

Maurice HALBWACHS (1923)

**« L'expérimentation
statistique et les
probabilités. »**

Un document produit en version numérique par Jean-Marie Tremblay,
professeur de sociologie au Cégep de Chicoutimi

Courriel: jmt_sociologue@videotron.ca

Site web: <http://pages.infinit.net/sociojmt>

Dans le cadre de la collection: "Les classiques des sciences sociales"

Site web: http://www.uqac.quebec.ca/zone30/Classiques_des_sciences_sociales/index.html

Une collection développée en collaboration avec la Bibliothèque
Paul-Émile-Boulet de l'Université du Québec à Chicoutimi

Site web: <http://bibliotheque.uqac.quebec.ca/index.htm>

Cette édition électronique a été réalisée par Jean-Marie Tremblay, professeur de sociologie au Cégep de Chicoutimi à partir de :

Maurice Halbwachs (1923)

« L'expérimentation statistique et les probabilités. »

Une édition électronique réalisée de l'article « L'expérimentation statistique et les probabilités ». Paris : *Revue philosophique*, 96, 1923, pages 340 à 371.

Polices de caractères utilisée :

Pour le texte: Times, 12 points.

Pour les citations : Times 10 points.

Pour les notes de bas de page : Times, 10 points.

Édition électronique réalisée avec le traitement de textes Microsoft Word 2001 pour Macintosh.

Mise en page sur papier format
LETTRE (US letter), 8.5'' x 11''

Édition complétée le 5 juin 2002 à Chicoutimi, Québec.



Maurice Halbwachs

« L'expérimentation statistique et les probabilités »

(1923) *

Entre le statisticien, penché sur des colonnes de chiffres et des graphiques, et le physicien en train de monter et de suivre une expérience, y a-t-il vraiment une différence essentielle ? On le croit communément, et on serait tenté de ranger le premier dans la catégorie de ceux qui n'entrent en contact avec les faits qu'à travers les livres ou annuaires où ceux-ci sont enregistrés, et, surtout, à lui refuser tout pouvoir d'agir sur ces faits, et de les modifier, ou de les simplifier. Volontiers, on le comparerait à un historien, avec cette différence qu'on n'a plus prise sur les faits historiques parce qu'ils sont passés, tandis que le statisticien n'a pas prise sur les faits relevés dans les statistiques parce qu'ils sont trop compliqués ou trop nombreux, ou l'un et l'autre.

Soit à chercher quelle influence l'âge exerce sur la mortalité dans un groupe. Si l'organisme était un composé chimique relativement simple, on pourrait procéder comme le chimiste qui soumet un composé connu à des réactions qui

* Extrait de la *Revue philosophique*, 96, Paris, 1923.

le décomposent. Mais quels sont les éléments d'un organisme, quels sont leurs rapports, quelles actions s'exercent sur lui, à chaque âge ? Problème infiniment complexe, non seulement parce qu'il s'agit d'un organisme, mais encore parce que tous les membres d'une espèce à un même âge ne sont pas identiques. On doit se borner à observer ce qui se passe, c'est-à-dire, sans exercer une action quelconque sur le groupe, à compter la proportion des hommes d'un même âge qui y meurent chaque année. Si cette proportion est constante, on arrive bien à un résultat, mais par une toute autre méthode, en apparence, que l'expérimentateur.

Cependant, est-ce bien une autre méthode ? M. Simiand n'en croit rien ¹. Il a mis en regard quelques exemples d'expérimentation physique et d'opération statistique, en vue de montrer que, dans les deux cas, on arrive aux mêmes résultats par des procédés très semblables. Examinons d'un peu près ces exemples.

« Voilà une série de données mensuelles, pendant un certain nombre d'années, sur le taux de chômage d'un certain ensemble ouvrier. La variation, telle quelle, de ces données apparaît, au premier examen, assez complexe et mêlant probablement une variation à période annuelle, selon le mois ou les saisons, et une variation à période plus longue, tendance à une hausse ou tendance à une baisse à travers plusieurs années. Par des procédés statistiques appropriés, nous éliminons, d'une part, la variation interannuelle, de façon à dégager ou isoler la variation intérieure à l'année ou variation saisonnière proprement dite, d'autre part, cette variation saisonnière, pour dégager et isoler la variation à période plus longue. Et, cela fait, nous étudions la relation que chacune de ces variations peut respectivement soutenir avec tel ou tel facteur. » En quoi procédons-nous autrement qu'un physicien qui décompose un phénomène en ses composants pour examiner l'action, sur chacun d'eux, de chacun des facteurs ?

Bien entendu, tout dépend ici du problème qu'on pose. On peut se demander, en effet, quelle est la perspective de chômage ou bien pour le groupe, ou pour un membre du groupe. Pour le groupe, on obtient un chiffre ou des chiffres qui mesurent exactement la fréquence relative des chômeurs dans telle ou telle condition : résultat aussi positif que ceux auxquels on arrive en n'importe quelle science. Pour un membre du groupe, il n'en est plus de même : rapportant le nombre des chômeurs au nombre d'ouvriers du groupe, on peut calculer ce qu'on appellera les chances de chômage pour un ouvrier quelconque. Ce n'est plus un résultat positif, en ce sens qu'on ne sait ni si tel ouvrier défini chômera, ni s'il ne chômera point. C'est pour cette raison que certains logiciens ont cru devoir distinguer la probabilité et l'induction, celle-ci seule fondant des lois, tandis que sur celle-là on ne peut fonder que des prévisions plus ou moins vraisemblables ².

Or, en physique, une loi permet de prévoir à coup sûr que tel corps particulier, tel « individu physique » se comportera, ou non, de telle manière. On ne nous dit pas que, sur cent barres de fer, tant se dilateront de telle longueur,

¹ F. Simiand, *Statistique et expérience. Remarques de méthode*, Rivière, 1922.

² Voir J. Venn, *Logic of chance*, 3e édition, 1888, p. 203 et suiv. et 265 et suivantes.

sans qu'on sache d'ailleurs lesquelles : la loi de dilatation s'applique à chacune des barres aussi bien qu'à leur groupe.

Nous verrons plus tard s'il y a lieu de parler de probabilité, dans le cas d'un groupe d'ouvriers exposés au chômage, ou d'un groupe de faits relevant de l'observation statistique, au même sens que dans le cas d'une série de faits qui se produisent, comme on dit, au hasard ¹. Il se pourrait que telle notion, comme la probabilité de vivre jusqu'à tel âge, n'ait été introduite par exemple dans la théorie des assurances que pour des raisons de commodité, et qu'elle suppose qu'on substitue, à un groupe naturel donné d'hommes de tel âge, un groupe fictif d'individus supposés identiques, ou même un seul type moyen. Mais tenons-nous-en à cette constatation : il y a des groupes tels qu'on peut établir avec certitude que tel phénomène, atteignant tant d'individus, s'y produira, sans qu'on soit en mesure de dire quels seront ces individus. Mais n'en est-il pas de même, d'une quantité d'objets physiques, qui possèdent des parties, et sont composés d'une quantité de molécules distinctes ? Serait-il légitime de demander au physicien en quel endroit précisément se rompra un fil métallique soumis à une forte tension, à quel point d'un corps sphérique se produira une décharge électrique, et comment se distribuera la température dans les plus petites parties d'un corps ? Cette incertitude dans le détail diminue-t-elle en quoi que ce soit la valeur des lois de l'élasticité des corps, de l'électricité et de la chaleur ? Et, pourtant, la détermination du détail n'offrirait pas moins d'intérêt dans bien des cas en physique que, pour l'homme, la prévision de la durée pour chaque individu, et, pour l'ouvrier, la prévision des certitudes de chômage en ce qui le concerne ².

Ceci posé, toute la différence entre l'exemple des statistiques du chômage indiqué par M. Simiand, et telle expérience de physique où l'on étudie la variation du volume d'un corps à l'état gazeux, dans des conditions de température et de pression différentes, consiste en ce que, dans le second cas, il semble qu'on élimine effectivement l'action d'un facteur, pour étudier celle d'un autre (en admettant qu'il n'y en ait que deux à envisager ³), tandis que,

¹ Sur les notions de fréquence, de hasard et de probabilité, cf. Fréchet et Halbwachs, *Le calcul des probabilités à la portée de tous*, Paris, Dunod, 1923.

² M. Borel a posé très nettement un problème de ce genre, dans un article « Radioactivité, probabilité et déterminisme », paru dans *la Revue du mois*, le 10 janvier 1920. Il indique l'intérêt qu'il y aurait pour nous à nous emparer de l'énergie contenue dans les atomes et par conséquent à savoir par quelle cause s'explique leur décomposition dans les corps radioactifs. Soit une masse infinitésimale de radium, renfermant par exemple 45 millions d'atomes. Des résultats expérimentaux et des considérations de probabilité nous font admettre que 45 environ se décomposeront par jour. « La question que l'on pose est alors : les 45 atomes qui doivent se décomposer dans la journée sont-ils dès maintenant déterminés ?... Y a-t-il des caractères extrinsèques par lesquels ils puissent être actuellement distingués ?... Si oui, la vie de chaque atome radioactif a actuellement une durée bien déterminée... » Mais on peut faire une seconde hypothèse, celle du non-vieillessement des atomes (et il y a, dit M. Borel, des raisons très fortes pour la préférer). Alors, la probabilité de décomposition est la même pour tous les atomes : les circonstances fortuites déterminent seules l'explosion.

³ En réalité on n'élimine pas un facteur, mais sa variation : en langage mathématique, on étudie la variation de la fonction par rapport à celle d'une seule des variables. Mais les autres variables n'en entrent pas moins dans la fonction avec la valeur fixe qu'on leur assigne. Par exemple, dans la formule : $p_v = p_v + (1 + a t)$, t (la température) ne varie pas : mais sa valeur fixe n'en conditionne pas moins celle de p_v , c'est-à-dire le rapport entre la pression et le volume.

dans le premier cas, on réussit à en faire abstraction (sans d'ailleurs la supprimer matériellement) « par des procédés statistiques ». Voici, par exemple, pour l'étude du chômage, comment nous procéderons. Pour éliminer l'élément de variation interannuelle, dans une industrie où nous savons d'ailleurs, que le caractère saisonnier de l'industrie ne change pas d'une année à l'autre, nous calculerons la moyenne du chômage pour tous les mêmes mois d'une période de plusieurs années successives (écartant au besoin les années où nous avons des raisons de croire qu'une perturbation annuelle a troublé le jeu des causes saisonnières), et nous comparerons ces moyennes, supposant qu'elles se rapportent aux mois d'une année fictive soustraite aux variations qui se produisent d'une année à l'autre. Pour éliminer l'élément de variation saisonnier, nous calculerons la moyenne du chômage de chaque année comprise dans la même période (écartant au besoin les mois où l'action saisonnière a pu troubler, pensons-nous, le jeu des causes annuelles), et nous comparerons ces moyennes, supposant que chacune d'elles se rapporte à une année de douze mois fictifs soustraits aux variations qui se produisent d'un mois à l'autre.

Certes il y a une différence entre l'élimination effective de telle espèce de variations qu'on réalise en physique, et l'opération qui, par un calcul de moyennes, permet d'étudier à part et sous forme abstraite deux aspects matériellement inséparables d'un même ensemble. Par moyennes des variations saisonnières, par moyennes des variations annuelles, nous représenterons des ensembles de faits répartis, les premiers sur une série d'années où s'exercent les variations annuelles, les autres sur une série de mois où s'exercent les variations saisonnières ; si nous les rapportons aux mois d'une année soustraite aux variations annuelles, à une année soustraite aux variations saisonnières, c'est à une année et à des mois fictifs, que d'aucune façon il ne nous est possible de réaliser. Comment comparerait-on les chiffres moyens du chômage aux mois de janvier, février, etc. (pour une période de dix ans), aux chiffres représentant le volume d'une masse d'un gaz à la température de 1°, 2°, etc. (sous une pression constante), alors que chacun des premiers chiffres représente une série de faits (par exemple, chômage en janvier à chacune des dix années) résultant à la fois de deux facteurs (cause de variation saisonnière *combinée avec* une part au moins d'une cause d'une variation annuelle), tandis que chacun des seconds chiffres représente un fait (volume de la masse de gaz) résultant d'un seul facteur (action de la température sans action de la pression).

Mais cette différence est peut-être toute relative. Après tout, nous ne savons point de quel jeu d'actions moléculaires ce que nous appelons la pression et la température résultent. « Dans la théorie cinétique des gaz, dit Poincaré, on envisage des molécules animées de grandes vitesses, dont les trajectoires, déformées par des chocs incessants, ont les formes les plus capricieuses et sillonnent l'espace dans tous les sens. Le résultat observé est la loi simple de Mariotte ; chaque fait individuel était compliqué ; la loi des grands nombres a rétabli la simplicité dans la moyenne. » Si nous supposons des physiciens à l'échelle de ces molécules, il est probable, en effet, que, pour dégager les lois de ces mouvements, il leur faudra appliquer la méthode statistique : ce qui, à nos physiciens, apparaît sous la forme de facteurs ou de forces réalisées tout entières dans un seul phénomène simple, se traduira pour ces observateurs microscopiques comme la moyenne d'un nombre considérable de variations

particulières¹. Et il se pourrait d'ailleurs qu'au terme de leurs calculs ils obtiennent les mêmes résultats que les physiciens terrestres au terme de leur expérience.

Il est vrai qu'ils n'auraient pas monté l'expérience eux-mêmes. Mais cela est-il bien une condition nécessaire de la recherche expérimentale ? En effet, on conçoit que nos physiciens, sans faire d'expériences, se fussent bornés à noter, au fur et à mesure des changements naturels de la température et de la pression, les changements du volume par exemple d'un gaz enfermé dans un ballon de substance élastique, fabriqué d'ailleurs pour des fins non scientifiques, et que, plusieurs fois, l'observation eût été faite par hasard à une température identique : d'un tableau où ces observations eussent été rangées suivant l'ordre de grandeur des volumes à même température et à pression variable, ou inversement, et par des calculs appropriés, n'eussent-ils pas dégagé, plus lentement, mais exactement sur les mêmes données et par le même genre de raisonnement, les mêmes résultats ?

M. Simiand rappelle qu'« il se présente certains cas - l'histoire des diverses sciences en témoigne - où, sans action de l'homme, par le seul concours de circonstances appropriées, se trouve réalisée une simplification suffisante pour permettre au savant d'apercevoir une relation » : cas d'expérience naturelle ou spontanée. Mais inversement, dans le domaine social en particulier, il n'est pas rare que l'enregistrement des faits statistiques suppose une action humaine sur la matière d'où sont extraites les données. Bien entendu on ne peut assimiler une telle intervention à celle de l'expérimentateur. Les lois qui prescrivent un recensement, ou une élection, ou des tarifs douaniers, ou la tenue de registres de l'état civil, de registres d'érou, etc., sont elles-mêmes des faits sociaux - elles font partie de la nature sociale, et répondent en général à des besoins non scientifiques. Par elles cependant prennent corps et consistance des catégories et des distinctions dont le savant peut tirer parti, bien qu'il ne les ait pas proposées lui-même, et que souvent elles ne soient pas telles qu'il lui eût le mieux convenu. Tout se passe comme si un physicien se servait pour ses recherches d'instruments industriels, on observait les opérations mises en train par quelque ingénieur ou quelque technicien : il y apprendrait tout de même plus que s'il observait simplement les faits naturels. Le travail du statisticien offre souvent un tel caractère. Il trouve avantage à étudier la matière sociale, aux moments où certains organes de l'administration, de l'État, ou qui représentent à quelque titre la collectivité, lui imposent de se distribuer dans certains cadres. Et, en même temps, il doit s'accommoder des lacunes, des confusions, des obscurités de classifications faites par d'autres que lui. Ainsi, pour établir les statistiques criminelles, on extrait les criminels et délinquants de groupes où ils sont confondus parmi d'autres hommes, on les réunit par catégories, on leur applique une épithète : du fait qu'on les isole des autres hommes, et qu'on les incorpore à un nouveau groupe où ils se retrouvent avec leurs semblables, on renforce et on met mieux au jour les traits propres de leur nature et de leur rôle dans la société. C'est ainsi

¹ On trouvera, dans le livre de M. Borel *sur Le hasard*, tout un chapitre consacré aux applications du calcul des probabilités en sciences physiques, p. 157 et suiv. Il est remarquable que la méthode statistique intervienne surtout dans l'étude des phénomènes d'ordre atomique, et dans l'étude de la mécanique céleste, c'est-à-dire dans deux domaines de la science où les corps, en raison de leur grandeur et de leur éloignement ou de leur petitesse, échappent à notre action, sinon à nos sens et à nos instruments d'observation.

qu'un chimiste extrait de minéraux telles substances, à peine perceptibles dans leur masse, mais, qui, prélevées sur des masses étendues, et rapprochées, offrent les caractères d'un corps nouveau, qui se prête à l'observation et à l'expérience. Seulement, comme la matière délinquante ou criminelle d'une société subit l'élaboration nécessaire à la juger et à la punir, à l'éliminer ou à l'isoler pour empêcher qu'elle ne nuise, et non pas à l'étudier à seule fin de la connaître, les statistiques criminelles souffrent de bien des insuffisances.

Si les données du chômage n'étaient pas recueillies d'abord par mois, par industrie, par pays, si on ne distinguait pas les chômeurs des vagabonds, des infirmes, etc., on n'aurait en effet aucune prise sur une masse de faits mouvante et confuse. Le travail qui correspond ensuite au calcul de moyennes consiste alors à recomposer sur un autre plan, sur plusieurs autres plans, l'ensemble dont les statistiques nous présentent les parties : opération intellectuelle, et non physique, comme dit M. Simiand, puisque nous réfléchissons sur des rapports, c'est-à-dire sur des abstractions. Est-ce là encore un trait qui distingue radicalement la statistique de la science expérimentale ? Mais qu'est-ce, par exemple, en physique, que l'accélération, sinon un rapport abstrait entre des vitesses, et où voit-on qu'il se réalise en dehors des termes d'où il est tiré ? Au reste, puisque, dans les moyennes, et dans les expressions statistiques, sont contenus les éléments extraits de la réalité, puisqu'une moyenne n'est qu'un point de vue sur une série de faits donnés, c'est bien sur les choses qu'opère l'esprit : puisque tout le contact du savant avec les choses se réduit à ce qu'il en observe, il suffit que nous ayons modifié l'ordre de nos observations, que nous les envisagions d'un nouveau point de perspective ; cela équivaut à changer l'ordre dans lequel nous passons en revue les parties de la réalité qui nous apparaît : à cela revient le rangement des faits en vue du calcul d'une moyenne ; mais il n'y a guère plus dans une expérience.

M. Simiand indique un autre exemple « où les deux processus se rapprochent encore davantage. Voici un ensemble d'opérations : semis de certaines plantes, fécondation des fleurs dans de certaines conditions, choix et semis de graines nouvelles, nouveau semis, nouvelle récolte, observations sur certains caractères de ces diverses générations de plantes, qui, par une élaboration appropriée, aboutissent à une des théories dites mendéliennes. Voici, d'autre part, un ensemble d'opérations sur diverses générations d'hommes ou d'animaux : observations sur les tailles ou autres caractères somatiques de ces diverses générations, traitement statistique de ces constatations pour en dégager des résultats simplifiés de certaine façon, qui, par une élaboration appropriée, aboutit à une des théories dites galtoniennes, Quelle différence essentielle y a-t-il entre les deux ensembles d'opérations initiales qui permettent à l'esprit humain d'aboutir à de certaines relations ? »

Ici, le premier ensemble d'opérations (semis, etc.) constitue une expérience, le second (observations sur les tailles, etc.), une opération statistique. En quoi consiste l'analogie ? Ce n'est pas, évidemment, en ce que, dans l'un et l'autre cas, on aboutit à des théories qui visent à expliquer le mécanisme ou à établir l'action de l'hérédité. C'est en ce que l'on compare les caractères de diverses générations. Examinons cependant d'un peu plus près comment se fait la comparaison dans les deux cas, et, pour préciser, rappelons en détail deux applications de ces méthodes.

I° Expérimentation. - Voici comment Mendel établit ce qu'il a appelé la loi de ségrégation des « gènes » ¹.

Il opéra par exemple sur des variétés du pois de jardin commun (*pisum sativum*) : certains plants sont longs, d'autres courts, certains produisent des pois verts, d'autres des pois jaunes, etc. Tenons-nous en à la taille des plants. Il mit du pollen provenant d'un plant long sur le stigmate d'un plant court (dont il avait enlevé au préalable les étamines et le pollen). Appelons P_1 ces premiers plants. Les plants hybrides, F_1 , qui sortirent de ces graines, étaient tous longs. Il les laissa se féconder eux-mêmes, recueillit leurs graines, et les sema : certaines produisirent des plants longs, d'autres des plants courts. Il compta ces plants F_1 : il trouva qu'ils étaient dans la proportion de 3 produisant des plants longs à 1 produisant des plants courts. Il laissa ensuite les nouveaux plants, F_2 , se féconder eux-mêmes : les plants courts produisirent seulement des plants courts. Quant aux plants longs, un tiers d'entre eux produisirent seulement des plants longs ; les deux autres tiers produisirent des plants longs et des plants courts ; il compta ces plants F_2 (ces deux derniers tiers) et il trouva qu'ils étaient dans la proportion de 3 produisant des plants longs à 1 produisant des plants courts, comme les hybrides de la première génération F_1 . Donc l'ensemble des plants F_2 comprenait des plants purs courts, des plants hybrides et des plants purs longs, dans la proportion : 1, 2, 1.

Voici maintenant par quelle hypothèse Mendel explique ces résultats. Il doit y avoir certains facteurs dans les plants longs originels P_1 , qui font que les plants de cette race sont toujours longs, et certains facteurs, dans les plants courts originels P_1 , qui font que les plants de cette race sont toujours courts. Appelons S le premier de ces facteurs, s le second. Quand on croise les deux plants, l'œuf fertilisé doit contenir les deux facteurs = Ss , et, puisque les hybrides qui en proviennent sont longs, S doit être dominant. Si maintenant les deux facteurs présents dans l'hybride F_1 se séparent quand les ovules et les grains de pollen se forment, la moitié des œufs contiendront le facteur S , et la moitié le facteur s . Et il en est de même de la moitié des grains de pollen, et de l'autre. Dès lors un simple calcul des probabilités permet de prévoir que ces œufs et ces grains de pollen, conjugués au hasard, donneront pour la génération P_2 , 1 SS (plants longs), 2 Ss (plants hybrides) et 1 ss (plants courts).

Dans cet ensemble d'opérations il y a lieu de distinguer ce qui est intervention matérielle de l'opérateur, calcul statistique, et raisonnement de probabilité. Le calcul statistique consiste, à chaque génération nouvelle, à compter le nombre des individus qui présentent ou ne présentent pas tel caractère, et à grouper les résultats obtenus en tableaux. Le raisonnement de probabilité consiste à comparer les chiffres observés, et les chiffres calculés dans l'hypothèse où ne joueraient que les lois du hasard ².

¹ Morgan, *The physical basis of heredity*, Philadelphia, 1919.

² Pour 7 paires de caractères étudiés par Mendel (1959 expériences), le rapport des caractères dominants aux récessifs était de 2,996 à 1004 (au lieu de 3 à 1). Pour la transmission de la couleur des pois de jardin (203 500 expériences), le rapport était de 3,004 à 0,996, soit une erreur probable 0,0026.

L'expérimentation consiste d'abord à choisir un plant long et un plant court, et à faire que le plant court soit fécondé par le plant long au lieu de se féconder lui-même (par ablation préalable des étamines et du pollen du plant court), ensuite à isoler les plants hybrides F_1 ainsi obtenus de façon à ce qu'ils se fécondent eux-mêmes, enfin à isoler encore les plants F_2 ainsi obtenus, de façon à ce qu'eux aussi se fécondent eux-mêmes. Le calcul statistique n'offrirait aucun intérêt, s'il portait sur des générations de plants qui se sont reproduits sans être ainsi isolés, et le raisonnement de probabilité n'interviendrait plus.

2' *Opération statistique.* Galton¹ considère 78 familles qui ne comprennent pas moins, chacune, de 6 frères et Sœurs, pour lesquelles il possède des données sur la couleur des yeux pendant trois générations (grands parents, parents, enfants). Il calcule successivement la proportion des yeux clairs pour les enfants de grands parents à yeux clairs ; pour les enfants de parents à yeux clairs ; pour les petits-enfants de grands parents à yeux clairs, en distinguant les cas où les parents eux-mêmes ont, ou n'ont pas les yeux clairs. Il trouve les chiffres suivants : 70,2 % ; 82,7 % ; 78 % (respectivement 86,4 et 58,3 %), tandis que les mêmes proportions, pour les enfants à yeux clairs de grands parents, et de parents qui n'ont pas les yeux clairs, et pour les petits-enfants, etc., sont : 13501 44,9 pour cent ; 54,2 % ; 60,3 % (respectivement 72,6 et 50,3 %). De ce calcul et de ces comparaisons il conclut qu'il y a transmission héréditaire de la couleur des yeux non seulement des parents aux enfants, mais des grands parents aux petits-enfants.

Dans cet exemple, comme dans le précédent, il y a des calculs statistiques et des raisonnements de probabilité. Dans un groupe défini de cas, pour plusieurs générations successives, on compte ceux où un caractère se présente, et ne se présente pas, et on établit ainsi un tableau de nombres. D'autre part, on admet que, s'il n'y avait pas transmission héréditaire, ces caractères se distribueraient au hasard, c'est-à-dire qu'ils ne se présenteraient pas plus souvent chez les enfants quand les parents ou les grands parents les possèdent que quand il ne les possèdent pas : l'écart constaté entre les chiffres calculés dans cette hypothèse et les chiffres observés permet d'affirmer qu'il n'en est rien. Mais il n'y a pas intervention matérielle de l'opérateur : les familles observées n'ont été soumises à aucune action artificielle de sa part : il les observe après coup, comme un historien décrit des faits passés.

Il ne s'ensuit pas que, dans cet exemple aussi, il n'y ait pas trace de procédé expérimental. On trouvera même, en recherchant quelles conditions supportent les observations et les calculs de Galton et quelles conditions lui permettent de saisir la relation qui l'intéresse, qu'elles ne diffèrent en rien d'essentiel de celles que Mendel dut réaliser artificiellement. Supposons en effet que Galton eût dû poursuivre ses recherches sur l'hérédité dans une société qui ne connaîtrait pas le mariage monogamique, dans une société telle que ces tribus australiennes que nous décrivont Spencer et Gillen où un groupe défini d'hommes a, légalement, des relations maritales avec un groupe défini de femmes, mais où ces groupes sont d'ailleurs assez étendus, il se fût trouvé dans le

¹ F. Galton, *Natural Inheritance*, 1889.

même embarras que Mendel si celui-ci n'eut pas réussi à isoler entièrement les plants qu'il étudiait. S'il a pu retenir un certain nombre de couples, et suivre leurs descendants pendant deux générations, c'est que ces couples forment au sein de la société autant de groupes fermés, c'est que les naissances des enfants s'expliquent par ces couples et par eux seuls. En d'autres termes l'isolement que Mendel réalise pour les plants un moyen d'opérations artificielles existe pour les couples humains dans nos sociétés, en vertu des lois du mariage. Or ces lois, bien que les savants n'y aient aucune part, n'en témoignent pas moins d'une intervention humaine qui modifie matériellement l'objet observé¹. On peut même dire, à la différence d'autres expériences mises en train spontanément par la société, et dont nous parlons ci-dessus (en particulier à propos des statistiques criminelles), que, dans le cas de l'expérience galtonienne, la société a fait exactement ce que le savant se fût proposé de faire pour rendre possible ses comparaisons et ses calculs.

Quel est, nous demanderons-nous maintenant, le caractère des opérations que nous venons de décrire ? Des éléments que nous y distinguons, expérience (au sens étroit d'intervention humaine dans la nature), calcul statistique et probabilité, lequel l'emporte dans chacune : est-ce, dans la première, l'expérience, dans la seconde, le calcul statistique, si bien qu'en dépit d'une ressemblance dans la forme, la première serait, seule, une « observation provoquée », où, en effet, l'observation précède les faits qui prennent corps en elle, et la seconde une « observation après coup », qui prend corps dans les faits déjà donnés ? En réalité, dans les deux cas, ne s'agit-il pas de résoudre un de ces problèmes qu'on appelle de « probabilité a posteriori », et qui reviennent, étant donnés des tirages effectués dans une urne, à retrouver la composition de l'urne ? Lorsque Mendel, en effet, trouve 1 plant long, 2 plants hybrides, 1 plant court, comme proportion caractéristique dans les plants de la troisième génération, c'est comme si, tirant d'une urne 2 boules à la fois, il trouvait que la proportion des tirages est : une fois 2 blanches et une fois 2 noires pour deux fois une blanche et une noire. D'où il conclut que l'urne contenait autant de blanches que de noires, que les facteurs « long » et « court » sont en nombre égaux dans les plants de la première génération. Il est vrai que l'hypothèse a pu venir d'abord, hypothèse qu'il y a dans l'œuf et le pollen des facteurs de ce genre, qui se « ségrègent » lors de la maturation. En tout cas, on ne pouvait la vérifier qu'en s'arrangeant de façon à ce que le contenu de l'urne, ici le nombre originel des facteurs, ne fût pas modifié durant toutes les observations : c'est dans cette mesure seulement qu'est intervenu l'expérimentateur.

¹ Il n'est pas sans intérêt d'indiquer que Galton, dans ses recherches sur la transmission héréditaire de caractères physiques tels que la taille, s'est servi de trois ensembles de données : en premier lieu de *Records of Family Faculties*, des témoignages (se rapportant au passé) sur les caractères de familles, obtenus de correspondants offrant des garanties sérieuses, et récompensés par voie de concours ; en second lieu d'observations « spéciales », qui ne portaient que sur les frères d'une même famille, effectuées à la demande de Galton ou dans son laboratoire ; enfin d'« expériences » sur des pois de senteur, choisis parce qu'ils ne se croisent pas, mais se fécondent eux-mêmes et eux seuls, condition assez rare dans le monde des plantes. Bien qu'il réserve le nom d'« experiments » à ce dernier cas (parce qu'il fait lui-même les semis, dans certaines conditions), on voit que l'avantage principal de telles observations (l'auto-fécondation) est commun aux autres, puisque les familles observées sont « extraites » d'un groupe plus large, et que les naissances qui s'y produisent résultent de couples aussi individualisés que les plants.

D'autre part, quand Galton calcule la proportion d'enfants issus de parents aux yeux clairs, de grands parents aux yeux clairs qui ont eux-mêmes les yeux clairs, il ne sait rien encore de l'influence de l'hérédité et des facteurs qui déterminent la couleur des yeux non plus que de leur rapport dans chaque groupe : il ne connaît pas le contenu des urnes. De ce qu'il trouve une plus grande proportion d'enfants aux yeux clairs, dans le groupe considéré, il conclut que les facteurs correspondants étaient plus nombreux que dans un groupe formé au hasard et, comme ce groupe est caractérisé par les yeux clairs des ascendants, que cette caractéristique tend à se transmettre par hérédité. Mais ici encore, ce qui lui permet de conclure ainsi, c'est que, grâce à l'organisation matrimoniale de nos sociétés, il a pu rattacher à un groupe d'ascendants fixe et fermé toutes les naissances, de même qu'on reporte à des urnes de contenu fixe tous les tirages.

Ainsi, dans ces deux exemples, la condition expérimentale n'est ni plus, ni moins nécessaire. Sans elle, nous n'aurions aucun moyen d'établir nos calculs statistiques, et de vérifier notre hypothèse. Si, au terme de nos calculs, nous obtenons des expressions qui représentent à nos yeux des groupes, ou les caractères de certains groupes, c'est qu'au point de départ de nos observations il y a une sélection préalable des faits, qu'elle résulte d'une opération matérielle par laquelle nous isolons certains d'entre eux, ou que nous profitons des divisions introduites dans la réalité par la société qui isole des autres groupes, par exemple, les familles, ou toutes les catégories d'hommes qu'elle recense séparément.

Mais ceci nous conduit à distinguer, avec M. Simiand, des comptages qui offrent un intérêt statistique, tous les autres. Critiquant ceux qui définissent la statistique : « un moyen d'étudier quantitativement des phénomènes qui se présentent comme pluralités de masses de cas susceptibles de varier sans règle assignable en toute rigueur », M. Simiand remarque qu'il ne suffit pas de compter un grand nombre d'unités ou de cas pour obtenir une statistique. Un calcul de ce genre, dit-il, n'offre en effet d'intérêt pour le savant que « s'il s'applique à quelque ensemble, à quelque groupe ayant une certaine consistance, ou soupçonné d'avoir une certaine consistance, en tant qu'ensemble, en tant que groupe ». Et il cite quelques exemples : « c'est le professeur Benin qui le remarque : le kilométrage d'une station de chemin de fer à toutes les stations du réseau n'est pas un fait statistique ; le nombre de fois où un certain jour de la semaine se rencontre au cours d'un mois n'est pas un fait statistique ». Cependant, n'y a-t-il pas tout un ensemble de recherches qui portent, précisément, sur cet ordre de faits dont on dit qu'ils se produisent sans règle fixe, et au hasard, et n'y emploie-t-on pas des méthodes de calcul qu'on appelle communément statistiques ? Supposons qu'on jette au hasard, sur une surface limitée par des droites, un certain nombre de pièces, et qu'on calcule la distance du centre de ces pièces à toutes ces droites : n'aura-t-on pas une série de chiffres qui expriment la façon dont elles se répartissent, quand on les jette ainsi au hasard, et ne sera-t-il point possible de montrer que cette répartition est conforme aux prévisions de la théorie des probabilités ? Ne pourra-t-on calculer de même, et pour le même objet, le nombre de fois que se présente,

dans la série des chiffres décimaux du nombre π , chacun des 9 premiers nombres, et ne démontrera-t-on pas que, par exemple, chacun des 6 premiers nombres se présente autant de fois que si, jetant un dé, on avait noté le nombre de fois qu'il amène chacun des six points ? Chacune de ces séries ne constitue-t-elle point, suivant l'expression de M. Simiand, « un groupe qui offre une certaine consistance », puisqu'on peut la représenter graphiquement par une courbe très régulière (courbe binomiale) ?

Nous sommes amenés ainsi à examiner dans quel rapport se trouvent les recherches statistiques et la théorie des probabilités. Malgré que M. Simiand ne parle point de probabilités, au cours de toute son étude, il nous semble bien qu'une des conséquences les plus certaines, et d'ailleurs les plus importantes de sa définition de la statistique est de distinguer nettement l'objet de ces deux disciplines.

Il envisage deux sortes de séries numériques, et montre, sur des exemples, en quoi les unes sont des statistiques, et les autres, non. Voici ces exemples, qui s'opposent deux à deux.

1° La statistique du mouvement général des prix est bien une statistique, parce que « derrière les prix individuels, seule réalité observable, il y a quelque chose qui, bien que ne se réalisant seulement ni pleinement en aucun d'eux, cependant... est bien une réalité ». Mais l'ensemble des observations faites sur un certain nombre d'étoiles en vue de déterminer le mouvement propre du système solaire, ne constitue pas une statistique, parce que ce mouvement propre est « un fait réalisé matériellement comme tel et qu'un observateur autrement placé pourrait constater par des moyens physiques », et sans doute par une seule observation. Ces exemples sont tirés du livre de M. Bowley ¹.

Examinons-les d'un peu plus près. « Comme le soleil et la terre, dit M. Bowley, se dirigent vers un point éloigné, situé dans la constellation d'Hercule, les étoiles ont un mouvement apparent, qui résulte du mouvement (non perçu) de l'observateur ; celles qui se trouvent dans la région de l'espace vers laquelle il s'avance semblent s'éloigner l'une de l'autre, celles qui se trouvent dans la région d'où il s'éloigne semble se rapprocher l'une de l'autre, et celles qui se trouvent sur les directions perpendiculaires à la ligne du mouvement semblent se déplacer en sens inverse de celui-ci. En même temps toutes ces étoiles ont leur mouvement propre, aussi rapide que celui du soleil, mais vers autant de directions différentes qu'il y a d'étoiles. Il y a donc une tendance générale à se mouvoir dans les directions déterminées par le mouvement du soleil, mais qui disparaît entièrement dans chaque cas particulier (pour chaque observation du mouvement d'une étoile). De même quand un changement dans la monnaie (dans son pouvoir d'achat) exerce une influence générale sur les prix, cette influence est cachée par les mouvements dus à des causes qui affectent seulement certaines des marchandises. Dans les deux cas il sera possible de découvrir la tendance générale si on dispose d'observations précises suffisantes [sans doute : en nombre suffisant] ». Ainsi M. Bowley distingue, qu'il s'agisse des observations sur le mouvement des prix ou sur le mouvement des étoiles, deux sortes de causes : l'une, générale, qui agit sur

¹ *Elements of statistics, 2e édition, 1902, p. 218.*

tous ces mouvements dans le même sens, les autres spéciales, dont certaines agissent dans un sens, certaines dans l'autre. « Si les effets séparés des causes spéciales sont petits par rapport à leur nombre, ils tendront à se neutraliser », et ce qui apparaîtra dans la moyenne, c'est le changement dû à la cause générale.

M. Simiand n'accepte pas cette assimilation. Voici, nous semble-t-il, pourquoi. Dans le cas du mouvement des étoiles, la cause générale c'est le mouvement propre du système solaire, qui apparaît dans la moyenne : or, elle est « réalisée matériellement » en dehors des observations particulières, des mouvements apparents des étoiles : laissant de côté la question de la relativité des mouvements, on conçoit qu'un observateur, placé sur une terre qui ne participerait pas au mouvement qui porte le système solaire vers la constellation d'Hercule, pourrait le reconnaître et le mesurer directement. Au contraire, la cause exprimée par la moyenne des mouvements des prix n'est réalisée que dans l'ensemble de ces mouvements, n'a point de réalité propre hors d'eux, et ne peut se saisir que dans l'ensemble des observations particulières. Un économiste qui détournerait son attention des prix particuliers n'aurait aucun moyen concevable d'observer, par exemple, le changement du pouvoir d'achat de la monnaie, puisqu'il n'apparaît que dans ces prix.

2° « La moyenne des observations de la densité d'un certain corps peut bien être obtenue par une opération mathématique identique à celle qui dégage, par exemple, d'un certain nombre d'observations sur des individus l'indice céphalique d'une race ; mais le caractère de ces deux données n'est-il pas tout différent ? » La première, en effet, (comme, ci-dessus, le mouvement propre du soleil) pourrait être établie, dans certaines conditions concevables, « par une observation directe et unique », mais non la seconde, qui n'a pas une existence physique ou matérielle hors de l'ensemble des individus. La détermination de la densité d'un corps résulte d'une série d'observations faites par des individus différents, ou par le même individu à des moments différents. Les chiffres obtenus diffèrent très peu les uns des autres, même si l'on a écarté toute cause d'erreur systématique ou si on en a éliminé l'action par un calcul, parce qu'on commet, lorsqu'on mesure une grandeur physique, de petites erreurs accidentelles. Or la théorie des probabilités prévoit que, puisque ces erreurs se produisent par hasard, elles consisteront aussi souvent à dépasser la mesure qu'à rester en deçà, et que leur nombre sera d'autant plus élevé qu'elles seront moindres en valeur absolue. L'expérience confirme d'ailleurs cette prévision. Mais il apparaît alors que, si on calcule la moyenne de toutes ces mesures, les mesures fautives par excès ou par défaut se neutraliseront : la moyenne exprimera les seules mesures qui ne se neutralisent pas, c'est-à-dire qui ne sont pas fautives, c'est-à-dire enfin la mesure exacte de la grandeur, telle que l'eût obtenue directement un physicien disposant d'instruments plus précis que les nôtres, et servi par des sens aussi parfaits que ses instruments. Cet exemple d'une série d'observations qui n'est pas une statistique correspond au précédent : on peut, en effet, considérer les étoiles comme autant d'instruments dont nous nous servons pour mesurer le mouvement propre du soleil, mais des instruments qui bougent : d'où autant d'erreurs accidentelles très petites, et qui, quand on a un grand nombre d'observations, se neutralisent.

On pourrait, il est vrai, faire à M. Simiand l'objection suivante. S'il distingue ces deux sortes de séries, série d'observations d'une grandeur physique,

série de prix, c'est que, dans le premier cas, il existe objectivement, en dehors des observations et des observateurs, une *chose*, « matériellement réalisée », et qu'on pourrait atteindre *directement* dans certaines conditions concevables. Mais c'est une vue toute théorique. En fait, de cette grandeur, nous ne pouvons connaître que ce que nous en représentent nos observations. Et comment sommes-nous assurés, d'autre part, qu'en dehors des prix il n'existe aucune réalité objective, dont ceux-ci ne sont que des appréciations ? Nous plaçant alors sur le seul terrain des faits, c'est-à-dire des données observables, nous ne relevons, entre les deux sortes de séries, qu'une différence : c'est que les unes se groupent régulièrement, en deçà et au-delà d'une valeur centrale, conformément aux prévisions du calcul des probabilités, et que les autres n'offrent pas la même régularité. Mais il existe, entre ces deux types de répartition, l'un régulier, l'autre irrégulier, bien des intermédiaires. Et, surtout, on a trouvé, dans le domaine biologique, des caractères humains tels que leur mesure, effectuée dans un grand nombre de cas, donne des résultats qui se répartissent autour d'une grandeur ou d'une mesure centrale avec la même régularité que les erreurs d'observation. Pourquoi n'admettrions-nous pas, dès lors, que les irrégularités relevées dans les autres séries s'expliquent par les conditions imparfaites de notre observation, ou parce que les faits eux-mêmes ou les tendances dont ils résultent se heurtent à des obstacles qui entravent le jeu des forces naturelles ? Il n'en reste pas moins que ces forces naturelles et générales sont les causes véritables, que nous devons essayer d'atteindre, et qui expliquent seules ce qu'il subsiste de régularité dans les effets.

On a signalé en effet un certain nombre de mesures qui paraissent ne pouvoir s'expliquer que par les probabilités ou la loi des erreurs. Le statisticien Quételet avait été très frappé de trouver une extraordinaire ressemblance entre la courbe des erreurs d'observation et la courbe qui exprime la répartition des tailles. Voici comment M. Borel résume sa thèse : « On peut dire que les tailles mesurées satisfont aux mêmes lois que les erreurs de mesure ; tout se passe comme si un même homme, dont la taille serait égale à la moyenne, avait été mesuré un grand nombre de fois par des observateurs assez maladroits ou ne disposant que d'instruments de mesure très imparfaits ¹. »

Il ne paraît pas douteux que, lorsqu'on mesure les tailles des conscrits d'un pays, et qu'on calcule leur moyenne, on fait de la statistique, exactement comme lorsqu'on calcule l'indice céphalique moyen d'une race. Mais lorsqu'on enregistre un grand nombre d'observations d'une même grandeur physique, et qu'on calcule leur moyenne, d'après M. Simiand, on ne fait pas de statistique. Or, Quételet a entièrement assimilé la moyenne obtenue dans le premier cas à la moyenne obtenue dans le second : « en prenant une moyenne, dit-il, on peut avoir en vue deux objets bien différents. On peut chercher à déterminer « un nombre qui existe véritablement » (c'est bien le cas de la densité d'un corps, visé par M. Simiand), ou bien à calculer un nombre qui donne l'idée la plus approchée de plusieurs quantités différentes, homogènes, mais inégales. On se trouve dans le premier cas lorsqu'on mesure la hauteur d'un édifice vingt fois de suite - c'est là calculer une véritable moyenne. On est dans le second cas, quand on calcule la hauteur moyenne des maisons qui se trouvent dans une

¹ *Le hasard*, p. 139. Voir aussi notre livre : *La théorie de l'homme moyen. Essai sur Quételet et la statistique morale*, Paris, Alcan, 1913, p. 28 et suiv.

rue déterminée : c'est une moyenne arithmétique (n'est-ce pas le cas visé par M. Simiand, lorsqu'il parle de ces comptages qui n'ont pas une valeur statistique ?). Entre les divers nombres qui donnent la moyenne proprement dite, il existe une continuité et, dans leur série, une régularité qui ne se découvre pas dans ceux d'où on tire la moyenne arithmétique (ce sont bien les caractères que M. Simiand désigne quand il parle de la consistance du groupe). Or le type de l'homme moyen s'exprime dans une moyenne véritable¹ ».

La ressemblance entre la courbe des tailles et courbe des erreurs d'observations n'est, cependant, pas une identité. Déjà Herschel disait à ce propos : « L'auteur assurément s'avance trop. L'erreur probable de la nature (comparée à un homme qui ferait les tailles d'après un modèle fixe) est à peu près une demi-fois plus grande que celle qui a été prise ici pour terme de comparaison (l'erreur probable d'observation), et il est clairement au-delà des bornes de toute négligence ou imperfection admissible dans la pratique de commettre des erreurs telles que les écarts extrêmes enregistrés, dans une série de pareils mesurages, quelque multipliés qu'on les suppose, ou même atteignant la moitié de ces nombres. » M. Borel exagère certainement, lorsqu'il dit que « tout se passe comme si, à un jeu convenablement choisi, une série de parties devait amener un nombre de parties gagnées égal au nombre de centimètres qui exprime la taille moyenne » et que « le tableau sur lequel on inscrirait le nombre de parties gagnées dans chaque série serait exactement comparable au tableau sur lequel seraient inscrites les tailles² ». Considérons quelques chiffres obtenus par Quételet il a essayé de vérifier par l'expérience les règles du calcul des probabilités : sur 4 096 tirages, il aurait dû extraire (d'une urne contenant un nombre égal de noires et de blanches), 2 048 blanches ; il en a tiré 2 066. Soit un écart de 18 entre le nombre calculé et le nombre observé. Il a d'autre part mesuré les tailles de 20 à 25 000 hommes, et il a trouvé, pour 1000, les nombres suivants, pour les catégories de hauteurs les plus fréquentes :

						Total
Observé	117	134	157	140	121	669
Calculé	107	137	153	146	121	664
Différence absolue	10	3	4	6	0	23

L'écart est de 23 pour 669 observations ; dans le cas du tirage de boules, il était de 18 pour 2 066 tirages, soit 3,5 p. 100 dans le premier cas, et moins de 0,9 p. 100 dans le second. On ne peut donc, sur ces résultats, assimiler la statistique des tailles à une observation de probabilité, ni la moyenne des tailles à une moyenne d'observations d'un même corps affectées d'erreurs accidentelles.

¹ La théorie de l'homme moyen, p. 36.

² Op. cit., p. 139, note.

Nous avons montré ailleurs ¹ que la répartition régulière des tailles, aussi bien que d'autres caractères physiques, autour d'une moyenne, si on veut l'expliquer par des considérations de probabilité, conduit à admettre que chaque cas (chaque formation d'un individu adulte d'une taille donnée) est indépendant de tous les autres. L'indépendance des causes de production des faits envisagés est le fondement d'un tel calcul ². Mais si, d'une part, en vertu de la transmission héréditaire des caractères physiques, les membres d'une même famille ou de familles associées, ou de telles variétés d'une même race, reproduisent ou tendent à reproduire ceux de leurs ascendants, si, d'autre part, les hommes d'une même époque et d'une même région subissent en commun l'action des mêmes forces, du même milieu physique et social, comment, au point de vue de leur stature, seraient-ils indépendants ? Loin que le désordre, ici, c'est-à-dire le hasard des naissances et évolutions individuelles explique la régularité, c'est la régularité, c'est-à-dire ce sont les conditions générales d'existence de l'espèce et du groupe, qui expliquent que les individus se disposent, sous ce rapport comme sous beaucoup d'autres, en séries bien réglées. En d'autres termes, les individus viennent seulement remplir les cadres déterminés par des lois physiques et sociales constantes. L'étude des individus ainsi « encadrés », c'est-à-dire des groupes, permet de définir ces cadres : mais aucun cas individuel ne découvre, à lui seul, l'existence et la forme d'aucun de ces cadres ³.

Reste l'argument tiré de ce que les deux sortes de séries distinguées par M. Simiand, séries de mesures d'une seule grandeur physique donnée (la densité d'un corps) et séries de prix se ramènent l'une et l'autre à des suites d'observations individuelles, et qu'il importe d'ailleurs assez peu qu'une « chose matériellement réalisée » existe ou n'existe pas en dehors d'elles, puisque nous ne connaissons que ces observations. Mais on ne tient pas compte de ce que celles d'une des séries diffèrent de celles de l'autre série par un caractère essentiel : dans la première, les observations sont ou tendent à être individuelles ; dans la seconde, elles sont collectives. Insistons sur ce point.

Supposons que les prix soient autant de mesures d'une grandeur, dont l'existence réelle est, d'ailleurs, hypothétique, qu'on appellera la valeur : les confondrons-nous avec les mesures de la densité d'un corps ? Mais une série de mesures peut être utilisée de deux manières bien différentes. Ou bien on se propose de connaître une grandeur qui a une existence au moins apparente

¹ Op. cit., p. 59.

² Galton remarque que la stature de l'homme n'est pas une donnée simple, mais la somme d'une quantité d'éléments, une cinquantaine d'os séparés, dans le crâne, l'épine dorsale, etc., des articulations, des cartilages, et il conclut : « Si les tailles d'une population, rangée par ordre de grandeur, forment une série si parfaitement régulière, cela résulte du nombre d'éléments variables et à peu près indépendants dont la taille est la somme. »

³ L'égalité des sexes à la naissance (avec légère prédominance des mâles) paraît offrir un exemple remarquable d'un résultat de probabilité. D'après les recherches les plus récentes (voir : Brachet, *L'œuf et les facteurs de l'ontogénèse*, Paris, Doin, 1917, p. 162 et suiv.) le sexe serait déterminé par les éléments sexuels mâles : il y aurait toujours un nombre pair de spermatoocytes, et la moitié d'entre eux contiendraient un chromosome supplémentaire (hétérochromosome) ce qui les prédisposerait, après conjugaison avec l'œuf, à donner des femelles, mais ce qui, en même temps, les retarderait un peu (d'où le léger excédent de naissances mâles). - Ceci n'est, toutefois, qu'une hypothèse; car « il est possible que certains œufs, ou même qu'un œuf aux divers états de sa maturation, attire une catégorie de spermatozoïdes de préférence aux autres » (Ibid., p. 174).

hors de nous. Alors, celui qui mesure devra s'efforcer de se subordonner à ce qu'il perçoit, et d'échapper, tandis qu'il observe, à toutes les influences que les membres de son groupe ont pu ou peuvent exercer sur lui. Il cherchera en même temps à se dégager de ses habitudes, à ne pas suivre la ligne du moindre effort, à ne pas se laisser influencer par tel nombre dont il a pu garder le souvenir, etc. C'est ainsi que, au jeu, d'un coup à l'autre on mêle les cartes, ou que, comme à la roulette, on use d'un dispositif matériel qui empêche que l'intention du joueur se manifeste d'une façon quelconque à travers les parties successives. On cherche, en somme, à isoler chaque observation et à la rendre indépendante de toutes les autres, à éliminer, de la série des observations, tout élément qui introduirait quelque continuité. Ou bien on s'intéresse aux observateurs, on cherche à déterminer suivant l'expression de M. Borel ¹, mais en un tout autre sens qu'il ne l'entend, « la sensation collective », c'est-à-dire la façon dont les hommes de différents groupes (groupes réels, et non constitués au hasard) perçoivent et mesurent un même objet apparent. Mais, alors, il ne faut plus que l'individu se subordonne étroitement à la grandeur, jusqu'à sortir de son groupe et à s'oublier lui-même : il vaut même mieux que l'objet soit tel qu'il évoque, dans l'esprit de l'observateur, des représentations familières, et qu'il lui donne occasion de commettre le genre d'erreurs qui sont propres à la société dont il est membre, ou aux divers individus successifs rattachés par le lien de sa personnalité. Ainsi, dans le cas des mesures d'une même grandeur physique, l'ensemble des mesures n'a aucune unité collective, aucune consistance interne : on pourrait supposer qu'elles ont été effectuées à des siècles de distance l'une de l'autre, par des hommes qui ne se sont pas connus et qui n'étaient point membres des mêmes sociétés. Le seul lien qui existe entre ces unités dissociées est extérieur à leur ensemble : c'est l'action matérielle de la grandeur, qui s'exerce isolément sur chaque observateur. S'il n'y avait pas un tel principe d'unité, matériellement réalisé hors des individus, on ne voit pas comment s'expliquerait que leurs observations s'accordent. Au contraire, dans le cas des prix, envisagés comme des mesures de la valeur, l'ensemble des prix d'un même produit, à une même époque et dans un même milieu, a une unité collective : ces prix sont autant de jugements formulés par des hommes qui se connaissent, ou en tout cas qui entrent en rapports précisément pour s'entendre sur les prix. Le lien qui existe entre ces prix, c'est la communauté d'appréciation d'un même objet ou d'un même service dans le groupe : chacun évalue une marchandise en tenant compte des jugements de prix antérieurs ou contemporains formulés par les autres ou par lui-même. Si chacun ne s'inspirait pas à cet égard de ce qu'on pense dans son groupe, on ne voit pas comment s'expliquerait que les membres du groupe s'entendent sur les prix. La série des prix individuels reconstitue et fait réapparaître un ensemble de tendances collectives, qui existent par elles-mêmes, et dont on se ferait une idée très inexacte, si on voulait en rendre compte par l'action d'une réalité extérieure au groupe.

¹ Borel, op. cit., p. 226 et suiv. M. Borel, d'après des renseignements qui lui ont été fournis par M. Guillaume, directeur-adjoint du Bureau international des poids et mesures, dit qu'« on est arrivé, en multipliant les lectures faites par des observateurs exercés, à atteindre dans les moyennes une précision *beaucoup plus* grande que la précision des observations individuelles », par exemple, au moyen de lectures faites à la loupe sur une échelle millimétrique, *le centième de millimètre*. Il ajoute qu'on a déterminé ainsi « la précision de la sensation collective ». Mais qu'y a-t-il là de collectif ?

Si la moyenne d'une série d'observations peut être envisagée ainsi tantôt comme la mesure exacte d'une grandeur physique matériellement réalisée, et tantôt comme l'expression d'un ensemble de cas individuels qui n'existe pas hors de cet ensemble, on peut prévoir que le choix du groupe d'observations devra répondre à des conditions différentes dans le premier cas et dans le second. Dans le premier il importera surtout que les observations soient assez nombreuses pour que la neutralisation des petites erreurs accidentelles puisse s'effectuer. Mais le lieu, le temps et l'ordre même où elles seront faites n'importent pas. Tout se passera comme dans le cas des jeux de hasard : qu'un grand nombre de joueurs jouent à la fois, qu'un seul joue pendant un temps très long, qu'on commence les observations à n'importe quel moment, qu'on les interrompe quand on voudra pour les reprendre quand on voudra, cela est indifférent. Dans le second, il vaut mieux sans doute que les observations soient nombreuses : les grosses irrégularités individuelles s'atténueront ainsi d'autant mieux, et il y aura un jeu de compensation entre les autres. Mais l'essentiel, comme l'a montré M. Simiand, c'est que l'observation s'attache à des ensembles homogènes et consistants. Dès lors il n'est pas indifférent que les cas relevés le soient en tel lieu ou en tel temps, ni de façon trop discontinue ¹.

Et il n'est pas indifférent non plus qu'au lieu de découper au hasard le champ de notre observation, nous nous efforcions de nous guider sur les « divisions » naturelles de la réalité. « Défions-nous des moyennes, dit M. Simiand : contrôlons, recoupons les indications de moyennes d'un type par d'autres indices, par des données complémentaires ; et ne retenons que celles qui, après ces épreuves, nous apparaissent avoir une consistance véritable et répondre à quelque réalité collective. »

Indiquons, sur un exemple, la portée de ces règles de méthode. Une théorie économique très répandue affirme que le taux du change entre les monnaies de deux États tend nécessairement vers le rapport entre le pouvoir d'achat de ces monnaies sur leurs marchés nationaux respectifs. La dépréciation du change dans un pays devrait donc se traduire par une élévation correspondante des prix. En pratique, cependant, cette correspondance n'existe pas. « il existe actuellement entre la Suède ou la Suisse, d'une part, et la Pologne de l'autre, qui se trouvent aux deux extrêmes parmi les pays européens, une différence dans le coût de la vie du triple au moins, et les prix montent à mesure que de Pologne on passe en Autriche, en Lettonie, en Esthonie, au Portugal, en Tchécoslovaquie, en Italie, en France, au Royaume-Uni, en Suisse ². »

¹ Voir là-dessus les chapitres VII et VIII du livre cité de M. Simiand, p. 39 et suiv. Nécessité de « voir le phénomène se produisant », au lieu de « s'en tenir à des états et des coïncidences qui peuvent être des effets ou des suites de variations intermédiaires fort diverses », d'embrasser « le phénomène dans son entier », de façon à en saisir « au moins une phase dans chaque sens », et à le suivre autant que possible dans toutes ses phases ; nécessité de répéter notre expérience le plus possible, etc.

² Nous l'empruntons à la remarquable étude de M. Gini Corrado, professeur de statistique à l'Université de Padoue : « Rapport sur la question des matières premières et des denrées alimentaires », *Publications de la Société des Nations*, 1922.

Voilà ce que nous apprend une statistique du mouvement général des prix dans chacun des pays à bas change, où l'on représente par une moyenne ou un indice unique l'ensemble des prix pour tous les produits. Mais une telle moyenne, parce qu'elle se rapporte à un seul pays, comprend un nombre d'observations peut-être trop limité. Supposons que nous calculions une moyenne générale des prix pour tous les pays, à haut change et à change moyen aussi bien qu'à bas change, et aussi le prix moyen de l'or (exprimé dans la monnaie de chacun de ces pays). Il est très possible que, dans cette moyenne, n'apparaissent plus, au moins au même degré, le manque de correspondance entre les prix et le change. Tout se passerait alors comme si la théorie classique indiquée ci-dessus correspondait à la réalité. L'équilibre entre les prix et le pouvoir d'achat de la monnaie tend à s'établir ; il y a dans certains pays des facteurs qui interviennent pour troubler ledit équilibre ; mais si on multiplie les observations, les actions de ces divers facteurs se neutralisent.

Mais on peut trouver, inversement, que la moyenne de tous les prix, même calculée pour un seul pays, correspond à un ensemble d'observations trop hétérogènes, à un groupe de faits trop peu consistants. Alors, on distinguera diverses catégories de prix, jugées homogènes et consistantes pour telle ou telle raison, et on calculera la moyenne de chacune d'elles. Il est naturel, en effet, de distinguer par exemple les marchandises que les commerçants achètent à l'étranger, les marchandises indigènes dont tous les éléments proviennent du marché national, et les marchandises fabriquées à l'intérieur, mais en partie avec des éléments importés. Tenons-nous-en aux deux premières. Nous constaterons que les prix des marchandises achetées à l'étranger suivent bien plus exactement les fluctuations du change, et bien plus vite, que ceux des marchandises indigènes. Si ces derniers témoignent d'une réelle inertie à cet égard, n'est-ce point parce qu'ils sont déterminés par les nationaux, tandis que les fluctuations du change dépendent d'un marché international ? Alors, divers facteurs psychologiques entrent en jeu : différence de sensibilité économique, marquée surtout entre les spéculateurs du change (étrangers ou nationaux), et la généralité de la population nationale : « les spéculateurs sur le change sont, en général, mieux en état de connaître les répercussions économiques que la généralité des ressortissants de l'État en question, et accordent une attention plus grande aux phénomènes qui peuvent les déterminer ¹ » ; d'autre part, « tandis que les étrangers ont une connaissance moins approfondie des ressources du pays que ses ressortissants (et tendent à les sous-estimer), le patriotisme qui anime ces derniers tend à les illusionner sur la situation réelle ». Sans doute d'autres raisons de ce retardement de la hausse pour une catégorie de prix comparée à une autre sont invoquées par M. Gini : mais dans celles-ci, comme dans celles que nous venons de reproduire, c'est la différence d'attitude de groupes définis qui s'exprime dans chaque espèce de prix. si le prix moyen est bien ici une donnée proprement statistique, c'est qu'elle correspond à une disposition psychique collective, qui disparaîtrait dans une moyenne des prix en général, à plus forte raison dans une moyenne s'étendant à beaucoup de pays. Elle n'apparaîtrait que comme un écart de la situation d'équilibre, comme une divergence ou une différence avec ce que représenterait l'état d'équilibre, ou l'état uniforme. Mais, en réalité, c'est la différence qui nous intéresse : si on élimine ou neutralise toutes les différences, en effet, on peut

¹ Gini, op. cit., p. 54.

se demander ce qu'il reste, et si, d'élimination en élimination, on n'en est pas arrivé à se retirer toute matière d'observation ¹.

Au reste, si, pour étudier le mieux un fait ou un objet, il fallait l'observer dans des conditions qu'on pourrait appeler les plus normales, c'est-à-dire alors qu'il ne varie pas ou ne varie que faiblement, au lieu de rassembler dans un vaste groupe toutes nos observations faites sur lui aux moments les plus divers pour en dégager la moyenne, nous pourrions limiter notre observation aux périodes d'équilibre, où les causes d'oscillations interviennent le moins. Reprenons l'exemple dit change : M. Gini distingue des périodes où « le cours des changes tend à ne subir que des ondulations périodiques, dues, par exemple, aux variations saisonnières du commerce international », et non des variations systématiques dans un sens donné ². Alors (avant la guerre) « les changes ne présentaient que des oscillations irrégulières autour d'un point d'équilibre ».

Dans ces conditions, « une assurance contre les risques provenant des variations des changes... eût été tout à fait réalisable ». Des spéculateurs pouvaient alors, sans hardiesse excessive, accaparer des devises étrangères au moment où elles étaient dépréciées, pour les revendre ensuite aux époques où la demande de ces mêmes devises devenait la plus forte ; ils contribuaient ainsi, d'ailleurs, à réduire les oscillations en l'un et l'autre sens. Dans d'autres périodes « des oscillations d'allure indéterminée, ou des variations systématiques dans un sens déterminé, viennent s'ajouter aux oscillations périodiques ;... par exemple les spéculateurs se procurent des devises étrangères quand elles sont en hausse, dans l'espoir d'une nouvelle hausse ultérieure ; ils contribuent ainsi momentanément à l'avalissement du change (de leur pays), sauf à l'atténuer ensuite, lors de la revente des devises, une fois qu'une cote suffisamment rémunératrice aura été atteinte ». Mais d'autres actions peuvent se produire : « par exemple, à l'annonce d'une défaite militaire, de troubles politiques, ou d'un conflit grave entre le capital et le travail, les étrangers qui commercent avec la monnaie du pays vaincu ou troublé prévoient que celui-ci aura besoin d'importer davantage, ou qu'il ne pourra exporter autant de marchandises qu'auparavant, et que, par conséquent, il y aura surabondance (de sa monnaie) à l'étranger ; ils font donc, dans cette prévision, baisser le change de ce pays ; ou, s'ils ont appris que le gouvernement a décidé d'augmenter son émission de papier-monnaie, ils prévoient pour cette raison une hausse des prix sur le marché intérieur, ou, d'une façon plus générale, ils jugent que la politique du gouvernement aggravera les conditions économi-

¹ Discutant la conception de M. Pareto, que « les oscillations des diverses parties du phénomène... sont des manifestations d'un seul et unique état de chose », d'un état d'équilibre représenté par la moyenne, nous disions : « On explique les oscillations par la moyenne : c'est la moyenne plutôt qui s'explique par les oscillations ; quant à celles-ci, il faudrait rendre compte non pas seulement du fait qu'elles s'écartent plus ou moins de la moyenne, mais les expliquer en elles-mêmes, puisque ce sont des phénomènes qui ont un sens et une intensité déterminée... La moyenne, loin de rien expliquer, supprime au contraire tout ce qu'il nous intéresserait de connaître, c'est-à-dire les variations et ce qui est derrière. » *Revue d'économie politique*, juillet-août 1920, p. 474.

² *Op. cit.*, p. 51.

ques du pays, ou pourra le conduire à une guerre, et ils déprécient d'avance la valeur de cette monnaie ¹ ».

Rapprochons maintenant, de ces observations faites dans une période où le change a varié *systématiquement* avec une grande intensité, si bien que la plupart des groupes économiques en ont été atteints, ont réagi diversement, ont contribué à accentuer ou à atténuer ces variations, rapprochons-en le tableau des oscillations du change en temps normal, et demandons-nous si, de celui-ci, nous aurions pu tirer, par un calcul ou un raisonnement quelconque, celles-là. Toute prévision de ce genre nous apparaîtra bien inconcevable. Nous ne savons pas si, après toutes ces variations, le change cessera de se trouver au premier plan de la vie économique. Toujours est-il que, s'il s'y est trouvé, s'il est entré comme élément essentiel dans la vie collective de tant de groupes consommateurs ou producteurs, et s'il existe, en ce sens, comme réalité économique distincte, c'est à partir du moment où il a cessé d'être en une situation d'équilibre. C'est ainsi que le temps (au sens météorologique) n'intervient dans les préoccupations courantes des hommes qu'à partir du moment et que dans les pays où il change souvent et par variations assez amples. M. Simiand disait ailleurs ², à propos d'un économiste de l'école mathématique : « toute sa construction tourne autour d'une théorie de l'équilibre ; mais à quoi nous sert cette théorie, même supposée parfaite, si c'est un perpétuel déséquilibre qui nous apparaît être l'essence de la vie économique réelle ? » et, plus loin : « pour avancer vraiment dans la connaissance économique, il faut s'attaquer directement, et d'abord, à des variations, c'est-à-dire à la forme dynamique des phénomènes ³ ».

Certes, si les observations de prix nous révélaient seulement les jugements ou opinions de groupes de vendeurs et d'acheteurs sur une réalité matérielle, qui existerait en dehors d'eux, savoir la valeur des marchandises, entendant par là le rapport entre les quantités disponibles et les besoins supposés mesurables, ou la quantité de travail incorporée dans le produit, ou toute autre donnée du même genre, alors les *variations* relevées dans ces jugements ou opinions pourraient être de simples erreurs, et le devoir du statisticien serait de faire en sorte qu'elles se neutralisent, et de chercher à atteindre, à travers et au-delà des prix, la valeur elle-même. Mais si la valeur n'a pas une réalité distincte des prix, alors les variations des prix sont tout ce qui existe de la valeur, et il faut les faire ressortir, pour les étudier directement.

Si, de ce point de vue, la méthode statistique s'oppose à la science des probabilités, elle s'appuie sur elle, au contraire, à un autre. En d'autres termes les lois que la statistique cherche à établir ne sont pas du même type que les lois de probabilité : ce ne sont pas des lois dit hasard, qui supposent l'indépendance des faits individuels : au contraire, il n'y a statistique que là où il y a un ensemble consistant, c'est-à-dire un système d'actions qui s'exercent simultanément sur tous les membres d'un groupe, et qui créent ainsi entre eux beaucoup de liens d'interdépendance : toute loi positive exprime des rapports de ce genre. Mais, pour établir de telles lois, la statistique doit s'aider très souvent des lois du hasard.

¹ Op. cit., p. 54.

² *La méthode positive en science économique*, p. 134.

³ *Ibid.*, p. 136.

Nous avons vu que l'on compare souvent des chiffres calculés, qui expriment ce qui se passerait si le hasard jouait seul, et les chiffres observés, afin, précisément, de mesurer le sens et la grandeur des variations qui ne s'expliquent point par le hasard, c'est-à-dire par un jeu d'actions individuelles. Par exemple, M. Perrin reproduit les vingt-cinq premiers coefficients atomiques des corps simples, dans l'ordre de grandeur croissante 4 ; 7 ; 9,1 ; 11 ; 12 14,01 16 ; 19 ; 20 23 ; 24,3 27,1 ; 28,3 ; 31 ; 32 35,47 39,9 ; 39,1 40,1 ; 44 48,1 ; 51 2 ; 52,1 ; 55. Tous ces nombres sont des nombres *observés*, résultant de mesures exactes à 0,01 près. Or « si les valeurs des coefficients atomiques (mesurées avec ce degré d'exactitude) étaient distribuées au hasard, on pourrait s'attendre à ce que 5 éléments sur 25 aient un coefficient entier à 0,1 près ¹ ». Mais, sans parler de l'oxygène (auquel on a imposé un coefficient entier) 20 éléments se trouvent dans ce cas. On pourrait s'attendre à ce que un élément ait un coefficient entier à 0,02 près ; mais 9 sont dans ce cas. D'où il conclut que « une cause encore inconnue maintient donc la plupart des différences de poids atomiques au voisinage de valeurs entières ». C'est la comparaison entre les résultats de l'observation et du calcul des probabilités qui met en lumière l'action d'une telle cause. Supposons, maintenant, que nous possédions, pour un grand nombre de ménages ouvriers, des données sur la répartition de leurs dépenses, et que nous ayons réussi à grouper tous ces ménages (à quelques exceptions près) en un certain nombre de catégories d'après le type de répartition en même temps que d'après le montant total de leurs dépenses. Si, entre ces catégories, il y a des intervalles très marqués (par rapport à l'intervalle qui correspond à chaque catégorie), nous pourrions aussi dire qu'il y a une cause qui maintient les ménages ouvriers ainsi groupés, parce que, s'ils étaient distribués au hasard, les intervalles de séparation entre les groupes n'apparaîtraient point ².

Dans ces deux exemples, tout se passe comme si on imaginait, à côté de la série observée, une série fictive qui diffère de la précédente en ce que la cause en question n'y intervient pas.

Mais il en est de même lorsque, comme nous l'avons montré ci-dessus, on décompose une moyenne d'ensemble en plusieurs moyennes composantes ³. Dans la première, les caractéristiques de celles-ci disparaissent, et tout se passait comme si on réalisait, à côté de plusieurs séries d'un même phénomène, dont chacune, sous l'action de causes propres, développerait des effets différents, une série de ce phénomène d'où toutes ces causes seraient éliminées. Cela revient à supposer encore que l'on étudie ce phénomène dans une

¹ « Car le cinquième seulement d'un grand nombre de points marqués au hasard sur une règle graduée en centimètres, subdivisée en millimètres, tombe dans les segments chacun de 2 mm qui contiennent les bouts de chaque centimètre », Perrin, *Les atomes*, 1913, p. 36.

² Voir notre livre : *La classe ouvrière et les niveaux de vie. Recherches sur la hiérarchie des besoins dans les sociétés industrielles contemporaines*, Alcan, 1912, p. 256 et suiv. ; 12 p. 100 seulement des ménages étudiés n'ont pu entrer dans ces groupes.

³ Quelquefois, d'ailleurs, dans la moyenne d'ensemble elle-même, telle ou telle moyenne composante ne disparaît qu'en partie. Dans les courbes pearsoniennes, on distingue quelquefois deux maxima, ce qui laisse supposer que, parmi les groupes associés, il y en a un qui est trop important pour que son action se fonde entièrement dans la moyenne de tous les groupes.

série de cas individuels rigoureusement indépendants les uns des autres, c'est-à-dire faisant partie de séries différentes : cela revient à se placer dans le cas où ne s'appliquent que les lois du hasard. Par exemple lorsqu'on calcule la fréquence des morts dans un groupe comprenant les deux sexes et toutes les classes d'âge, c'est comme si on la calculait pour un ensemble qui ne serait soumis ni aux causes de mort propres à chaque sexe, ni à chaque âge, et sur lequel par conséquent ne s'exercerait aucune des actions collectives qui définissent un groupe.

Si on réussissait à isoler matériellement chacun des composants d'un groupe d'ensemble, on étudierait l'action de chacun des facteurs sans être gêné par l'action des autres. Mais y parvient-on, même dans les sciences physiques ? Reprenons l'exemple indiqué au début : voici un ensemble d'observations sur les variations de volume d'une même quantité d'un gaz. Puis-je isoler l'action de la pression, puis l'action de la température, et les étudier séparément ? Oui, mais par un artifice, c'est-à-dire en fixant dans le premier cas la température, dans le second la pression ; mais si je les fixe, je ne les supprime point : j'oppose en réalité aux actions qui tendraient à élever la température d'autres actions qui l'abaissent : je réalise à ce point de vue un état d'équilibre, c'est-à-dire une moyenne entre des températures plus ou moins élevées que celle où je maintiens le corps. Mais, d'autre part, quand je calcule une moyenne du chômage d'un mois, puis d'un autre, pour toute une série d'années, je suppose que les variations interannuelles, si je prends une période assez longue, se neutralisent, j'oppose aux variations en un sens à partir de la moyenne les variations à partir de la même moyenne en un sens opposé. La différence entre les deux cas est seulement que, dans le premier, je fixe d'avance ma moyenne, et que, dans l'autre, j'admets que les causes de variation de chaque année ainsi ajoutées exercent une influence égale à chaque mois, si bien que je réussis à *fixer* la variation annuelle ; mais c'est là une application des lois du hasard. Je n'ai aucune raison d'admettre, en effet, que l'action de la variation annuelle se fasse sentir, plus ou moins, d'une façon systématique, en tel mois qu'en tel autre : en d'autres termes je dois supposer qu'elle se distribue au hasard sur tous les mois, et qu'en la calculant pour un même mois dans toute la période j'opère sur des grandeurs indépendantes, qui obéissent aux lois de probabilité, c'est-à-dire qui se distribuent également autour d'une moyenne. Je ne peux, en résumé, éliminer l'action d'un ensemble de facteurs, pour étudier isolément l'action des autres, qu'en appliquant une loi de probabilité¹. Mais, inversement, je ne peux attribuer un sens à l'action de ces autres facteurs, et je n'ai des raisons de l'étudier, que parce que les variations correspondantes forment un ensemble consistant, que parce que, dans notre exemple, les variations du chômage d'un mois à l'autre dépendent les unes des autres, c'est-à-dire ne sont pas soumises à une simple loi de probabilité.

¹ On pourrait, ainsi que nous l'a suggéré M. Simiand, calculer d'abord pour chaque année le rapport du nombre des chômeurs de chaque mois au nombre des chômeurs de cette année supposé uniformément égal à 100, ce qui revient à éliminer tout de suite la variation annuelle. Mais, même alors, on n'obtiendra pas chaque année, pour le même mois, la même valeur relative, parce que la variation saisonnière peut être sujette, d'une année à l'autre, à de petites variations accidentelles. Il faudra donc éliminer autant que possible ces petites variations, en calculant, pour chaque mois, la moyenne de la valeur relative dans toute la série d'années, c'est-à-dire appliquer une loi de probabilité.

Ainsi les opérations statistiques présentent tous les caractères d'une méthode expérimentale ; et, toutefois, elles sont en de si étroits rapports avec la théorie et le calcul des probabilités qu'on a pu se demander récemment encore si « le calcul des probabilités ne constitue pas la base de toutes J'es prévisions statistiques ¹ ». Que ces deux propositions ne se contredisent point, c'est ce qui résulte de notre examen du rôle assigné, en statistique, aux raisonnements de probabilité. Les lois statistiques se distinguent des lois du hasard, et peut-être conviendrait-il, pour éviter toute équivoque, de renoncer à des expressions telles que : « probabilités statistiques » et « équilibre statistique », qui désignent des résultats sans rapport avec ceux que poursuit l'étude expérimentale et positive des groupes. Mais le statisticien est obligé, à chaque instant, de se servir du calcul des probabilités, pour analyser les objets collectifs qu'il observe, et déterminer isolément les variations de chacun de leurs éléments, dans leurs rapports avec d'autres objets collectifs ou avec d'autres objets quelconques. En ce sens, le calcul des probabilités joue, en statistique, à peu près le même rôle que *les instruments* dans l'expérimentation physico-chimique.

¹ Medolaghi, « La previsione statistica e il calcolo delle probabilità », *Metron*, décembre 1920.