Journ@l Electronique d'Histoire des Probabilités et de la Statistique

Electronic Journ@l for History of Probability and Statistics

Vol 6, n°1; Juin/June 2010

www.jehps.net

Statut de la Dispersion : de l'erreur à la variabilité

Michel ARMATTE[1]

Abstract

The concept of dispersion, ubiquitous in the history of statistics, is the object not only of different measures but also different meanings. An important moment is when one moves from the observations of the same object (in the theory of errors) to the statistical analysis of different individuals of a certain population. By allowing a confusion of mean and average in these two cases, which permitted him to establish his social physics, Quetelet does not forget that the error, becoming variation, changes in nature and magnitude, and that this variation has spatial and temporal components. The German School of variability of chances illustrated by the work of Lexis and Bortkiewicz, while rejecting the overly simplistic model of urn with constant composition, will develop this intuition of a decomposition of variability in two dimensions: consistency and stability.

Résumé

La notion de dispersion, partout présente dans l'histoire de la statistique, est l'objet non seulement de mesures différentes mais surtout de significations différentes. Un moment important est celui où l'on passe d'observations d'un même objet (en théorie des erreurs) à des relevés statistiques sur des individus différents d'une certaine population. En autorisant une confusion des moyennes objectives et subjectives dans ces deux cas, ce qui lui permet de fonder sa physique sociale, Quetelet n'oublie pas que les erreurs changent de nature et d'ordre de grandeur en devenant des variations, et que cette variation a des composantes spatiales et temporelles. L'école allemande de la variabilité des chances illustrée par les travaux de Lexis et Bortkiewicz, tout en récusant le modèle trop simple de l'urne à composition constante, développera cette intuition d'une décomposition de la variabilité en deux dimensions : homogénéité et stabilité.

C'est à l'école biométrique anglaise, celle de Francis Galton, Karl Pearson et surtout Ronald Fisher, père de la statistique mathématique, que l'on attribue le plus souvent l'irruption d'une mesure universelle de la variabilité statistique sous la terminologie communément adoptée aujourd'hui de variance et d'écart-type. Et la notion de dispersion n'apparaît guère dans les traités de statistique avant 1900[2]. On voudrait ici non seulement faire référence à la préhistoire de ces notions au XIXe siècle, mais évoquer un moment majeur de la réflexion sur la sémantique qui se cache derrière l'idée de variabilité. Ce moment est celui du glissement de la notion de dispersion des mesures dans l'observation des phénomènes

1 michel.armatte@dauphine.fr

2 C'est ainsi par exemple que le chapitre « dispersion and skewness » du traité de William Bowley que l'on trouve dans l'édition de 1920 n'est pas dans la première édition de 1902.

naturels – en astronomie par exemple – vers celle de variation en statistique sociale ou économique. Ce passage d'une science de la nature à celle de l'homme que le marin et historien des sciences Michel Serres avait nommé le passage du Nord-Ouest[3] désignait une possibilité longtemps fermée aux trafics de tout genre entre ces deux continents du savoir. Si l'ouverture de ce passage marin pourrait bien avoir un effet accélérateur sur la mondialisation du commerce, l'ouverture d'un passage entre la théorie des erreurs et la statistique au milieu du XIX^e^ siècle a permis un large développement des sciences de l'homme. Mais dans les deux cas l'ouverture nous a fait – et nous fera - oublier les différences initiales. Le commerce mondial uniformise les modes de consommation, et l'écart-type a uniformisé les significations différentes de la variabilité en astronomie et en sociologie. C'est ce paysage différencié que nous voulons restituer.

Dispersion et déviation dans la théorie des erreurs

Lorsque Johann-Heinrich Lambert (1765) introduit le terme de théorie des erreurs comme « étude des relations entre les erreurs, leurs conséquences, les circonstances de la prise de mesure et les qualités des instruments » il savait que la question était débattue depuis les débuts de la science astronomique et qu'elle empoisonnait la vie des astronomes : comment construire en effet une science universelle, comment lire le Livre de la Nature, alors que trois observateurs différents d'une même grandeur liée à un phénomène astronomique, voire trois observations faites par le même observateur, en rapportaient des mesures différentes. Il était si évident pour eux que le phénomène n'avait qu'une seule réalité et la grandeur une seule « vraie valeur », que la *dispersion* des observations ne pouvait se comprendre que comme l'effet d'une *erreur* de mesure, une erreur constituant un véritable poil à gratter qui empêchait de construire *la* science de l'univers. Lambert savait aussi comme mathématicien que la question des erreurs n'était en rien mathématique au départ, mais plutôt une question de pratique de l'observation, de conditions atmosphériques, de qualité des instruments, d'équation personnelle, bref de la *métaphysique*, au sens cartésien de ce qui n'est ni physique ni formel, ou encore au sens kantien de ce qui relève de la raison pure (objets de la foi rationnelle), et des principes régulateurs de la pensée scientifique[4]. Lambert est l'un des premiers mathématiciens d'une longue série qui vont transformer le problème métaphysique (et épistémologique) de l'erreur en un problème mathématique.

Comment vont-ils le faire ? Par le moyen d'une opération appelée *réduction* des erreurs, qui vise à purifier l'erreur de ses constituants métaphysiques pour n'en retenir qu'une composante, modélisable par un calcul (voire deux calculs, comme on le verra). Cette purification s'obtient en séparant deux composantes de l'erreur ; l'une dite systématique, l'autre dite accidentelle. La première est définie comme le résultat d'un long procès fait aux observations, un procès au sens technique de processus, et au sens juridique de mise en accusation avec droits pour la défense. L'analyse a priori et a posteriori des conditions de l'observation par un professionnel de l'astronomie doit lui permettre d'identifier certaines causes d'erreur systématiques, que celles-ci soient dues à l'observateur, aux instruments, à leur interaction (la lecture d'une division) ou à d'autres facteurs (par exemple la parallaxe).

[3] Passage de l'Arctique permettant de relier de façon très réduite atlantique et pacifique, puisqu'il était obstrué par la banquise…jusqu'en septembre 2007 où l'on a constaté son ouverture à la navigation. « Les sciences exactes ne sont pas liées aux sciences humaines par un simple intervalle, une interface, ou un espace lisse. Le passage du Nord-Ouest correspond en image, à leurs relations compliquées. La route est coupée, elle est libre, le voyage est une aventure, il dépend des lieux, du temps, des circonstances. Tracé sur une carte, le chemin est cependant spécial, original, chaque traversée le trouve différent. »

[4] Lalande, *Vocabulaire technique et critique de la philosophie,* PUF, 1993, p.613

En supprimant ou contrôlant ces causes bien identifiées, on peut éliminer leurs effets et donc réduire l'erreur à sa partie « non expliquée ». C'est cela le grand partage qui permet d'extraire le minerai de l'erreur fortuite de sa gangue métaphysique. Mais comme tout dans le monde a une cause - c'est le *principe de raison suffisante* invoqué par Jacques Bernoulli, Leibnitz et John Locke - il est clair que les disparités d'observations qui demeurent après ce nettoyage ont elles aussi des causes, qui nous sont tout simplement inconnues. L'erreur qui demeure après cette « réduction » est attribuée à des causes dite accidentelles, ou encore « au hasard », si ce mot est bien le cache sexe de notre ignorance, de notre méconnaissance partielle de ce qui a troublé l'observation. Dans sa nature métaphysique, l'erreur accidentelle est une erreur épistémique.

En quoi consiste dès lors le fait de mathématiser cette erreur accidentelle ? Il s'agit tout simplement d'en faire une « variable » au sens du calcul analytique. Mais deux mathématisations possibles vont alors entrer en concurrence. Ce que Lambert ne manquera pas d'ailleurs d'étiqueter comme deux branches de la théorie des erreurs, mais qu'il voit comme complémentaires plus que concurrentes.

Prenons le cas d'une somme de n mesures y_i d'une même grandeur, de valeur vraie constante mais inconnue V, susceptibles chacune d'une erreur e_i, avec le modèle implicite $y_i=V+e_i$. Quelle est l'erreur sur la somme des observations $S = \sum y_i = nV + \sum e_i$ et sur la moyenne arithmétique des observations $M=S/n = V + \bar{e}$?

Si l'on s'appuie sur le calcul différentiel inventé quelques décennies plus tôt par Newton sous la forme du *calcul des fluxions*, et si l'on traduit cette erreur e_i par une différentielle, alors la règle de Roger Cotes (1722), élève de Newton, fournit le résultat suivant :

Pour une quantité $x = F(a,b,c,\ldots)$ dont les déterminants $a, b, c\ldots$ sont affectées des erreurs $da, db, dc\ldots$, on obtient pour erreur sur x, par développement de la fonction F, et en se restreignant à l'approximation de premier ordre :

$dx = F'_a\,da + F'_b\,db + F'_c\,dc + \ldots.$

Dans laquelle les F'_a F'_b F'_c sont les dérivées de F par rapport à a,b,c,….

Dans le cas particulier ou x = a + b + c +… ces dérivées sont égales à 1 et

$dx = da + db + dc + \ldots$[5]

Malheureusement les erreurs effectives (et algébriques) *da, db, dc* sont inconnues.

Une première solution est de raisonner en terme d'erreurs maximales ou « limites d'erreurs » comme le disent plusieurs auteurs du XIX^e^ pour donner des majorants d'erreur. L'erreur *maximale* (en valeur absolue) sur x est alors au pire égale à la somme des erreurs *maximales* sur a, b, c, … et si ces limites sont égales (à M) – cas de n mesure avec le même instrument dans les mêmes conditions – alors l'erreur maximale sur la somme S sera bien au plus de nM[6]: les erreurs (absolues maximales) s'ajoutent. Si je me trompe *au plus* de 1cm sur chaque mesure de largeur de cette table, je me trompe bien *au plus* de 10cm sur la mesure 10 tables (si je procède par addition de 10 mesures individuelles).

Si au contraire on s'appuie sur le calcul des probabilités, alors on sait par expérience depuis fort longtemps que se tromper ainsi de propos délibéré, en tombant n fois sur une

[5] soit $dS = \sum dy_i = \sum e_i$ dans nos notations

[6] Car $|\sum dy_i| \leq \sum |dy_i|$

erreur de même grandeur, serait aussi peu probable que d'obtenir n fois la même valeur aux dés. Les erreurs successives ne s'ajoutent point : elles se compensent. C'est ce que dit Pierre Bouguer (1749):

" S'il est possible que nous nous trompions sur chaque angle, il ne l'est pas moralement que toutes les erreurs soient égales, ni dans un sens précis qui font qu'elles s'accordent à s'ajouter ensemble dans le résultat. Il faudrait pour cela se tromper de propos délibéré (...) Dans quelques triangles il n'y aura pas d'erreur du tout, dans les autres, il y en aura un peu, mais elle sera de différents sens, et il se formera de tout cela une espèce de compensation."

Joseph Fourier (1829, p.568) reprend le même raisonnement :

«Il est extrêmement vraisemblable que, l'erreur négative étant supposée aussi facile que l'erreur positive, il s'établira dans un grand nombre d'erreurs partielles, une compensation qui tend à diminuer l'erreur totale. Il n'est pas rigoureusement impossible que les erreurs partielles, même en très grand nombre, soient toutes positives ou toutes négatives ; mais on ne doit point, dans la pratique, supposer qu'un tel événement a lieu, parce que sa probabilité est trop petite… »

Ces deux calculs – calcul différentiel, calcul aléatoire - sont ils si incohérents entre eux ? Pas exactement si l'on précise à la fois l'idée de somme et la sémantique de l'erreur. Que les quantités e_i désignent une « limite d'erreur » une « erreur absolue maximale à craindre », comme le dirait Laplace, et que l'on suppose « un même soin aux diverses mesures », donc une égalité de ces erreurs à craindre, et dès lors l'erreur maximale à craindre sur la somme sera bien n fois plus grande que sur chaque mesure comme le veut le calcul différentiel. Mais si par contre les erreurs e_i sont des « erreurs fortuites ou éventuelles » comme le dira Fourier, des erreurs conçues comme des variables *algébriques*, les unes positives les autres négatives, certaines grandes, d'autres petites, suivant donc une certaine loi de distribution sans doute paire et donnant autant de chance aux erreurs négatives qu'aux erreurs positives, alors l'erreur sur la somme sera bien la somme de ces erreurs, mais une somme algébrique dans laquelle la compensation va jouer de manière à presque annuler cette somme. Dès lors le calcul différentiel est réconcilié avec le calcul des probabilités.

Et c'est cela que réalise en 1755 le mathématicien autodidacte Thomas Simpson, vulgarisateur de l'œuvre de De Moivre, lorsqu'il démontre « l'avantage qu'il y a à prendre la moyenne des observations ». Il montre, par exemple pour $n = 6$ observations et pour une loi triangulaire, que $prob\ (|\overline{X}| \leq 2) = 0{,}967$ alors que $prob(|X| \leq 2) = 0.667$, et il peut conclure que "*Le fait de prendre la moyenne de plusieurs observations diminue grandement les chances de petites erreurs et élimine presque toute possibilité de grandes erreurs (...) Plus on fera d'observations ou d'expériences, moins les conclusions souffrirons d'erreur, pourvu que ces observations puissent être répétées dans les mêmes circonstances."*

Rappelons que des savants aussi importants que Léonard Euler en 1749 et Thomas Bayes en 1755 avaient refusé de l'admettre : le dernier écrivait : « *Que les erreurs provenant de l'imperfection des instruments et des organes des sens puissent être ainsi réduites à rien ou presque rien seulement en multipliant le nombre des observations me semble extrêmement incroyable. Au contraire plus vous faites d'observations avec un instrument imparfait, plus il semble certain que l'erreur dans votre conclusion sera proportionnelle à l'imperfection de l'instrument utilisé.»* Sans doute étaient ils marqués par les raisonnements sur les erreurs systématiques, non réduites à la seule composante accidentelle. D'autres, comme Joseph-Louis Lagrange, Pierre-Simon de Laplace ou Daniel Bernoulli reprendront la démonstration de cet avantage dans les années 1770. Et Carl Friedrich Gauss bien plus tard, imposant la mesure de dispersion par l'écart quadratique moyen (notre écart-type σ) montrera[7] que :

$$\sigma_S = \sqrt{n}\, \sigma_X \text{ et } \sigma(\overline{X}) = \sigma_X / \sqrt{n}.$$

Plus généralement, le théorème dit de Gauss-Markov énonce que pour tout modèle linéaire $y_i = \sum a_j x_{ij} + e_i$, la méthode des moindres carrés fournit les estimations les plus précises (de plus petite variance) des coefficients a_j inconnus, quelle que soit la loi des erreurs e_i. La formule de Gauss-Bessel permet une estimation a posteriori de la variance des erreurs par la variance corrigée des résidus observés, in fine, entre observations et vraie valeur estimée.

Dans le cas des observations directes (modèle $y_i = V + e_i$) la meilleure estimation de V est la moyenne arithmétique $\bar{y}$ des y_i et cette estimation est entachée d'une erreur $\sigma_{\bar{y}} = \sigma_e / \sqrt{n}$ [8]. Pour une certaine précision des mesures supposée la même pour toutes les observations et donnée par l'écart-type σ_e, la vraie valeur V est estimée avec une précision bien plus grande, proportionnelle[9] à $\sqrt{n}$, aussi grande que l'on veut en multipliant les observations. « *La précision du résultat moyen augmente comme la racine carrée du nombre des observations* » écrit Fourier (1826) et il en conclut que « *la multiplicité des observations supplée en quelque sorte à la connaissance des causes* ». Il ne craint pas d'écrire dans un second mémoire, de 1829, qu'en répétant indéfiniment le nombre des mesures on peut faire disparaître toutes les erreurs fortuites.

Dans ce dernier mémoire Fourier revient sur le rapport entre erreur maximale et erreur aléatoire. Dans ses notations cela donne ceci : Supposant une grandeur $x = F(a,b,c\ldots)$ et son cas particulier $x=a+b+c+\ldots$, il distingue

- (1) les quantités *da, db,dc…et dx*, qui représentent les erreurs effectives inconnues, et qui sont reliées par l'équation différentielle $dx = F_1 da + F_2 db + F_3 dc\ldots$
- (2) D*a*, D*b*, D*c*…, *Dx* qui sont « les limites des erreurs que l'on peut supposer » dans les mesures de *a,b,c,…x* , c'est-à-dire encore celles qui n'ont qu'une chance sur vingt mille d'être dépassées »
- et enfin (3) les quantités d*a*, d*b*,d*c*, ..d*x* qui désignent les erreurs *moyennes* sur ces mêmes mesures, c'est-à-dire au sens de Fourier, celles qui ont une probabilité ½ d'être dépassées en valeur absolue, et qui dépendent de la loi des erreurs. Cette erreur moyenne est donc ce que les auteurs contemporains appellent erreur probable, jusqu'à ce que Cournot lui préfère l'expression moins trompeuse d'erreur médiane.

Fourier démontre ensuite que « « En ne faisant ainsi aucun usage de l'Analyse des probabilités, on estimerait mal la précision du résultat…L'erreur que l'on déduirait de l'expression différentielle serait excessive » car la somme des valeurs absolues est toujours

[7] En fait comme on le sait, il préfère raisonner sur le module de précision $h = 1/\sigma\sqrt{2}$.
[8] Je pourrais tout aussi bien écrire de façon plus moderne que e_i étant une variable aléatoire, y_i est aussi une variable aléatoire dont la *variabilité* autour de V se caractérise par une dispersion σ_e tandis que $\bar{y}$ a une *variabilité* plus faible de dispersion $\sigma_{\bar{y}} = \sigma_e / \sqrt{n}$, mais cette interprétation en terme de variabilité serait anachronique.
[9] L'écart-type qui mesure la dispersion, donc l'inverse de la précision, est proportionnel à $1/\sqrt{n}$

supérieure à la racine de la somme des carrés (il en donne une preuve géométrique). Dans les cas (2) et (3) il s'appuie en effet sur une moyenne quadratique, qu'il donne dans ses deux mémoires sous la forme suivante :

$$g = \sqrt{\frac{2}{n}\sum_i (x_i - \bar{x})^2} \quad ,$$

équivalente au facteur $\sqrt{2}$ près à l'écart-type introduit bien plus tard par Karl Pearson.

Dès lors, les erreurs limites (D) et probables (d) sont liées aux erreurs sur a, b, c, … par une équation de forme quadratique $(dx)^2 = (F_1 da)^2 + (F_2 db)^2 + (F_3 dc)^2$….. Reprenant les valeurs de la table qu'il donne dans le premier mémoire sous l'hypothèse de normalité (il dit « d'une loi de Bernoulli »), Fourier propose de considérer comme erreur « moyenne » - c'est-à-dire médiane - la quantité *0,47708.g* et comme limite de la plus grande erreur la quantité *3g* ce qui correspond à des intervalles de confiance de 50% et de 99,998 %).

Et ce qu'il propose est donc d'appliquer la règle du $\sqrt{n}$ qu'il formule ainsi : Dans le cas de la mesure d'une longueur composée d'un grand nombre de parties pour lesquelles on considère les erreurs moyennes comme égales ($da = db = dc$), « si l'on désigne par *e* la limite de l'erreur positive ou négative à laquelle on est exposé en mesurant chaque partie, il faut multiplier cette limite *e* par la racine carrée du nombre *n* des parties et le produit $e\sqrt{n}$ est précisément la limite de l'erreur positive ou négative que l'on peut commettre dans la longueur entière… ». C'est notre formule $\sigma_S = \sqrt{n}\ \sigma_X$. L'application qu'il fait de ce principe à la hauteur de la pyramide de Cheops est un des exploits les plus connus de notre savant dans l'expédition d'Egypte : Cherchant à évaluer cette hauteur par la somme des hauteurs prises sur les 203 marches de la pyramide, Fourier demande quelle était la précision à attendre d'une telle mesure? 14 fois celle des erreurs partielles sur les « assises » répond-il sans nous préciser ce qu'était l'ordre de grandeur de ces erreurs partielles. Assez faible assurément pour qu'il puisse se réjouir de la congruence de ce résultat avec celui de la méthode concurrente du sextant.

Curieusement, ce mémoire se termine ainsi : « Nous pensons que la publication de ces théorèmes sur les erreurs de mesure et sur la précision des résultats du calcul contribueront à perfectionner les applications des sciences mathématiques. Ces considérations appartenaient *naturellement* à une Collection qui a pour objet d'observer et de constater tous les principaux éléments de la prospérité publique.» Rien de moins naturel pour nous que cette publication mathématique se fasse comme la précédente dans les *Recherches statistiques de la Ville de Paris* par le Préfet Chabrol. Mais ce serait sans compter le rôle que Fourier et Chabrol attribuent à la méthode statistique dans la formation des élites politiques et la saine administration, comme en témoigne par exemple la gestion par eux du prix Montyon (Brian, 1995). Stéphane Callens (1994) écrit « le pacte entre le préfet de la Seine et son directeur des travaux statistiques est basé sur la compatibilité entre deux propositions : « les faits modèrent les opinions » dit le Préfet et « les données possèdent des régularités à découvrir » dit le directeur des travaux statistiques.

Adolphe Quetelet, l'homme moyen et le facteur 33

Louis-Adolphe Quetelet est d'abord un jeune et brillant mathématicien et poète chargé de mettre en place l'Observatoire de Bruxelles et qui a ce titre visite les principaux mathématiciens français engagés dans les applications astronomiques, le calcul des probabilités, et la statistique telle que Fourier la défend dans le prix Montyon. De retour à Bruxelles Quetelet y mène une quadruple carrière de mathématicien, de directeur de l'Observatoire, d'enseignement propagandiste du calcul des probabilités et de chercheur en statistique. Ces quatre facettes sont illustrées par la publication de la *Correspondance mathématique* (1825), du *Traité d'astronomie élémentaire* (1826), des *Instructions populaires sur le calcul des probabilités* (1828), et de l'*Essai de physique sociale* intitulé d'abord *Sur l'Homme et le développement de ses facultés* (1835). Dans cet ouvrage, il se pique d'étudier d'un point de vue de la science naturelle et avec l'appui des statistiques les qualités physiques (anthropométrie), biologiques (démographie) et morales (criminologie) de l'homme. Dans une quatrième partie, il développe sa théorie de l'homme moyen. Or celle-ci n'est rien d'autre qu'un vaste projet de transposition de la théorie des erreurs astronomiques à une sociologie qu'il appelle physique sociale, comme en témoigne cette lettre à son éditeur Sylvain van der Weyer du 22 août 1834 :

"Je me fais peut-être illusion, mais j'ai lieu de croire que le travail fera plaisir ... La partie la plus curieuse du travail sera, je crois, la théorie de la population. Je suis parvenu à la <u>transporter</u> entièrement dans le domaine des sciences exactes au moyen de deux théorèmes dont l'un m'appartient. (...) On pourra résoudre les grands problèmes des mouvements de population comme ceux des <u>mouvements des corps célestes</u>; et ce qu'il y a de plus remarquable, c'est l'étonnante <u>analogie</u> qui existe entre les formules qui servent à ces calculs. Je crois avoir réalisé en partie ce que j'ai dit depuis longtemps sur la possibilité de faire <u>une mécanique sociale comme l'on a une mécanique céleste</u>; de formuler les mouvements du corps social comme on a formulé les mouvements des corps célestes et d'en reconnaître toutes les propriétés et les lois conservatrices. »

Cette transposition analogique, il va la préciser tout au long des ouvrages suivants, mais elle n'est nulle part aussi bien exposée que dans ses *Lettres à S.A.R. le Duc Règnant de Saxe-Cobourg et Gotha, sur la théorie des probabilités appliquées aux sciences morales et politiques,* publiées en 1846, et véritable « cours à distance» de probabilité aux qualités pédagogiques inégalées. La Lettre 20 est sans doute celle qui montre le mieux comment sa théorie de l'homme moyen se rattache à la théorie des erreurs.

La démonstration tient en 4 étapes. Il commence d'abord (1) par imaginer que l'on cherche à mesurer la poitrine de la statue du gladiateur, qui représente pour lui à la fois le type de l'homme et le comble de la beauté.

Si l'on avait le courage de recommencer mille fois, on finirait par avoir une série de nombres qui différeraient entre eux selon le degré de précision qu'on aurait mis a les recueillir. La moyenne de tous ces nombres s'écarterait certainement très peu de la véritable valeur. De plus, en classant toutes les mesures par ordre de grandeur, on ne serait pas médiocrement étonné de voir les groupes se succéder avec la régularité la plus grande. Les mesures qui s'écartent le moins de la moyenne générale composeraient le groupe le plus considérable ; et les autres groupes seraient d'autant plus petits qu'ils contiendraient des mesures plus en désaccord avec cette même moyenne. Si l'on figurait la succession des groupes par une ligne, Votre Altesse a déjà deviné que cette ligne serait la courbe de possibilité ; ce résultat, en effet, était à prévoir. En sorte que la maladresse, ou le hasard, si nous aimons mieux ce mot pour couvrir notre amour-propre, procède avec une régularité qu'on ne serait guère tenté de lui attribuer.

Je suppose maintenant qu'on réunisse les cinq cents mesures qui s'écartent le moins de la moyenne; la demi différence qui se trouvera entre la plus grande et la plus petite de toutes *ces mesures, sera le module de la précision ou l'erreur probable. Il pourrait se faire que, dans les circonstances actuelles, cette erreur probable ne fut que d'un millimètre en sorte que, sur les mille mesures, cinq cents seraient en erreur de moins d'un millimètre , et cinq cents autres seraient en erreur de plus d'un millimètre. On aurait ainsi 1 contre 1 a parier qu'en prenant une nouvelle mesure, on ne s'écarterait pas d'un millimètre de la moyenne de toutes les mesures, laquelle peut être considérée comme la véritable circonférence qu'on voulait apprécier.*

Nous sommes ici dans un exercice qui est exactement celui de la répétition des mesures astronomiques, avec un objet de référence plus accessible que les astres. Mais on y retrouve la distribution en cloche des mesures, leur moyenne comme meilleure estimation, et l'erreur probable (de 1 mm) ayant une probabilité de 50% d'être dépassée.

Toute la puissance du raisonnement de Quetelet vise à nous faire passer de cette situation de la mesure astronomique à celle de la mesure statistique des tailles de 1000 individus. Il y parvient en passant par deux étapes intermédiaires. La seconde étape (2) consiste à remplacer la statue par une personne vivante :

Si l'on avait à mesurer la poitrine d'une personne vivante au lieu de celle d'une statue, les chances d'erreur seraient beaucoup plus nombreuses ; et je doute fort qu'après mille mesures, on trouvât encore une erreur probable de 1 millimètre. Le seul acte de la respiration qui fait varier a chaque instant la forme et les dimensions de la poitrine, ajouterait une puissante cause d'erreur à toutes celles qu'on rencontre déjà en opérant sur une statue parfaitement immobile. Malgré ce désavantage, les mille mesures groupées par rangs de grandeur procéderaient cependant encore d'une manière très régulière. La ligne qui les représenterait, serait toujours la courbe de possibilité, mais dilatée dans le sens horizontal, proportionnellement à l'erreur probable.

La troisième étape (3) consiste à remplacer les mille mesures du gladiateur par mille mesures de mille copies différentes du gladiateur :

Modifions encore notre hypothèse, et supposons qu'on ait employé un millier de statuaires pour copier le gladiateur avec tout le soin imaginable. Votre Altesse ne pense certainement pas que les mille copies qui auront été faites, reproduiront chacune exactement le modèle, et qu'en les mesurant successivement, les mille mesures que j'obtiendrais seraient aussi concordantes que si je les avais prises toutes sur la statue du gladiateur même. Aux premières chances d'erreur viendraient se joindre les inexactitudes des copistes ; en sorte que l'erreur probable serait peut-être très grande. Malgré cela, si les copistes n'ont pas travaillé avec des idées préconçues, en exagérant ou en diminuant certaines proportions d'après des préjugés d'école, et si leurs inexactitudes ne sont qu'accidentelles, les mille mesures, groupées par ordre de grandeur, présenteront encore une régularité remarquable et se succéderont dans l'ordre que leur assigne la loi de possibilité.

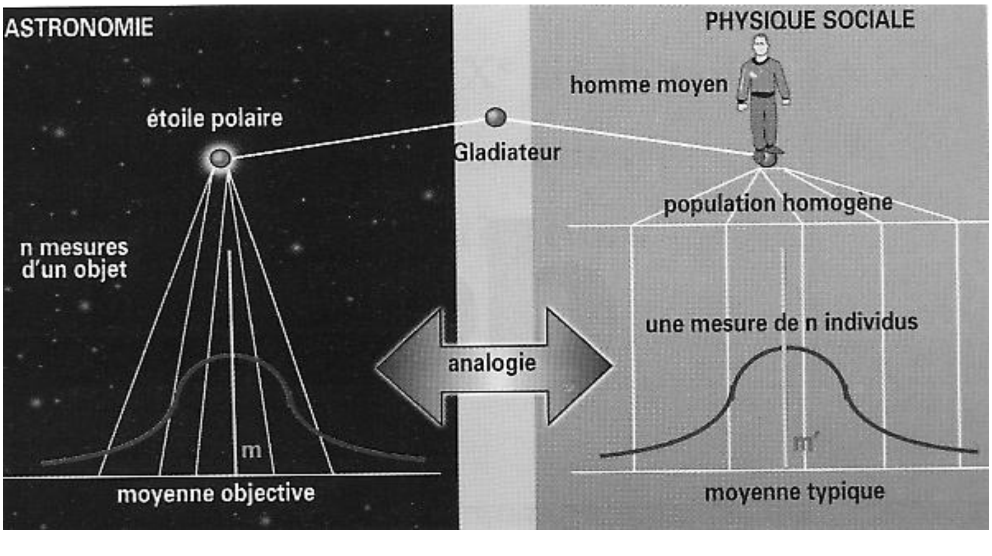
La quatrième et dernière étape du raisonnement consiste à combiner les deux derniers déplacements analogiques en étudiant le cas de mille mesures de mille personnes vivantes différentes, les fameux soldats écossais :

Je vois sourire Votre Altesse : elle me dira sans doute que de pareilles assertions ne me compromettront pas, attendu qu'on ne sera pas disposé à tenter l'expérience. Et pourquoi pas? Je vais peut-être bien l'étonner, en disant que l'expérience est toute faite. Oui vraiment, on a mesuré plus d'un millier de copies d'une statue que je n'assurerai pas être celle du *gladiateur, mais qui, en tout cas, s'en éloigne peu : ces copies étaient même vivantes. En sorte que les mesures ont été prises avec toutes les chances d'erreur possible : j'ajouterai, de plus, que les copies ont pu se déformer par une foule de causes accidentelles. On doit donc s'attendre ici, à trouver une erreur probable très sensible.*

J'en viens au fait. On trouve, dans le 3e volume du journal médical d'Edimbourg, les résultats de 5738 mesures prises sur les poitrines des soldats des différents régiments écossais. Ces mesures sont exprimées en pouces anglais et groupées par ordre de grandeur, en procédant par différences de 1 pouce. La plus petite mesure est de 33 pouces environ, et la plus grande de 48, la moyenne de toutes les mesures donne un peu plus de 40 pouces pour circonférence de la poitrine d'un soldat écossais : c'est aussi le nombre qui correspond au plus grand groupe de mesures ; et, comme la théorie l'indique, les autres groupes, diminuent de grandeur à mesure qu'ils s'éloignent de celui-ci , l'écart probable est de 1P,312 ou de 33mm,34. Je prie Votre Altesse de ne pas perdre de vue cette valeur.

La conclusion de Quetelet est simple et radicale : malgré les différences, et à un facteur d'échelle près, on peut faire *comme si* la dernière expérience, statistique, était similaire à la première, celle des observations entachées d'erreurs :

Eh bien, en admettant cette erreur probable [de 33 mm], 5 738 mesures prises sur une même personne ne se grouperaient certainement pas avec plus de régularité quant à l'ordre de grandeur, que les 5 738 mesures prises sur les soldats écossais. Et si l'on nous donnait les deux séries de mesures sans les avoir désignées d'une manière particulière, nous serions très embarrassés de dire quelle série a été prise sur 5 738 soldats différents, et quelle série a été obtenue sur une seule et même personne, avec moins d'habitude et des moyens d'appréciation plus grossiers.



J'ai longtemps présenté ce texte comme la forme la plus pédagogique pour introduire la notion si ambiguë mais si importante de l'homme moyen, notion qui aura un grand avenir comme paradigme statistique des sciences sociales qui naissent à la fin du 19e siècle. En effet dans l'analogie entre les mille mesures de la statue et les milles mesures des mille soldats, on a d'abord considéré que l'essentiel était, comme dans toute analogie la transformation d'une distinction - de la moyenne objective (ayant un objet référent unique) de la moyenne subjective (l'homme moyen, n'ayant aucune réalité) - en une assimilation, coup de génie de Quetelet, qui est aussi un coup de force.

Il m'apparaît maintenant qu'un élément a résisté à l'analogie : c'est la quantification de la dispersion qui fait les frais de l'assimilation des deux situations. Son ordre de grandeur n'est vraiment pas le même: 1mm pour une erreur probable sur la mesure de la statue ; 33 mm pour l'erreur probable sur la taille des soldats écossais ! Ce facteur 33 n'est pas anodin. On pourrait dire qu'à ce niveau le changement qualitatif est devenu un changement quantitatif. Et tenter de nous convaincre comme le fait Quetelet que nous ne sommes que de pâles copies d'un modèle préexistant, et que cette oscillation des individus autour d'un type fixe est du même ordre que celle des mesures autour de la vraie valeur, c'est ignorer ce facteur 33, c'est vraiment opérer un coup de force de grande ampleur.

A quoi faut-il attribuer ce facteur ? Caractérise-t-il une quelconque essence des objets des sciences de la nature et des sciences de l'homme ? L'infranchissable passage du Nord Ouest ? On peut esquisser un premier raisonnement dans ce sens : dans la mesure physique ou astronomique d'objets inanimés, la variabilité de l'erreur est limitée: on ne saurait se tromper de plus de quelques millimètres en mesurant un objet comme ma table avec un mètre à ruban, et de quelques microns avec des appareils plus sophistiqués. Le progrès constant des instruments et de la méthodologie ont sans arrêt réduit la plage de variation des erreurs de mesure, nous faisant passer par exemple pour un angle astronomique de 30' pour le bâton de Jacob à 2' pour le quadrant mural de Tycho Brahé vers 1570, 15'' pour le quart de cercle de Picard en 1667, 2'' pour celui de Cassini en 1744 etc…Pour ce qui est des populations humaines, aucune réduction de la variabilité ne peut se faire par élimination des causes constantes de déviation comme dans le cas des mesures d'un angle, et la variabilité des propriétés physiques, mentales et comportementales des humains n'a pas cette propriété de diminuer sans cesse avec le progrès technique. Les fluctuations de la dispersion des revenus ou des patrimoines caractérisent des états de la société qui peuvent être très différents d'un groupe à l'autre, d'une époque à l'autre. Dans un cas l'erreur est au hasard par rapport à la vraie valeur qui reste au centre des fluctuations de l'observation. Dans l'autre cas la variabilité des caractéristiques humaines dans un groupe social n'est pas un bruit qui brouille un ordre immuable, c'est un bruit total, un hasard absolu, un phénomène en soi qu'il convient d'expliquer. Il nous faut donc revenir à la nature de cette variabilité, à ce qui sépare les deux éléments de l'analogie, à ce que Quetelet distingue pour mieux les confondre par la suite.

Observations et statistique

Les statisticiens d'avant 1830 pouvaient-ils faire la distinction entre la dispersion des *n* mesures d'un même objet - voire d'un même sujet – et celle de n mesures prises 1 fois sur chacun des *n* objets ou sujets d'un groupe ? La différence de nature entre ce que nous appellerons avec Francis Ysidro Edgeworth (1885) des *observations* et des *statistiques* est pourtant limpide :

"*Observations and statistics agree in being quantities grouped about a Mean : they differ in that the Mean of observations is real, of statistics is fictitious. The mean of observations is a cause, as it where the source from which diverging errors emanate. The mean of statistics is a description, a representative quantity put for a whole group (...) Measurements by reductions of which we ascertain a real time, number, distance are observations. Return of prices, exports and imports, legitimate and illegitimate marriages or births and so forth, the averages of which constitute the premises of practical reasoning are* *statistics. In short observations are different copies of one original; statistics are different originals affording one generic portrait. Different measurement of the same man are observations; but measurements of different men, grouped round l'homme moyen are primâ facie at least statistics*"

Mais cette différentiation n'est pas forcément explicitée avant 1846. La Statistique des préfets conduit à des monographies de départements et ne se donne pas les moyens d'une mise en série des observations statistiques. Dans les travaux d'arithmétique politique les estimations numériques se font le plus souvent sur la base d'hypothèses (taux de mortalité, multiplicateur des naissances, rapport des charrues et attelages à la production…) que sur la base de résumés statistiques. Condorcet développe une vision du progrès des sociétés qui s'inspire d'une application de l'analyse mathématique et du calcul des probabilités qu'il ne met pas en œuvre sur des « données ». Dans son *Essai,* Laplace assimile aux tirages dans une urne avec des boules indiscernables sauf par leur couleur aussi bien les résultats d'une « répétition des observations » que d'une « multiplication indéfinie des événements ». La coupure faite entre le calcul des probabilités et une statistique, maintenue par l'Académie dans l'état d'une description comparative des États limite les usages conjecturaux des statistiques. Ceux qui manipulaient à la fois des observations naturelles et des observations humaines et disposant de nombreuses observations étaient par exemple les pasteurs ou les astronomes qui travaillèrent sur les tables de mortalité. Et bien évidemment si la répétition des observations sur un *même* corps céleste était possible, conduisant d'ailleurs à une dispersion des résultats qu'il pouvait attribuer soit au mouvement de l'astre soit au mouvement de l'observateur, la répétition des observations de durée de vie ne pouvait s'effectuer sur le *même* homme. Il faut attendre la création de la SGF et la multiplication des services de statistique des administrations pour que se produise « une avalanche de chiffres », comme le souligne Ian Hacking, qui 1°) permet des mises en série statistiques des observations 2°) amène à se poser des questions sur l'induction qu'autorise la nature de ces séries.

Or le point clé qui permet de différencier le cas des observations de celui des statistiques est l'interprétation que l'on peut faire de l'unicité du référent : soit la mesure se fait sur le *même* objet ou homme (qui est unique et garde son intégrité) soit sur des objets ou des hommes qui sont les *mêmes* c'est-à-dire qui sont similaires d'un certain point de vue. En cela la différence ne porte pas tant sur le fait que l'objet est naturel ou social, mais sur le fait que les objets ou individus mesurés peuvent être considérés comme indiscernables ou non. Ce qui est une affaire de convention ou de jugement a priori. Par exemple toutes les feuilles d'un chêne auront à peu près la même taille et la même forme. Il se trouvera même un statisticien allemand, Gustav Rümelin, pour nous expliquer que « *dans la nature, tout cas particulier peut servir de type, en sorte qu'un seul cas bien observé autorise déjà une induction (..) Mais que notre assurance diminue déjà quand nous passons aux plantes et aux animaux qui subissent l'action de l'homme, et qu'elle s'évanouit tout à fait quand nous entrons dans le domaine de l'âme humaine* » Puis nous expliquer que « *la même graduation se continue jusque dans le règne humain. Un sauvage est plus typique que l'homme civilisé; le Nègre et le Mongol le sont davantage que l'Européen; l'homme de l'Antiquité l'est plus que celui du Moyen-Age, et ce dernier que l'homme moderne. L'homme est plus individuel que la femme, l'adulte que l'enfant, l'homme cultivé que l'inculte, l'homme noble que l'homme vulgaire* ». Les régularités de Quetelet ne sont alors que les attributs de certaines espèces.

En affaiblissant et relativisant l'idée de similitude, on va trouver la notion d'homogénéité du groupe humain étudié, qui est pour Quetelet la pierre de touche de son analogie, et la raison d'être d'une assimilation possible des statistiques à des observations. Le statut de cette homogénéité n'est pas si clair : pour le Quetelet de la Lettre XX il n'est attesté que par une caractéristique de la distribution, la courbe des possibilités, en forme de cloche. Dans une des Lettres suivantes, Quetelet précise que si l'on avait mélangé « un régiment de cuirassiers ne contenant que des hommes très grands et très vigoureux et un régiment de chasseurs composé d'homme beaucoup plus petits » ou encore « des patagons et des lapons », dans de tels cas, écrit-il, « on pourrait toujours prendre une moyenne arithmétique (Edgeworth écrirait *average*) mais on n'aurait pas une véritable moyenne (Edgeworth écrirait *mean*).

Le statisticien démographe Louis Adolphe Bertillon reprendra dans l'article « Moyenne » (1876) cette distinction entre « moyennes objective et subjective », et , tout en brocardant la chimère de l'homme moyen, reprendra en le distordant le message de Quetelet, sur la condition d'homogénéité ; mais tandis que cette condition était chez Quetelet dans la forme même de la loi en cloche qui revient comme un leitmotiv dans la Lettre XX, elle devient chez Bertillon une propriété a priori des « groupes naturels » qu'il faut établir de manière exogène au modèle, par une chasse aux « collectivités factices » comme les conscrits du Doubs[10], suspectés de mélanger deux sous populations de souches distinctes. Si l'homogénéité est une condition du transport analogique, son statut varie grandement : c'est pour Quetelet une évidence générale, fondée sur un a priori monogéniste, et appuyée sur l'analogie formelle des courbes en cloches, c'est pour Bertillon une propriété supposée de certaines données liées à la nature du groupe naturel, et ce sera pour les statisticiens allemands de l'école de Lexis une hypothèse à tester sur les données.

Si la différence entre observations et statistiques est due pour l'essentiel à la présence d'un référent de la mesure qui est dans le premier cas l'objet mesuré, tandis que les mesures statistiques n'ont pas de référent puisqu'il n'y a plus de référent là sauf à en inventer un dans ce cas qui est l'homme moyen, cette question rejoint alors celle du réalisme. L'objet préexiste-t-il vraiment à sa mesure dans le cas des observations, comme le pensent en général les astronomes où bien l'objet n'existe-t-il que par sa mesure, comme l'affirment les constructivistes, de Jean Ullmo (1969) - « la grandeur naît du moment où elle est mesurée » - à Pierre Bourdieu et Bruno Latour - « l'objet est construit » -, le réel ne pouvant finalement se définir qu'a posteriori comme ce qui résiste aux entreprises de déstabilisation de la construction ? Symétriquement, l'homme moyen qui est le nom de ce que désigne a posteriori la mesure chez Quetelet, est il un simple concept, « centre de gravité du corps social », comme aime à le dire Quetelet, une pure construction de l'esprit, aussi monstrueuse qu'irréelle disent ses détracteurs, ou bien désigne-t-elle quelque essence préexistante à la mesure, ancrée au cœur du réel, qu'elle soit le résultat d'un dessein divin, ou d'une constitution propre du monde ? Trancher sur ces deux points permettrait de mieux distinguer la nature de la variabilité des observations et des statistiques, mais le débat sur les deux points semble encore ouvert.

Dans sa thèse (1994, 1997), Stéphane Callens s'était intéressé à la position de Quetelet par rapport aux positivismes, celui de Comte bien sûr mais aussi celui de Littré, des Bertillons, et de Broca. Leur reconnaissant comme caractéristiques communes un attachement viscéral au réel observable et une volonté « de faire succéder un état positif à un état métaphysique et abstrait », mais en privilégiant la seule méthode historique, Auguste Comte s'est totalement démarqué de Quetelet. Callens (1994, p.234) finit par déclarer que « le rejet de Quetelet fédère les positivismes », le désaccord étant total, on le sait, sur le rôle attribué au calcul des probabilités, mais aussi sur l'idéalité de l'homme moyen, son étrange statut

[10] Voir E. Brian, 1991 pour les différents artefacts à l'œuvre dans cet exemple.

d'abstraction, de fiction désincarnée, décrochée de l'influence des races et religions, des milieux ambiants, de l'environnement naturel et social.

Ted Porter (1997) se pose la même question – *Was Quetelet a positivist ?* – et montre que les arguments sont balancés : si le terme est utilisé dans son sens commun pour caractériser une science qui renonce à tout a priori et à l'inaccessible « chose en soi », et qui ne procède que des faits observés, de leur mesure et des régularités qu'elles peuvent manifester, alors incontestablement Quetelet est un positiviste. Mais si l'on suit Quetelet lorsqu'il déifie sa créature de l'homme moyen pour en faire l'essence même d'une population et la base de la physique sociale, alors on ne peut manquer de voir dans cette création la main d'un réaliste pour qui les apparences ne suffisent pas, puisqu'il faut recréer des entités nouvelles pour les comprendre; sa mécanique sociale s'apparente à un réductionnisme, assez métaphysique, ayant besoin des notions de cause et de propension, en rupture donc avec la thèse des positivistes. Porter s'en tire en caractérisant Quetelet comme un « positiviste malgré lui » avide d'explication par des entités essentielles, mais assez sage pour se contenter de régularités. Ce faisant il conclut à l'inverse de Callens.

Notre position serait plus proche de cette vision, mais consisterait à lui donner une expression historique. Quetelet est dans ses premiers travaux totalement redevable à l'Ecole de Laplace et Poisson dont le mécanisme est totalement prégnant. La rhétorique du « lieu vrai des corps (célestes) » et des causes constantes, variables et accidentelles domine leurs discours. Mais l'avalanche de données conduit Quetelet à lire celles-ci comme des mouvements de l'espèce humaine dans le temps ou dans l'espace et y trouver des lois qui n'ont point d'autre fondement que les régularités. La volée de critiques que provoque son concept d'homme moyen l'amène à restreindre le credo mécaniste. Et dans ce sens, si les positivistes de Emile Littré à Paul Broca reprennent le credo de la moyenne, il faut concéder à Callens qu'ils en rejettent l'interprétation en terme d'homme moyen.

Louis Adolphe Bertillon (1876) et dans une moindre mesure Emile Cheysson (1886) abandonnent en effet totalement et le calcul des probabilités et le concept d'homme moyen pour ne retenir que le rôle magique de filtrage des moyennes et en répandre une théorie très idéologique : « *C'est la moyenne qui règle notre vie, notre loyer, nos plaisirs, nos toilettes, en un mot notre budget. Les phénomènes démographiques, les faits sociaux eux-mêmes subissent l'empire de la moyenne (…) L'univers aspire à la moyenne , qui n'est qu'une autre expression de l'équilibre.* » Avec cette évolution vers un discours de la moyenne omniprésente, que devient le souci d'une variabilité des comportements ? Il est évidemment complètement éludé. Il n'y a plus ni variation autour d'un type, fut il idolâtré, ni ce que j'ai appelé hasard absolu d'une variation sans référentiel, d'une variation pour soi.

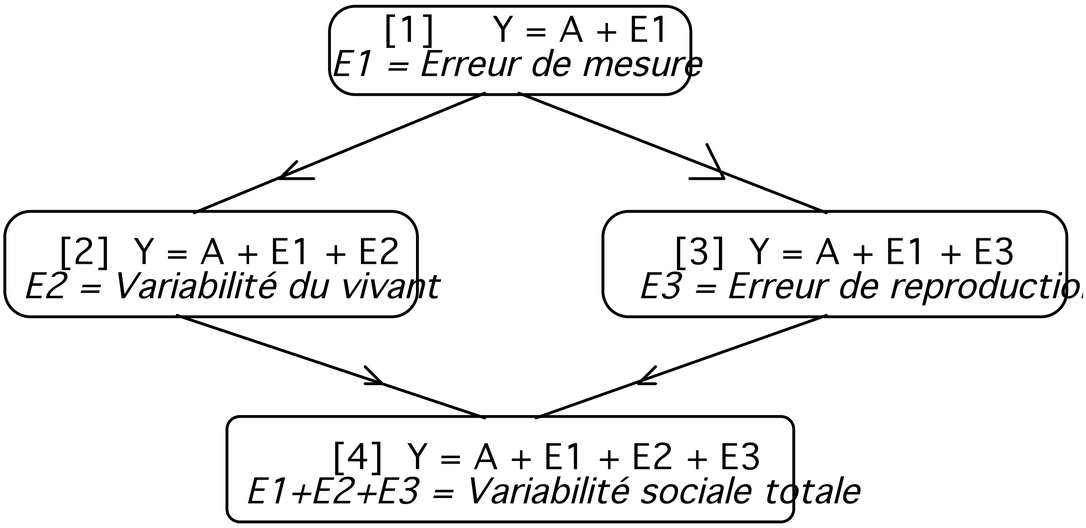
La résurrection d'une certaine préoccupation « pour le facteur 33 » est l'objet du travail de Karl Pearson autour de la notion de contingence. Cette notion, développée abondamment par les classiques et en particulier par Jacques Bernoulli dans son *Ars Conjectandi*, avait été reléguée au second plan par l'école de Laplace. En faisant du hasard le cache sexe de notre ignorance, il avait réduit la contingence à un supplément d'âme du mécanisme absolu règnant par principe sur le monde. Grâce au calcul des probabilités, l'erreur d'observation qui perturbe légèrement l'action des causes constantes par des causes accidentelles pouvait être éliminée. Cette notion de contingence devient au contraire pour Karl Pearson l'alpha et l'omega de la philosophie phénoménaliste et antiréaliste exprimée dans sa *Grammaire de la Science* - la science ne traite que des phénomènes compris comme des routines de perception – et aussi la condition même qui fait de la méthode statistique la clé des sciences d'observation. Le renversement de la notion de hasard est complet. L'astronome concevait d'abord la vraie valeur, ou la vraie loi qui relie par exemple linéairement deux variables (par exemple le sinus carré de la latitude et la longueur d'un degré sur la géode supposée conforme à un ellipsoïde) puis constatait son brouillage par l'erreur aléatoire produisant une dispersion des observations autour de cette référence, et dès lors il cherchait à extraire cette vrai valeur ou vrai loi par une prise de moyenne ou une procédure des moindres carrés. Le statisticien selon Pearson n'a pas de vraie valeur ou de loi hypothétique au départ, et par ses tables ou ses graphiques en nuages de points, il accède directement à une dispersion importante des observations, à une variabilité intrinsèque qu'il interprète comme contingence pure. Identifier au milieu de cette table ou de ce nuage de points un centre, une droite, bref une loi qui ne sera qu'un résumé d'une « routine de perception » du phénomène, est une opération seconde qui permettra uniquement de résumer et caractériser cette contingence par une moyenne ou un coefficient de corrélation, sans rien avancer de sa nature profonde.

Les composantes de la variabilité sociale

La position de « positiviste contrarié » prise par Quetelet le situe clairement comme un intermédiaire dans ce chemin qui va de Laplace à Pearson, du traitement des observations brouillées par des erreurs à celui des statistiques reflétant une contingence pure. Lorsque Quetelet se concentre sur l'émergence de l'homme moyen, les fluctuations autour de la moyenne sont dès lors perdues de vue, la contingence ne l'intéresse pas en elle-même, et la régularité dans le temps ou dans l'espace constitue a elle seule le programme de sa physique sociale. Le rapport de 1 à 33 des variabilités ne joue d'autre rôle dans son analogie que celui d'une simple différence d'échelle dans l'assimilation qui est faite entre observations et statistiques. Tout cela est vrai si l'on s'en tient au point de départ et au point d'arrivée de sa démonstration. Mais revenons maintenant aux intermédiaires qu'il se choisit pour y arriver, aux étapes par lesquelles il passe de (1), la moyenne des observations prises sur la statue, à (4), la moyenne des statistiques des soldats écossais.

Reprenant son raisonnement, on est frappé par le fait qu'il n'a pas choisi au hasard les situations intermédiaires (2) et (3) entre observations et statistiques. Quelle est la petite différence introduite entre (1) et (2) ? Le passage de l'objet inanimé au vivant, tout simplement. Le seul acte de la respiration, dit-il, ajouterait une nouvelle et puissante cause d'erreur à la précédente. Mais s'agit-il encore d'une erreur ? Évidemment non. C'est une nouvelle cause de *variabilité* de la mesure, sous l'influence d'une cause variable – la respiration – propre au vivant et qui s'ajoute à la cause constante qui serait ici l'existence du même individu derrière les 1000 mesures. Cette variabilité est temporelle ; c'est celle d'un *mouvement* du corps. Cette variabilité temporelle caractérise la stabilité ou l'instabilité des phénomènes. Quetelet ne dit rien de son ordre de grandeur sous la toise de l'erreur probable. 10, 15, 20 mm ? Il serait en tout cas facile de la connaître et cette donnée figure sans doute dans les différentes études d'Anthropométrie qu'il a collationnées dans la *Physique sociale*. Dans la situation intermédiaire (3), on abandonne le vivant, on repart de la situation (1) de la mesure d'une statue inanimée mais celle-ci n'est plus unique ; elle est reproduite à 1000 exemplaire par mille statuaires copistes, qui « n'ont pas travaillé avec des idées préconçues, ou des préjugés d'école », et ne sont pas du genre de Fernando Botero à vous faire des grosses dames à partir de modèles communs. A l'erreur de mesure de ces copies s'ajoute donc une erreur de reproduction due « aux inexactitudes des copistes ». Évidemment il y a derrière cette idée farfelue des copistes une théorie sociale, celle de la variabilité des individus d'un même groupe, que Quetelet explicitera complètement dans d'autres pages : nous autres humains, nous ne sommes que très peu différents les uns des autres, car nous ne sommes que les copies légèrement altérées d'un même modèle, le type humain, qui a présidé à notre fabrication, que l'on désigne sous ce terme un processus divin, un processus biologique, ou une pure fiction : « tout se passe comme si… ». Cette variabilité dans l'espace caractérise l'homogénéité ou l'hétérogénéité du groupe. Là aussi Quetelet se garde de donner un ordre de grandeur absolu de l'erreur probable pour les principales caractéristiques humaines. Seul importe la dilatation de l'échelle de la « loi de possibilité ».

Pour finir, dans la situation de la statistique des soldats écossais, l'erreur que nous préférons appeler variabilité sociale apparaît bien comme la somme de trois variabilités de statut bien différent : la première est une *erreur* de relevé et de mesure, qui vient d'ailleurs en dernier dans l'ordre des choses, et dont la théorie des erreurs a quelque légitimité à rendre compte ; la seconde est une variabilité temporelle et intra-individuelle qui caractérise l'aptitude de chaque vivant au mouvement (du corps et de l'âme pourrait-on dire) et la troisième est une variabilité spatiale et inter-individuelle qui caractérise la dispersion des individus, au même sens que celui que pointe le terme de diaspora.



Une telle décomposition de la variabilité par Quetelet est remarquable à double titre. Elle montre que sa pensée n'est pas obnubilée par la seule moyenne et qu'il a encore en tête les essences diverses des « erreurs » qui composent la variabilité sociale. Elle montre aussi qu'au milieu du XIX^{e} siècle, on a déjà les prémices d'une sorte de décomposition de la variabilité sur deux facteurs – le temps et l'espace – qui sera à l'œuvre dans la statistique mathématique de Bortkiewicz et dans celle de Fisher, puis en économétrie où le traitement de données de panel permet de saisir simultanément les variations dans l'espace et dans le temps. Cela est d'autant plus important que le calcul des élasticités, c'est-à-dire des sensibilités d'une variation de y à celle de x, par exemple de la consommation aux prix, donne des résultats différents sur des données en coupe instantanée et en série chronologiques.

Variabilité des chances, hétérogénéité, instabilité

Après une trentaine d'années d'exploitation tous azimuts du paradigme queteletien de l'homme moyen (Bertillon, Cheysson, mais aussi Henry-Thomas Buckle en histoire, Emile Durkheim et Maurice Halbwachs en sociologie….) et d'âpres discussions de la légitimité d'une physique sociale fondée sur la théorie des erreurs, la fin du XIXe siècle fut marquée par une remise en cause assez radicale de ses hypothèses, et en particulier de l'assimilation entre théorie des erreurs d'observation et statistique. Les premières corrections de la théorie de Laplace et de ses prolongements par Quetelet ont lieu dans *l'Exposition sur la théorie des chances* de Cournot. On y trouve par exemple un rejet de la probabilité subjective et épistémique de Laplace et Condorcet au profit d'une probabilité objective, liée à la possibilité physique, et seule digne de mathématisation. On y trouve une révision du déterminisme laplacien et de la rhétorique des causes constantes ou accidentelles. L'urne de composition constante n'est pas la cause mécanique d'un processus binomial, elle en est d'un point de vue rationnel la condition structurelle. Enfin, pour finir de relativiser l'omniprésence du schéma de l'urne constante, Cournot invente avec Bienaymé et Lexis de nouveaux protocoles de tirage dans des urnes de compositions variables. Le traité de Joseph Bertrand (1889) est une nouvelle entreprise de démolition des théories de Quetelet, et bien davantage de la statistique de Laplace, Poisson et Bienaymé sur laquelle il se fonde. « *C'est avec grande défiance, dit-il, qu'il faut, sur les traces de Condorcet, éclairer les sciences morales et politiques par le flambeau de l'Algèbre…* » L'assimilation n'est pas permise entre probabilité objective et probabilité subjective. Dans les questions démographiques et morales, l'assimilation à l'urne de composition constante est abusive. La formule de la probabilité des erreurs n'est qu'un leurre, au mieux un axiome. La méthode des moindres carrés a « procuré plus d'une déception (masse de Jupiter…)». Quetelet n'a fait qu'ajouter à ce désordre en confondant homme moyen et homme idéal. « l'homme type sera donc sans passion, sans vices, ni fou ni sage, ni ignorant ni savant, souvent assoupi,…bref médiocre en tout ».

Mais la contestation vient principalement des écoles allemande et anglaise. Du côté des anglais on trouve toute une série de travaux de l'école dite fréquentiste, qui ne se contentent pas de rejeter la probabilité comme « raison de croire ». Ils relativisent les régularités statistiques trouvées par Quetelet dans de nombreuses données, en attribuant ces régularités au seul effet de les avoir observées sur des individus ou objets de même genre. L'homogénéité produit *logiquement* ces régularités. Pour John Stuart Mill, Robert Leslie Ellis et John Venn, comme d'ailleurs pour Augustin Cournot, les notions d'espèces et de genres, qui résultent de classifications rationnelles et logiques, suffisent à expliquer certaines régularités. On retrouve ici l'argument de Rümelin : selon les genres et les espèces l'homogénéité peut être absolue, forte, ou faible.

John Venn (1888), quant à lui, reproche aussi à Quetelet d'avoir assimilé l'idée de variation que l'on trouve dans les jeux (dés, cartes, urnes…), dans les mesures physiques, et dans les comportements sociaux. « *Rien n'est moins absurde que l'idée d'un type humain invariable, d'une part parce qu'aucune population vraiment homogène n'existe (une nation est un corps artificiel dont les membres sont sujets à un nombre considérable de causes perturbatrices) et d'autre part parce que toute personne qui adhère un tant soit peu à la doctrine de l'Evolution ne peut regarder le type , au sens de Quetelet, comme doté d'une réelle permanence et fixité* »

Cet argument du darwinisme joue encore dans le discours de Francis Galton. Celui-ci, cousin de Darwin, et promoteur en Angleterre d'un programme de recherche eugéniste et biomathématique qui va accoucher des deux outils les plus importants de la Statistique, la régression et la corrélation, ne manquera pas une occasion de se moquer du goût des statisticiens pour les moyennes. S'il glorifie la loi des erreurs comme « une forme d'ordre cosmique qui règne sur les plus grands chaos » il n'en fait pas le même usage. Il écrit (1908) : « *Les premiers objectifs de la loi des erreurs étaient exactement opposés, en un sens à ceux auxquels je l'appliquai. C'était pour éliminer ou réduire les erreurs. Mais ces erreurs étaient* *justement les choses que je voulais préserver et connaître* » La courbe en cloche ne lui sert donc point à révéler un quelconque type, symbole pour lui de la médiocrité, mais plutôt à *"ranger les hommes en fonction de leurs capacités naturelles, en les mettant dans des classes séparées par des degrés égaux de mérite, et de montrer le nombre relatif d'individus inclus dans les différentes classes"*. La loi se Laplace Gauss est devenue un instrument de différenciation et d'interclassement des individus selon leurs « capacities ».

La révision du credo de Quetelet s'accompagne d'un certain nombre de tests de ses résultats. William Guy ouvre les cérémonies du Jubilee de la Royal Statistical Society (1885) par un discours qui démonte point par point les soi-disant régularités de Quetelet en soumettant les données du maître à un test assez grossier, et qui aboutit à un renversement des propositions de la physique sociale : les fluctuations l'emportent largement sur les stabilités. De plus les régularités sociales les plus fortes ne sont pas celles qui relèvent de facteurs du monde physique. Le libre arbitre ne s'oppose pas au déterminisme, il se combine à une sorte d'habitus pour stabiliser les comportements.

L'actuaire français Émile Dormoy avait mené une démarche de même type dans la décennie précédente avec un test plus précis reposant sur l'idée que si des causes indépendantes produisent une loi des écarts de type binomiaux ou gaussien, alors on peut envisager réciproquement que des écarts observés qui auraient une dispersion différente de ce que donne la loi binomiale étaient le signe d'une certaine solidarité des causes, positive ou négative. Le rapport entre écart absolu moyen observé et écart espéré dans un cadre normal fournissait la statistique d'un véritable test d'homogénéité. Celui-ci fut redoutablement sévère puisque seules les statistiques sur le taux de masculinité à la naissance fournissent une valeur proche de 1, indiquant que « les écarts entre naissance des deux sexes suivent une marche normale ». Pour tous les autres cas, « l'assimilation à une urne de composition invariable n'est pas acceptable ».

Wilhelm Lexis, actuaire et professeur de statistique puis d'économie à Fribourg appartient à la fameuse école historique allemande. Sa vision d'une population est qu'elle est un système de groupes homogènes imbriqués, chacun caractérisé par un certain modèle probabiliste. Lui aussi va concevoir un test analogue à celui de Dormoy dans lequel il va comparer dispersion empirique et dispersion normale. Mais à la différence de Dormoy il dispose de modèles alternatifs de tirage dans des urnes différentes. Soit une urne de boules blanches et noires de composition fixe 50/50: la variance du nombre X de boules blanches obtenues en n tirages avec remise est n/4. Elle n'est plus que de 6n/25 si l'on tire alternativement dans deux urnes de composition 20/30 et 30/20. Elle est au contraire supérieure à n/4 si l'urne dans laquelle on tire dépend de la couleur de la boule précédente. Ces résultats, prolongeant ceux de Poisson, Bienaymé, et Cournot, définissent pour Lexis deux modèles - respectivement "hyponormal" et "hypernormal" - complémentaires du modèle binomial, dit "normal". La quantité Q^2 est le rapport entre la variance totale de la série temporelle, et sa composante aléatoire de type binomial, et devient largement supérieure à l'unité en cas d'hypernormalité. Lexis ne connaît pas la loi de sa statistique Q^2 et fixe arbitrairement un seuil à 2. Markov et Fisher établiront que c'est une loi du χ^2. Son test n'est pas indépendant du nombre des observations et surtout il présuppose l'homogénéité pour tester la stabilité.

Ladislaus von Bortkiewicz (1868-1931) a étudié à St Petersbourg puis à Göttingen où il fait sa thèse avec Lexis. Il enseigne ensuite à Strasbourg avec Georg Friedrich Knapp (1895-1899), à St Petersbourg puis à l'université de Berlin (1901) où il est professeur d'économie et statistique (1920). Dans un joli papier de 1931 intitulé « *relations entre stabilité et homogénéité* », il va reprendre et prolonger cette « analyse de variance » faite par Lexis bien avant Fisher. Bortkiewicz fait un calcul analogue en considérant que les n observations instantanées ne sont pas homogènes mais relatives à n' composantes, ou territoires, de tailles diverses. Il peut alors calculer une valeur Q_i pour chacun des n' territoires, et obtenir une variation dans le temps et dans l'espace des probabilités p_{it}. Bortkiewicz traduit l'idée de réaction positive ou négative des causes les unes sur les autres par un "*coefficient de syndromie*" qui n'est rien d'autre qu'une moyenne pondérée des coefficients de corrélations - ou plutôt de covariation - entre p_{it} et p_{jt} , les séries temporelles de *p* pour deux territoires *i* et *j*. Les valeurs-clés de ce coefficient γ définissent différents régimes de covariation moyenne: "antidromie" ($\gamma<0$) , "paradromie" ($\gamma=0$), "homodromie" ($\gamma>0$) et "isodromie" ($\gamma=1$) qui caractérisent des évolutions de l'urne des dans les différents territoires. Le cas de l'homodromie étant le plus fréquent dans les groupements naturels, compte tenu d'une certaine *solidarité des causes*. Appliquant par exemple cela au taux de suicide dans 40 états et provinces d'Allemagne (de tailles très différentes) entre 1902 et 1911, Bortkiewicz trouve qu'il y a homodromie ($\gamma=0,38$), c'est-à-dire existence de facteurs communs à la fluctuation, qui explique une valeur de Q_o supérieure pour la population totale à la moyenne de toutes les valeurs Q_i provinciales. Tandis que l'application de la méthode aux taux de mariage dans 6 villes (aussi différentes que Barcelone, Birmingham, Boston, Leipzig, Melbourne et Rome) conduit à une légère antidromie ($\gamma=-0,054$) et une relation d'ordre inverse entre Q_o et la moyenne des Q_i: la population totale n'est plus une "population naturelle".

« *Donc, conclut-il, il existe entre homogénéité et stabilité une relation antagoniste - une petite homogénéité s'accorde avec une grande stabilité. Par exemple, les provinces selon lesquelles un pays peut être divisé montreront, en moyenne, une plus grande homogénéité et en même temps une plus faible stabilité par rapport à un événement A que ne le fera le pays tout entier. A leur tour, les districts selon lesquels les provinces peuvent être divisés montreront en moyenne une plus grande homogénéité associée à une plus petite stabilité.*"

Bortkiewicz avait déjà montré dans son article de *l'Encyclopédie des sciences mathématiques* de Klein et son adaptation pour l'Encyclopédie française de Molk que « *tout le système construit par P.S. Laplace devenait illusoire dès lors qu'il est nécessaire d'admettre que les probabilités varient dans chaque série d'observations* », et il avait cherché non pas comme son prédécesseur à Strasbourg (Knapp) à se passer du calcul des probabilités, mais à en inventer un autre qui rende mieux compte des diverses situations observées. Ce travail marque l'aboutissement de ce programme, initié chez Lexis par la théorie des *Chancensystem*. La variabilité n'est plus dès lors une vibration parasite autour du type, elle est pleinement constitutive du monde qui est le nôtre. Le modèle laplacien et queteletien des erreurs n'étant plus valide, l'homme moyen étant brisé en morceaux, il faut des modèles statistiques plus complexes qui prennent en charge toute la variabilité dont le monde fait preuve.

Conclusion

Dans ce voyage au XIXe siècle autour de la notion de dispersion, nous avons visité trois grands paradigmes de cette notion. Celui de la théorie des erreurs de Laplace, Gauss et Fourier. Celui de la théorie des moyennes de Quetelet et Bertillon. Celui de la variabilité des chances de Lexis et Bortkiewicz. Chaque fois, ce qui fait problème n'est pas tant la formule choisie pour la mesure de la dispersion que la signification donnée à la propriété de variation que cette mesure révèle et quantifie. Dans le premier cas nous voyons se mettre en place une approche probabiliste de l'erreur et de sa loi qui prend en compte la seule composante accidentelle de cette erreur. Dans l'approche de Quetelet, nous assistons à une transposition de la théorie des erreurs de mesure dans la statistique sociale qui se traduit par une assimilation forcée des statistiques à des observations indépendantes d'un même objet, et des variations à des erreurs. Dans le troisième épisode, nous voyons se constituer une école allemande et une école biométique anglaise qui tiennent pour essentielle la variation et la réinterprètent en proposant sa décomposition en composantes individuelles, spatiales et temporelles.

Pour mieux ajuster notre vocabulaire à ce qui vient d'être raconté, je crois qu'il conviendrait d'une part de reprendre les dénominations d'*observations* et de *statistiques* introduites par Edgeworth pour distinguer les mesures qui ont un référent réel explicite et celles qui n'en ont point, et d'autre part de donner deux noms différents à la fluctuation des séries selon qu'elle désigne l'ensemble des déviations par rapport à une valeur qui se pose comme référent, ce qui est le cas des *observations*, ou qu'elle désigne dans le cas des *statistiques* une variation, dans le temps et/ou dans l'espace, propre à un groupe d'individus plus ou moins homogène, et dont les propriétés et comportements sont plus ou moins stables sur une certaine période. Traiter toutes les séries de nombres par une même formule – par exemple celle de l'écart-type - n'a pas plus de sens que de les traiter par une même formule de milieu – par exemple celle de la moyenne arithmétique. Dans les deux cas nous perdons le sens des choses c'est-à-dire leur signification, et nous opérons une mise en équivalence très souvent abusive qui conduit à des mauvais usages. L'assimilation des statistiques aux observations et des déviations aux variations n'est pas sans risque. On en sait quelque chose encore aujourd'hui avec les séries financières, pour lesquelles l'hypothèse d'une valeur fondamentale de référence semble après la crise tout aussi problématique que celle de l'homme moyen[11].

Bibliographie

ARMATTE M., 1989, "Modèles statistiques de l'homogénéité et de la stabilité d'une population au XIXème siècle", in *Les ménages; mélanges en l'honneur de J.Desabie*, Paris, INSEE

ARMATTE M., 1995, *Histoire du Modèle linéaire. Formes et usages en Statistique et en Économétrie jusqu'en 1945*, Thèse EHESS, sous la dir. de J. Mairesse.

ARMATTE M. et DROESBEKE J.J., 1997, Quetelet et les probabilités : le sens de la formule, in *Actualité et universalité de la pensée scientifique d'Adolphe Quetelet,* Actes du colloque des 24-25 oct. 1996, Académie royale de Belgique, Mémoire de la classe des sciences.

ARMATTE M., 1998, "Erreurs et moyennes", *Les Cahiers de Sciences et Vie,* N°48, Décembre.

ARMATTE M., 2006, Les images de la Statistique à travers ses traités, *Journal électronique d'histoire des probabilités et de la statistique*, Vol.2/N°2, déc. 2006. www.jehps.net.

BERTILLON L. A., 1876, "Moyenne", in *Dictionnaire Encyclopédique des Sciences Médicales*, A.Dechambre,II-10, Paris:Masson,296-324; et *JSSP*, 1876, p. 265-271 et 286-308.

BERTRAND J., 1971,(1889), *Calcul des probabilités*, 3ème édition, New York, Chelsea, XLIX + 317 p.; 1ère ed. 1889; 2ème ed. 1907.

BORTKIEWICZ L.von , 1931, The relation between stability and homogeneity. *Ann.Math.Stat.*, 2, 1-22.

[11] Voir la « critique de la valeur fondamentale » chez Walter et Brian 2008.

BOUGUER P., 1749, *La Figure de la terre, déterminée par les observations de MM. Bouguer et La Condamine, de l'Académie royale des sciences, envoyés par ordre du roi au Pérou pour observer aux environs de l'équateur*, Paris.

BRIAN E., 1991, Des courbes qui parlent dans un brouhaha de chiffres, *Bulletin de l'association française pour l'histoire de l'informatique, mémoire vive, Vol. 5,* http://clioweb.free.fr.

BRIAN E., 1995, Le prix Montyon de statistique à l'Académie royale des sciences pendant la Restauration, *Revue de Synthèse,* 116, 2/3, Avril 1995.

CALLENS S., 1994 (1997), *La valeur pratique du calcul des probabilités selon Emile Borel*, Thèse EHESS, publiée sous le titre : *Les maîtres de l'erreur : mesure et probabilité au 19e siècle*, Paris, PUF, 1997.

COTES R., 1722, "Aestimatio Errorum in Mixta Mathesi, per Variationes Partium Trianguli Plani et Sphaerici", in *Opera Miscellanea*, Cambridge, R. Smith

DORMOY E., 1878, *Théorie Mathématique des Assurances sur la Vie*. , Paris, Gauthier-Villars

EDGEWORTH F. Y., 1885, "Observations and statistics. An Essay on the theory of errors of observations and the first principles of statistics", *Trans. of the Cambridge Phil. Society*, 14, p. 138-169.

FOURIER J., 1826, Mémoire sur les résultats moyens déduits d'un grand nombre d'observations, in *Recherches statistiques sur la Ville de Paris et le département de la Seine,* Vol. III, et *Oeuvres*, 1890, p. 525-545

FOURIER J., 1829, Second Mémoire sur les résultats moyens et sur les erreurs de mesure, in *Recherches statistiques sur la Ville de Paris et le département de la Seine,* Vol. IV, et *Oeuvres*, 1890, p. 549-590

GAUSS C. F. 1855 (1809) , "Theoria motus...", in *Méthode des moindres carrés. Mémoire sur la combinaison des observations*, Trad. J. Bertrand , Mallet-Bachelier, Paris

GAUSS C. F. 1855 (1821-26), "Theoria Combinationis...", in *Méthode des moindres carrés. Mémoire sur la combinaison des observations*, Trad. J. Bertrand , Mallet-Bachelier, Paris; tr. angl. H.F. Trotter in *Gauss's Work (1803-1826) on the theory of Least Squares*, Princeton Univ.

GUY W., 1885, "Statistical developpment with special reference to Statistics as a Science", *Jubilee of the R.S.S.*, p. 72-86.

LAMBERT J. H., 1765, "Anmerkungen und Zusätze zur praktische Geometrie", *Beiträge zum Gebrauche der Mathematik und deren Anwendung*, Berlin.

LEXIS W., 1879, "Ueber die Theorie der Stabilität statistischer Reihen", *Jahrbuch für Nationalöfonomie und Statistik*, Vol. 32, p. 60-98.

PORTER T., 1997, Was Quetelet a Positivist?, in *Actualité et universalité de la pensée scientifique d'Adolphe Quetelet,* Actes du colloque des 24-25 oct. 1996, Académie royale de Belgique, Mémoire de la classe des sciences

QUETELET A., 1835, *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale*, Paris: Bachelier

QUETELET A., 1846, *Lettre à S.A.R. le Duc Règnant de Saxerr-Cobourg et Gotha, sur la théorie des probabilités appliquées aux sciences morales et politiques*, Bruxelles, Hayez.

RUMELIN G., 1896 (1863) , "De l'objet de la statistique" in *Problèmes d'économie politique et de statistique*, Paris, Guillaumin.

SERRES M., 1980, *Le Passage du Nord Ouest*. Hermes V, Editions de Minuit.

ULLMO J., 1969, *La pensée scientifique moderne*, Paris, Flammarion.

VENN J., 1888 (1866), *The Logic of chance*, Londres, 1ère ed. 1866, 2ème ed. 1876, 3ème éd. 1888.

WALTER C. et BRIAN E., 2008, *Critique de la valeur fondamentale,* Springer.