Logique et Connaissance Scientifique
Dirigé par J. Piaget, 1967, 1097-1113.

Épistémologie du hasard dans les sciences sociales. Invariance des lois et vérification des prédictions

Benoit Mandelbrot

Bien que les encyclopédies doivent en règle générale se limiter à l'exposé de résultats fermement acquis, nous ne pouvons éviter de consacrer ce chapitre à des questions encore très ouvertes. On ne peut en effet parler de la signification épistémologique du hasard sans se référer aux modèles de la réalité basés sur ce concept. Or, dans les sciences sociales ainsi que dans les sciences physiques à l'échelle de la terre, le domaine des modèles stochastiques est en flux si rapide que son présent deviendra déjà passé avant que cette étude soit publiée; il nous a donc semblé préférable de parler de ce qui nous paraît être une des voies de l'avenir. De plus, la prédiction étant par nécessité une chose subjective, nous espérons dans son contexte être mieux excusé de parler surtout de l'aspect que deux des problèmes centraux de l'épistémologie prennent dans le contexte des recherches que nous poursuivons actuellement.

Nos considérations épistémologiques, étant nées de l'étude de problèmes très ouverts, ont des conséquences pratiques directes. On verra en particulier pourquoi, même aux mains de l'expert le plus qualifié, certains problèmes importants de statistique sociale ne peuvent être résolus par les techniques aujourd'hui disponibles. Ce travail est donc un appel au développement de techniques plus appropriées. Certaines de nos conclusions ont paru choquer lorsque nous les avons présentées pour la première fois. Elles deviendraient moins surprenantes si toutes les qualifications requises étaient insérées, mais la nécessité d'être bref ne nous le permettra pas, et l'expression de nos thèses en prendra peut-être un « ton » plus extrême que nous ne le désirons. Pour une présentation précise de notre pensée, nous devons encore renvoyer à de nombreux articles, dont quelques-uns sont cités en bibliographie.

Il est tout d'abord bien entendu que l'objet des lois scientifiques est de décrire les influences que les variations de diverses « observables » exercent l'une sur l'autre. Mais, avant de chercher en quoi une variation contrôlée d'une quantité *A* influe sur la valeur de la variable *B,* il est évidemment indispensable d'examiner les variations dites spontanées de cette dernière, c'est-à-dire celles que l'on observe lorsque l'on n'essaie de contrôler aucune des quantités susceptibles de la modifier. Par exemple, avant d'examiner l'efficacité d'un remède, il est indispensable de recueillir des données sur le développement habituel de la maladie à laquelle ce remède s'adresse.

Il est bien connu que le concept flou de « hasard » a été en quelque sorte « domestiqué » par les mathématiciens pour fournir un bon modèle de ces activités spontanées ; il est difficile cependant d'éviter les connotations non formalisées des termes de la langue usuelle et on préfère donc dire que les activités non contrôlées sont « engendrées par des processus stochastiques ». Ceci nous permet d'énoncer notre première question: étant données les difficultés notoires de l'observation dans les sciences humaines, comment se fait-il que divers observateurs puissent jamais se mettre d'accord sur le même modèle stochastique pour représenter l'activité spontanée des systèmes auxquels ils s'intéressent?

Sous certaines conditions assez générales, nous pensons que la réponse à cette question est que seules les lois d'une certaine famille sont susceptibles d'être jamais observées, parce qu'elles sont les seules à être invariantes par rapport aux diverses transformations de la réalité que l'on rencontre lorsque l'on passe d'un observateur à un autre. Ces lois sont liées à une célèbre expression statistique dite loi de Pareto, dont l'ubiquité est d'ailleurs telle que de nombreux exemples, de ses possibilités d'application sont habituellement désignés par d'autres termes. Tel est par exemple le cas pour le problème de la récurrence dans le jeu de pile ou face, qui se présente comme suit :

Pierre et Paul jouent à pile ou face avec une pièce dont on suppose qu'elle reste indéfiniment juste (bien que nos partenaires aient commencé à jouer au début du XVIII^{e} siècle !). A chaque fois que la pièce tombe sur pile, Pierre gagne un sou, et inversement si la pièce tombe sur face. De plus, les fortunes des deux joueurs sont infinies, et leur partie ne pourra donc pas être interrompue par la ruine de l'un ou de l'autre. Cela étant, commençons donc à suivre leur partie à un moment *M,* et attendons jusqu'au premier moment *T* tel que, entre *M* et *T,* la pièce soit tombée un nombre égal de fois sur pile et sur face (nous mesurerons le temps écoulé par le nombre de fois que la pièce a été jetée). On démontre dans la plupart des livres de calcul des probabilités que l'on a la relation suivante :

« Il y a une probabilité proportionnelle à $1/\sqrt{t}$ que le nombre de lancements de la pièce entre les moments *M* et T soit supérieur à *t*. »

Pointons donc sur du papier doublement logarithmique le nombre *t* et la probabilité *Pr* [*T* — *M* > *t*], c'est-à-dire pointons log (Probabilité) et log (*t*) sur du papier ordinaire. La loi que nous avons énoncée signifie que l'on doit s'attendre à ce que notre graphique soit une ligne droite de pente égale à — 1/2. Une loi de probabilité ayant cette propriété est appelée « loi de Pareto d'exposant 1/2 ».

Plus généralement, la loi de Pareto d'exposant α est définie et caractérisée par le fait que la courbe qui la représente sur papier doublement logarithmique est une ligne droite de pente — α. On peut d'ailleurs montrer que les pointages de ce genre ne sont dignes de confiance que si a est petit, au plus égal à 3 ou 4; donc la loi de Pareto doit pratiquement être limitée à cette zone de valeurs de α.

Étant données les propriétés d'invariance de cette loi, sur lesquelles nous allons revenir, son ubiquité cesse d'être étonnante et le problème classique de son « explication » est considérablement modifié. Si telle ou telle famille de faits n'est pas représentable par une loi de probabilité de ce type, on doit en effet craindre qu'elle ne soit pas représentable par quelque loi générale que ce soit, mais apparaisse sous des formes différentes à divers observateurs. Les lois invariantes constituent dans ce chaos des oasis de régularité qui ne pouvaient manquer d'attirer spécialement l'attention. On peut s'estimer heureux qu'elles recouvrent tant de problèmes pratiquement importants.

La deuxième question que nous allons nous poser est une suite immédiate du problème des variations spontanées d'une variable *A.* Supposons que celles-ci sont parétiennes, donc susceptibles de réunir l'unanimité des observateurs, et revenons au problème central de la dépendance entre les variations d'une autre quantité *B* et celles de *A.* Par exemple, le docteur *D* trouve que le remède *T* améliore visiblement l'état d'une bonne proportion de ses patients souffrant de la maladie *M.* L'amélioration est cependant rarement immédiate et le nombre total des cas examinés est malheureusement petit; ne se pourrait-il donc pas que l'opinion du docteur *D* soit une illusion, en ce sens qu'il devait de toute façon s'attendre à une amélioration spontanée de l'état d'une bonne partie de ses malades ? La question se complique évidemment si on examine de plus près la signification du terme « spontané » : il se pourrait par exemple que le « hasard » qui

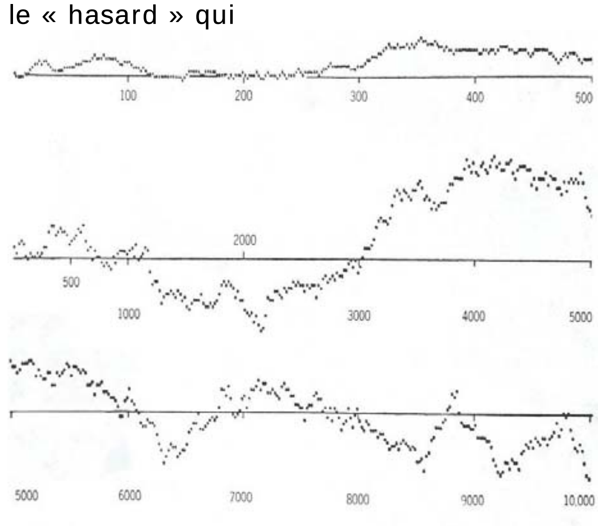

Figure 1 *Exemple de la variation dans el temps des gains de Pierre après qu'il a jeté un jeton* T – M *fois. Horizontalement, valeur de* T - M *(sur les lignes 2 et 3, nous avons seulement indiqué les valeurs multiples de 20 ; sur la ligne I, nous avons indiqué les valeurs multiples de 2 afin de donner plus de détails) ; verticalement, valeur du gain. Ce graphique est extrait du traité Probability Theory, de W. Feller, New York, 1957, avec l'autorisation de Tauteur et de l'éditeur.*

conduit à l'amélioration de certains malades soit l'absorption du remède *T* dans une nourriture prise dans d'autres buts. Mais, avec ou sans telles complications, le problème de la signification statistique d'une relation entre variables *A* et *B* est une des questions centrales de la science, et on trouve malheureusement que cette signification statistique ne pourra être mesurée dans le cas parétien que lorsque des outils statistiques nouveaux auront été développés.

En fait, le caractère essentiel du hasard parétien est de reposer de façon plus réaliste la vieille question de la signification des « valeurs typiques » des populations statistiques. En d'autres termes, l'effort qui a été dépensé pour inventer et enseigner le concept d' « homme moyen » a si bien réussi que l'on a fini par dépasser les limites de validité de cette idée, oubliant que les lois qui caractérisent l'activité spontanée de nombreux systèmes intéressants sont si extraordinairement « dispersées » qu'aucune « valeur typique » n'est représentative. Il est même fréquent d'observer que quelques résultats extrêmes dominent l'étude d'un phénomène, et qu'il importe par contre peu de savoir que huit résultats sur dix se trouvent être très voisins d'un nombre qui est à la fois une moyenne, une médiane et une valeur la plus probable. Pour d'autres phénomènes, la moyenne et la valeur la plus probable diffèrent totalement; la valeur probable peut même être infinie; tel est le cas pour le temps de retour au point de départ dans le jeu de pile ou face ! Tout ceci aurait pu inquiéter, car il semble clair qu'un concept de valeur typique n'a d'importance pratique que dans la mesure où l'on obtient des résultats approximatifs mais raisonnables en l'utilisant pour remplacer une variable aléatoire. On sait même qu'il est toujours possible de décomposer une quantité aléatoire gaussienne en deux parties : sa moyenne, qui ferait l'objet d'un premier stade d'approximation de la théorie, et sa variance, qui ferait l'objet d'un deuxième stade. Mais tel n'est pas en général le cas en sciences sociales. En pratique, on réussit souvent à se contenter de descriptions basées sur des valeurs typiques; on y arrive en décidant de considérer que les très grandes valeurs de la variable en question doivent être décrites séparément, seules les variables moyennes devant faire l'objet de théorie statistique. Malheureusement, la séparation entre le très grand et le moyen est — dans les cas qui vont nous concerner — nécessairement vague ou artificielle; jusqu'à nouvel ordre, on doit considérer que les modèles aléatoires ont pour responsabilité de donner une idée des grandes valeurs tout autant que des petites.

Empressons-nous de donner des exemples variés de ces divers phénomènes. Commençons par l'économique, en examinant la variation temporelle du prix d'une denrée telle que le coton brut. Ce fut autrefois l'objet d'une spéculation active, et son cours variait fortement au moins une fois par jour, et souvent des centaines de fois. Du point de vue du petit spéculateur, sans contrôle sur le monde extérieur et sans bonne information à son sujet, les changements du prix ont le caractère de cette « activité spontanée » dont nous nous occupons ici; conformément à l'idée que s'en fait le public, les cours changent en général très peu de jour en jour, mais de temps en temps ils bougent considérablement et très soudainement, et il est clair que les valeurs moyennes des petites fluctuations ne disent rien au sujet des rares grands changements qui font les grandes fortunes et les grandes ruines. Il semble bien cependant que l'on puisse décrire ces grands sauts comme étant engendrés par le même mécanisme aléatoire que les changements « ordinaires » de prix. La découverte de ce mécanisme est toute récente (voir la bibliographie de nos travaux); mais il paraît désormais bien établi qu'il existe et qu'il est parétien, donc invariant par rapport aux diverses transformations qui sont associées à des changements du « point de vue » d'observation.

Quant au problème de la vérification des prédictions dans le contexte des prix, inutile de dire qu'il est aussi vieux que la profession de conseiller boursier.

Puisque nous en sommes au chapitre de la spéculation, il n'est pas besoin de parler de choses aussi compliquées que le prix du coton : même pour le problème de pile ou face, on rencontre des opinions rationnellement indéracinables quant à l'existence de déviations systématiques, donc prévisibles et exploitables, par rapport à l'hypothèse d'une activité spontanée régie par le hasard pur. Emile Borel a consacré de nombreuses pages à ce sujet, et ses remarques restent très actuelles; on peut même dire que leur importance s'est accrue, car on ne pouvait se douter vers

1900-1930 que des difficultés similaires se rencontrent si souvent dans la construction des modèles stochastiques du réel.

Au lieu d'examiner les applications classiques de ce jeu, nous allons immédiatement le retraduire en termes de la psychologie du Rat. Supposons qu'il soit lâché à l'origine d'un plan infini, quadrillé au moyen de droites équidistantes et parallèles à l'un ou l'autre des deux axes de coordonnées, et supposons que de la nourriture soit placée à l'origine (juste après le départ du Rat) ainsi qu'à tous les points pour lesquels x et y sont égaux au même nombre entier positif ou négatif. Le « problème » que le Rat doit résoudre est simplement de se nourrir.

Supposons qu'il l'attaque en décidant absolument au hasard à chaque carrefour dans quelle direction il doit continuer. Comme il a quatre choix possibles, sa décision sera équivalente au résultat de deux tirages au sort successifs d'un jeton; deux de ses choix le rapprocheront de la nourriture et deux l'en éloigneront, et l'on peut admettre que ce sera le premier des tirages au sort qui en décidera, le deuxième pouvant être négligé ici comme étant sans effet sur la distance entre la nourriture et le Rat. Ce dernier aura par suite des chances égales de s'approcher et de s'éloigner de la nourriture, ce qui nous ramène visiblement au problème classique du retour à l'origine dans le jeu de pile ou face. Le Rat se nourrira donc dès que le nombre de « pile » (gains de Pierre) égalera le nombre de « face » (gains de Paul) et il y a une probabilité 1/2 qu'il réussisse en deux pas. Mais supposons qu'il commence par « tomber mal » et par s'éloigner beaucoup de la droite $x = y$; la chance le fera-t-elle revenir à la nourriture ? La réponse est : « Oui, mais... ». Mais nous savons déjà que la probabilité que le temps mis à résoudre le problème soit supérieur à t est proportionnelle à $1/\sqrt{t}$, cette décroissance très lente exprimant la loi de Pareto d'exposant 1/2. En pratique, ceci signifie qu'un certain nombre de Rats vont atteindre les limites physiques de l'appareil expérimental sans se nourrir; si l'appareil avait été infini, ils seraient épuisés avant d'arriver au but.

Il est cependant intéressant de considérer le cas hypothétique où l'appareil est de dimension infinie et où les rats aussi bien que l'observateur sont infatigables. Supposons que l'on fasse cent séries de cent expériences chacune, toutes de la forme envisagée ci-dessus, et calculons le temps moyen pour chaque série d'expériences. On trouverait que ces temps moyens varieraient de série à série de façon affolante, le fait de prendre la moyenne ne régularisant rien.

Ceci est une conséquence du fait suivant : si l'appareil était infini et le Rat infatigable, la valeur probable de la durée d'expérience aurait été infinie. Il en résulterait aussi que le Rat le plus malchanceux occuperait le psychologue pendant la moitié environ de la durée d'une série d'essais, et il est difficile de concevoir que le psychologue ne « censure » pas des sujets qui s'éloignent de toute norme de façon en apparence si aberrante. Et pourtant, tous les sujets dont nous parlons suivent exactement le même mécanisme, que les aberrations sont susceptibles d'éclairer. (C'est par exemple ce qui nous est arrivé lors de l'étude de la récurrence des erreurs de communication digitale par téléphone : les données expérimentales mises à notre disposition conduisaient pour ce problème à une loi de Pareto incontestable, *sauf que* le nombre de très longs intervalles entre erreurs était trop petit; renseignements pris, il se trouva précisément que les trop longs intervalles avaient été censurés volontairement par notre source comme étant « sûrement sans signification » !)

L'expérience précédente sur le Rat est de celles dont on conçoit volontiers qu'elles soient faciles à « simuler » mathématiquement au moyen d'une machine dite « à calculer ». Pour cela, notre Rat (dont la décision était déterminée par un tirage au sort) est remplacé par un *stat-rat* entièrement réduit à un mécanisme aléatoire dépourvu de chair mortelle. Pour de tels « animaux », il semblerait que les expériences puissent être poursuivies sans les limites imposées par la fatigue de la matière vivante. Malheureusement, les limites imposées par la dimension des machines à calculer resteront toujours finies. La difficulté sera repoussée de plus en plus mais jamais éliminée.

Les détails du problème que nous avons décrit se trouvent importer peu. La nourriture pourrait par exemple être disposée le long de deux axes de coordonnées; il y aurait alors un peu moins de Rats très aberrants, puisque le nombre de ceux mettant plus de t pas à se nourrir décroîtrait comme $1/t$ au lieu de $1/\sqrt{t}$, c'est-à-dire plus vite. Mais les difficultés de principe seraient semblables. De même, il importe peu que le problème à résoudre soit relatif à un Rat, à une machine à décoder des messages perturbés par des bruits, ou à un chercheur espérant résoudre tout autre problème compliqué, en répétant autant de fois qu'il le faut certaines opérations plus élémentaires. L'essentiel est que l'on ait affaire à une forme, ou à une autre, de cheminement dans un « labyrinthe » infini.

On trouve de plus dans beaucoup de cas potentiellement très intéressants, qu'il y a une probabilité non nulle que le problème ne puisse Jamais être résolu par la chance seule, quelle que soit la dimension des appareils de simulation et la résistance des participants.

Que dire alors des labyrinthes bien plus complexes encore que l'on rencontre lorsque les efforts du Rat ou de l'ingénieur sont remplacés par les fondions supérieures de l'intelligence ? L'étude psychologique expérimentale de ces problèmes est comme on le sait très difficile, et la technique de simulation mécanique de ces fonctions n'a pas encore dépassé les balbutiements de ce que l'on appelle 1' « intelligence artificielle ». Quel que soit le problème que l'on se pose, on doit *a priori* s'attendre à ce que la réalité et la simulation donnent toutes deux des résultats si extraordinairement variables de cas à cas, que avant de se prononcer — même provisoirement — on devra attendre que le nombre de bonnes expériences et de bonnes simulations dépasse de plusieurs ordres de grandeur ce qu'il est aujourd'hui. Il ne suffit pas qu'une personne dont on respecte le jugement dise « qu'il lui semble » que la simulation est réaliste.

De la psychologie, passons à la sociologie des groupes. Négligeant leur structure, c'est leur dimension qui joue le rôle que la longueur d'une expérience psychologique — négligeant ses étapes intermédiaires — joue dans l'exemple précédent. Pour avoir un exemple probant, nous devons immédiatement considérer des groupes très complexes, tels que ceux définis par les populations des villes d'un pays. Une fois de plus, on constate qu'il n'y a rien de tel qu'une « ville moyenne » : les petites villes sont très nombreuses, les grandes villes sont très, très grandes et les villes ayant la population moyenne ne sont en rien représentatives.

Passons maintenant à un problème de caractère tout différent, qui se pose dans l'étude statistique du discours (nous en parlons plus en détail dans le volume de cette encyclopédie consacré à la *linguistique).* La dispersion extrême des valeurs d'une variable se traduit ici comme suit : Sur un long échantillon donné de discours, un bon tiers des mots appartient à quelques formes spécialement fréquentes, et un autre tiers appartient à des formes que l'on ne rencontre que très peu de fois. Il ne reste donc qu'un tiers pour les « mots-formes de fréquence moyenne », et il est évident qu'on déformerait totalement la structure du discours en les prenant pour représentatifs du tout. Cependant, la question des déviations entre le discours et une suite purement aléatoire de mots ne pose pas ici de problème, en ce sens qu'il n'y a aucun doute au sujet de l'existence de lois telles que celles de la grammaire. C'est dû au fait évident que le grammairien expérimentateur peut conduire un nombre illimité d'expériences linguistiques sur le sujet dont il étudie la langue maternelle; cette circonstance unique peut même faire espérer que les lois de la linguistique soient établies d'une façon qui, à la limite, deviendrait, cent pour cent certaine. Nous avons voulu insister sur cette possibilité pour bien montrer que les difficultés épistémologiques auxquelles ce texte est consacré ne sont aucunement sans solution.

Les exemples que nous avons jusqu'ici discutés se rattachent à des sciences sociales (respectivement l'économique, la psychologie, la sociologie et la linguistique). Mais on ne doit pas croire que l'étude du comportement soit caractérisée par des moyennes dépourvues de sens ou d'utilité, tandis que les sciences physiques seraient caractérisées par des distributions statistiques dans lesquelles tous les faits sont très proches d'une valeur typique,

Par exemple, la distribution des dimensions des nappes d'eau naturelles est très voisine de celle des populations des villes. Les petites nappes sont en effet très nombreuses et les grandes nappes sont très très grandes. Ce fait n'est pas sans influence sur notre langue, la conduisant à représenter la réalité par les trois termes distincts de « lac », « mer » et « océan », impliquant qu'entre nappes d'eau les différences de « quantité » de taille sont telles qu'elles en deviennent des différences de « qualité ». Des remarques analogues existent pour les surfaces des bassins des cours d'eau, conduisant ainsi aux concepts qualitatifs de fleuve, rivière et ruisseau.

Les remarques que nous venons de faire au sujet de la terminologie peuvent d'ailleurs suggérer que le savant contemporain n'est pas le seul à craindre les distributions très dispersées et à chercher à éviter d'avoir affaire à la fois à des objets de tailles par trop disparates : il en était sans doute de même des Anciens dont la science plus ou moins formalisée est à la racine du vocabulaire de la langue courante.

Nous avons assez vu d'exemples pour avoir une idée de l'extraordinaire dispersion des distributions associées aux valeurs naturelles, ou « spontanées », des variables de la statistique sociale, et des limites de la notion d' « homme moyen ». Mais il ne suffit pas de démontrer l'impossibilité d'application de la loi gaussienne habituelle de distribution; il reste surtout à la remplacer, et il reste d'abord à démontrer qu'une loi (quelle que soit sa forme) est susceptible de s'appliquer, étant donné les difficultés notoires de l'observation dans les sciences sociales !

Il « se trouve » heureusement que les lois de Pareto, très dispersées par rapport à leurs valeurs typiques, sont formellement invariantes par rapport à diverses transformations qui jouent un rôle central dans la théorie de l'observation.

Tout d'abord, elles sont invariantes par rapport à l'opération suivante de « mélange ». Supposons — ce qui n'est que réaliste — que l'on ne puisse pas exactement reconstituer les conditions d'une expérience, et que l'on ne puisse obtenir des échantillons de bonne taille qu'en mélangeant des résultats d'origines diverses. La valeur de la variable observée sera dans ce cas déterminée en deux stades, le premier consistant à déterminer par une loterie quelle sera la loterie du deuxième ordre dont on observera les résultats finals. Fort heureusement, sous des conditions assez larges, si chaque expérience conduit à une loi de Pareto pour les résultats, la loi de Pareto s'appliquera aussi au mélange. Par exemple, les données sociologiques ci-dessus sont un mélange de données relatives à des « villes » définies différemment en différentes parties d'un pays ; mais ce n'est pas grave du tout. De même, « le » prix est défini pour des qualités variables d'une denrée; ce n'est pas grave du point de vue de la théorie (pour la pratique, c'est tout autre chose !)

Deuxièmement, la loi de Pareto est invariante par rapport à la considération du maximum. Supposons en effet qu'au lieu de posséder les résultats de toutes les expériences d'une longue série on ne connaisse que les expériences hors de l'ordinaire. Une telle méthode est bien entendu peu satisfaisante en théorie, mais il est quelquefois inévitable que le folklore d'un domaine ne retienne que les résultats extraordinaires. Cela n'est pas le cas pour les expériences hypothétiques sur les Rats dont nous avons parlé : le danger, dans ce cas, est que l'on ne tienne pas assez compte des expériences conclues par l'échec. De même, nous avons vu que les ingénieurs omettent quelquefois des faits en apparence vraiment trop aberrants. Mais l'histoire ne se réduit-elle pas souvent à une suite de bons rois régnant sur des sujets repus par les récoltes, et de mauvais rois régnant sur des famines ? Il se trouve heureusement (sous des conditions encore assez larges) que si la distribution des données suit la loi de Pareto, le cas sera le même pour la distribution des données exceptionnellement grandes.

Troisièmement, la loi de Pareto est invariante par rapport à la considération des « agrégats ». Remplaçons, en effet, une expérience simple par une suite d'expériences devant être réalisées dans un ordre donné pour atteindre le but recherché. Il se trouve encore que, sous des conditions assez larges, si les durées des expériences partielles suivent la loi de Pareto, il en sera de même pour la durée de l'expérience totale.

« Agrégation » est d'ailleurs synonyme de combinaison linéaire ou purement additive. Comme la mathématique n'offre guère d'autre assortiment complet d'outils bien rodés, les premiers modèles que l'on essaye dans tout domaine nouveau sont en général linéaires; la loi de Pareto doit jouer un rôle central dans ces modèles de première approximation.

Réciproquement, il se trouve qu'il n'y a aucune autre famille de lois qui soient invariantes par rapport à nos trois transformations. Il y a, cependant, quelques exemples isolés de lois invariantes seulement par rapport à la maximalisation ou par rapport à l'agrégation. Par exemple, si la loi classique de Gauss s'applique à des parties, leur somme sera également gaussienne; par suite, si l'on ne considère que l'agrégation, il est possible, malgré la transformation des données, de constater qu'un « ordre » non parétien s'applique; ce fait a été exploité à fond dans les modèles statistiques classiques.

Un cas particulièrement important est celui des limites des agrégats d'un grand nombre de parties. On connaît l'opinion habituelle selon laquelle la somme d'un grand nombre de parties suit nécessairement la loi de Gauss, à condition que ces parties ne soient ni « trop » différentes ni trop interdépendantes. Les théorèmes qui expriment ces idées sont même habituellement invoqués comme « explication » lorsque la loi de Gauss est observée empiriquement. Bien sûr, il faut éventuellement étudier l'interdépendance entre les parties d'un « système » social. Par exemple, il serait trop naïf aujourd'hui d'affirmer que les opérations psychologiques complexes sont des simples accumulations de « briques » dont chacune est une boucle « stimulus-réponse » : ce sont des édifices bien plus complexes, et ce que l'on veut surtout découvrir, ce sont les plans de leur « architecte ». Il est légitime cependant de commencer par examiner ce qui arriverait si ces plans étaient sans effet sur les statistiques, et on s'attend donc à trouver que les distributions des agrégats soient gaussiennes et qu'en plus, la dimension relative de toute partie d'un agrégat soit toujours très petite. C'est une généralisation de l'idée selon laquelle les erreurs d'observation sont gaussiennes « parce qu'elles sont faites de la somme » d'un grand nombre de contributions dont chacune est petite en valeur absolue.

En réalité, la loi de Gauss n'est pas la seule qui puisse s'appliquer à des agrégats additifs, et l'on se limite donc de façon excessive si l'on approche le problème des modèles statistiques avec les seuls outils gaussiens. Si c'est la loi de Pareto qui s'applique, on doit s'attendre à des « phénomènes de concentration ». Si un agrégat est formé de la somme d'un grand nombre de « parties », ayant a priori la même distribution, il se peut que la partie la plus grande de cet agrégat ne soit pas une portion négligeable du tout (comme nous l'avons vu dans le cas où nous comparions le temps pris par l'animal le plus malchanceux et la durée totale d'une série d'expériences sur les Rats). Il s'ensuit que, si l'on considère une loi non gaussienne, ou que l'on observe que la plus grande des parties n'est pas négligeable par rapport au tout, on ne doit pas se hâter d'invoquer des arguments non statistiques ou d'autres complications.

Passons maintenant à un autre problème, celui de la confirmation d'une loi de prédiction affirmant l'existence de déviations systématiques par rapport à une « activité spontanée » parétienne. Comme base de notre discussion, nous garderons l'idée que l'ultime confirmation d'une loi consiste dans le fait que ses prédictions sont systématiquement vérifiées par les faits observés. De ce point de vue, une activité spontanée parétienne pose des difficultés plus grandes encore qu'une activité spontanée gaussienne, car on y perçoit des « formes » si intéressantes d'apparence et si riches de traits, que l'on a grand peine à ne pas y « lire » des structures.

Tout ceci nous amène à nous poser dans le contexte des modèles statistiques une question que les mécaniciens n'ont résolue que vers 1900, grâce au concept de « problème bien posé » dû à Jacques Hadamard. Pour qu'une science soit utilisable en tant que système de prédictions, il est en effet évidemment nécessaire que des petits changements dans les données initiales conduisent à des petits changements dans les prédictions — sauf peut-être en un petit nombre de cas dits critiques. Hadamard a montré que, de ce point de vue, certains problèmes de mécanique sont simplement sans réponse : quelle que soit leur importance pratique apparente, ils avaient été « mal posés ». A l'aide de ce concept, on peut résumer notre discussion en disant que bien des problèmes qui sont exactement posés dans un univers gaussien ne le sont plus dans un univers statistique parétien.

Pour terminer cette discussion trop schématique, nous voudrions dire qu'elle est l'aboutissement, quelque peu inattendu pour l'auteur, de la charge que nous avions reçue d'écrire un article sur les « modèles quasi physiques en sciences sociales », en motivant bien la combinaison entre « pareil » et « différent » que recouvre le terme « quasi ». Tandis que nous continuions par ailleurs nos travaux de recherche, le poids des différences ne cessait malheureusement d'augmenter, et la rédaction de ce texte ne cessait d'être reportée. Nous trouvions en effet de plus en plus de raisons de croire que des progrès décisifs peuvent être espérés d'une application systématique aux sciences sociales des méthodes de construction des modèles de la physique; nous constations cependant chaque jour un peu plus qu'en adaptant ces méthodes à leur nouveau contexte, nous aboutissions à des résultats de forme extrêmement différente. Cette tendance ne cessa que le jour où nous commençâmes à nous convaincre que les lois statistiques qui caractérisent la physique à l'échelle du laboratoire (ou à l'échelle plus petite) cessent d'être applicables à plus grande échelle. Il était bien sûr raisonnable de supposer — au moins provisoirement — que ce sont ces derniers phénomènes qui influencent le plus les inégalités observées dans les sciences sociales. Par ce détour compliqué et inattendu, nous en sommes revenu à insister sur les similitudes que recouvre le quasi » de notre titre primitif. Mais tout ceci n'est que très très provisoire !

Bibliographie

Deux de nos articles d'exposition, dont un ancien et un récent, développent des thèmes voisins de ceux que nous exprimons ici, mais en les étendant dans d'autres directions. Voyez donc

Benoît Mandelbrot, *Les Lois statistiques macroscopiques du comportement,* « Psychologie Française », t. III, octobre 1958.

Benoît Mandelbrot, *New Methods in Statistical Economics,* « The Journal of Political Economy », t. 71, octobre 1963.

Voir aussi notre contribution à la *Linguistique* dirigée par André Martinet, « Encyclopédie de la Pléiade ».

La loi de Pareto fut énoncée par cet auteur dans le contexte des revenus personnels, en 1897. Un grand nombre d'exemples ont été réunis dans l'ouvrage suivant, dont il est recommandé de ne pas lire le texte mais d'examiner les figures et la liste des références :

George Kingsley Zipf, *Human Behavior and the Principle of the Least Effort,* Cambridge, Mass., 1949 (épuisé).

L'étude des phénomènes parétiens exige des outils mathématiques qui ont été inventés — sans but pratique apparent — par Paul Lévy, qui fut entre 1920 et 1959 professeur d'analyse à l'Ecole Polytechnique. L'univers personnel de Lévy — qui a désormais cessé d'être imaginaire — est, en particulier, décrit dans deux ouvrages classiques :

Paul Lévy, *Calcul des probabilités,* Paris, 1925.

Paul Lévy, *Théorie de l'addition des variables aléatoires,* Paris, 1937, 2e édition 1954.

Pour les arguments techniques qui ont conduit aux thèses du présent travail, le lecteur pourra glaner dans nos propres travaux, dont voici quelques échantillons :

Benoît, *The Pareto-Levy Law and the Distribution of income,* « International Economic Review », t. 1, pp. 79-106, mai 1960, et t. 4, pp. 111-115, 1963.

Benoît, *The Variation of certain speculative Prices,* « Journal of Business of the University of Chicago », t. 36, pp. 394-419, octobre 1963. Réimprimé dans l'anthologie. *The Random Character of Stock Market Prices,* (réunie par Paul H. Cootner), pp. 307-332, Cambridge, Mass., 1964.

Benoît, *Self-similar Error-Clusters and the Concept of conditional Stationarity,* « I.E.E.E. Transactions on Communications Technology », t. C.O.M.-I3, pp. 71-90, juin 1965.

Benoît, *Very long-tailed Probability Distributions and the empirical 'Distribution of City Sizes',* dans « Mathematical Explorations in Behavioral Science » (réuni par F. Massarik et P. Ratoosh), pp. 322-332, Homewood, 111, 1965.

Benoît, « On the Theory of Word Frequencies and on Related Markovian Models of Discourse », *Proceedings of the American Mathematical Symposium on the Structure of Language and its Mathematical Aspects* (réuni par R. Jakobson), pp. 190-219, Providence, R.L, 1961.