Le ROBERT, 1993, p.946) est la suivante :

Arabe : **ZAHR** = *fleur*



AZ-ZAHR = *jeu de dés* parce que les dés portaient une fleur sur l'une des faces



Espagnol : **AZAR** = *jeu de dés, coup défavorable au jeu de dés*



Français : **HASART** (vers 1150)

↳ **HASARD** (1200) = *un jeu de dés*



coup heureux à ce jeu (6)

(« jeu de hasard » : 1538)



XIII^e Sens figuré : mauvais coup



XV^e : risque, danger

(seule acception qui reste aujourd'hui : « les hasards de la guerre »)



Début XVI^e : cas, évènement fortuit

Milieu XVI^e : cause qu'on attribue à ce qui arrive sans raison apparente (cf en sciences « les lois du hasard »)

Si la preuve est du côté de la probabilité, la collecte et le traitement des données sont bien recouverts par le mot statistique et cela montre la nécessité d'avoir deux cordes (au moins) dans l'apprentissage de la statistique, l'une utilisant la pratique d'enquête par questionnaires participe au sens de l'apprentissage, l'autre à travers la simulation permet de multiplier le nombre d'essais réalisant en cela les affirmations de Galilée et Pascal concernant la lettre au Duc de Toscane pour l'un et la question du Chevalier de Méré pour l'autre.

Table des matières

UNE APPROCHE HISTORIQUE DE LA STATISTIQUE : DES PREMIERS PAS A LA	
« GEOMETRIE DU HASARD » ET AUX « MAITRES DE L'ERREUR ».....	1
1. DES DEBUTS PROMETTEURS	3
1.1. Les systèmes artificiels à mémoire (SAM).....	3
1.2. Des actes fondateurs : les listes, les nombres.....	4
1.2.1. L'usage de listes.....	4
1.2.2. Une invention de comptables.....	5
1.3. Une origine de la statistique descriptive.....	5
1.4. D'autres approches de la statistique.....	6
2. LE CADASTRE	8
2.1. Liminaire lapidaire.....	8
2.2. Les périodes antique et médiévale.....	8
2.3. De la Renaissance à la Révolution française.....	9
2.4. De la Révolution française à la fin du XIX ^{ème}	10
2.5. En guise de conclusion de cet aperçu de l'histoire du cadastre.....	11
3. DE LA THEOCRATIE.....	12
4. CALCULER LE FUTUR SE POSE COMME UN OBSTACLE EPISTEMOLOGIQUE... 15	15
4.1. Une pensée débarrassée des superstitions ?.....	15
4.2. Un paradoxe par rapport à la pensée d'Aristote : prendre et rejeter à la fois.....	15
4.3. Franchir l'obstacle épistémologique.....	17
5. ALEAE GEOMETRIA.....	18
5.1. Introduction.....	18
5.2. La géométrie du hasard.....	18
5.3. Cardano et Ludo Aleae (1663) ; Galilée et le « Problème du Duc de Toscane » (1620)...	20
6. LES MAITRES DE L'ERREUR	22
7. CONCLUSION.....	24