ANNALES DE L'I. H. P.

CORRADO GINI

Sur quelques questions fondamentales de statistique

Annales de l'I. H. P., tome 14, nº 4 (1955), p. 245-364

<http://www.numdam.org/item?id=AIHP_1955__14_4_245_0>

Sur quelques questions fondamentales de statistique (*).

par

Corrado GINI.

CHAPITRE I.

NÉCESSITÉ ET DANGERS DE LA STATISTIQUE.

Pour discuter profitablement de la nécessité et des dangers de la statistique, il faut d'abord s'entendre sur ce qu'elle est réellement, car ce mot recouvre des concepts très différents. Le commis du recensement et le secrétaire de la municipalité ne le voient certainement pas sous le même angle que le professeur qui l'enseigne à l'école de commerce, et celui-ci en a une notion bien différente de celle des spécialistes de Physique atomique. Ces divergences tiennent au fait que l'on comprend sous le nom de Statistique un ensemble de techniques extrêmement variées au point de vue des moyens qu'elles mettent en œuvre, des difficultés qu'elles soulèvent et des connaissances qu'elles présupposent. Ces techniques vont des besognes matérielles de la collection des données et de leur contrôle, dépouillement et tabulation, pour lesquelles suffisent une instruction élémentaire et de la diligence, à l'élaboration des résultats, qui nécessite au moins une bonne instruction secondaire et notamment la connaissance des mathématiques relatives à ce niveau, et enfin à l'évaluation de la confiance qu'il convient d'attri-

(*) Cet article réunit six conférences données entre le 27 avril et le 5 mai 1954 à la Faculté de Sciences de l'Université de Paris, dont quatre (reproduites aux chapitres II, IV.1, IV.2 et V) à l'Institut Henri Poincaré et deux (reproduites aux chapitres I et II) à l'Institut de Statistique.

buer aux dits résultats et à leur comparaison, qui exige l'application des mathématiques supérieures et en particulier du Calcul des Probabilités; sans parler de l'interprétation des résultats dans le cadre de la science dans laquelle ils rentrent, qui demande pour le moins l'assimilation des éléments de la science en question et qui implique souvent des problèmes délicats de la théorie philosophique de la connaissance.

Toutes ces techniques ont en commun ce caractère qu'elles sont nécessaires, ou tout au moins utiles, pour étudier au point de vue quantitatif certains phénomènes qui n'apparaissent avec l'exactitude désirée que dans une masse ou collection d'observations et qui, par conséquent, sont appelés *phénomènes de masse ou collectifs.*

C'est là le fondement, la raison d'être de la Statistique : la nécessité d'y recourir pour l'étude quantitative des phénomènes collectifs.

*
* *

D'où vient cette nécessité? Pourquoi l'observation ordinaire ne permet-elle pas d'étudier quantitativement les phénomènes collectifs sans l'aide de la technique statistique?

Cela provient de plusieurs causes ([1]).

D'abord d'une capacité très limitée de notre esprit d'opérer des synthèses quantitatives. Chacun de nous connaît un certain nombre, de dizaines, peut-être de centaines, de familles et, pour chacune d'elles, le nombre et le sexe des enfants, mais, si l'on nous demande quel est le nombre moyen des enfants ou le rapport des filles aux garçons dans l'ensemble des familles en question, nous ne savons pas répondre. Ici il ne s'agit que de dizaines ou tout au plus de centaines d'unités. Comment pourrait-on, sans la technique statistique, se rendre compte du nombre des habitants d'un pays ou d'une région ou d'une ville, ainsi que de leur composition par sexe, par âge, par profession, etc.? Toutes les fois qu'il y a une masse nombreuse de cas, la technique statistique devient donc nécessaire pour en tirer des notions quantitatives. Lorsque Graunt

---

([1]) Plusieurs de ces causes avaient été déjà signalées dans le volume : ***Il sesso dal punto di vista statistico***, Sandron, Palermo, 1908, p. 12-17, ainsi que dans les cours de Statistique professés depuis 1909 aux Universités de Cagliari, Padoue et Rome. ***Voir***, pour tous, le ***Corso di statistica***, *a cura di S. Gatti e C. Benedetti* (*Nuova edizione aggiornata*) Année académique 1954-1955, Rome, Veschi, 1955.

a donné le premier exemple d'investigation statistique, on attribuait à Londres deux millions d'habitants : il les a réduits à 400 000 ([2]).

Notre mémoire aussi est limitée. Il s'ensuit que, même si la masse n'est pas nombreuse, l'aide de la statistique est nécessaire lorsque les événements se succèdent à de longs intervalles. Par exemple, les centenaires qui meurent dans un pays sont en nombre limité, mais ces décès sont tellement espacés dans le temps qu'après une année l'on ne saurait en dire le nombre. Pour l'exprimer il faut en prendre note, c'est-à-dire dresser une statistique pour rudimentaire qu'elle soit.

Il faut de même se méfier de l'impression courante lorsqu'elle se base sur d'anciennes statistiques, car dans l'intervalle des modifications peuvent être survenues dont, sans statistique nouvelle, on ne saurait dûment tenir compte. C'est ainsi que lorsqu'en 1927 on a fait le premier recensement de la population de la Turquie, on a trouvé un chiffre plus ou moins double de celui généralement admis. Par contre on a constaté à la première enquête statistique que les chiffres généralement admis pour certaines régions de l'Afrique du Nord — telles la Tripolitaine et le Maroc — étaient excessifs. C'est que la population de la Turquie avait subi dans les années précédentes un essor démographique remarquable, dont on n'avait pas pu se rendre compte faute de statistique du mouvement de la population, tandis que pour la Tripolitaine et le Maroc on répétait de vieilles estimations sans réaliser les conséquences des guerres et des émigrations successives.

D'autres sources d'erreurs dans les évaluations de la population faites sans l'aide de la statistique proviennent soit de doubles emplois soit de l'omission des habitants. C'est ainsi que les navigateurs européens qui, les premiers, ont jeté l'ancre devant les petites îles du Pacifique en ont rapporté des chiffres fabuleux quant à leur population, et ceci parce que les indigènes suivaient les bateaux d'un port à l'autre, jamais rassasiés d'admirer ces îles flottantes, et que les navigateurs — qui ne s'en rendaient pas compte — additionnaient mentalement les foules qui les recevaient. Une exagération semblable s'est vérifiée — d'après l'aimable information d'un collègue français — dans les premières

([2]) *Cf.*, à ce sujet, la conférence : *The first steps of Statistics* (*Educational Research Forum, Proceedings*, Endicott, New York, 1948, p. 37 et suiv.). Dans ladite publication la typographie avait omis plusieurs périodes, mais le texte a été retabli dans son intégrité dans les tirés à part.

évaluations des Touaregs de l'Afrique du Nord. Lorsque les Français pénétrèrent dans leur territoire, ils les rencontraient partout sur leur chemin, et en rapportèrent l'impression qu'ils se montaient à des centaines de milliers, tandis que, lorsque l'on a pu en faire un relevé systématique, ils se réduisirent à quelques milliers. Évidemment, c'étaient toujours les mêmes groupes qui se déplaçaient suivant les troupes françaises que probablement ils surveillaient. D'autres tribus indigènes s'effacent au contraire devant le voyageur civilisé dont elles ont parfois de bonnes raisons de se méfier, de sorte que, se fiant à ses rencontres, le voyageur sous-estimera la population locale et, s'il traverse un pays habité par des nomades, il risquera même de le croire dépeuplé.

Les sources d'erreurs que je viens de rappeler, et que seule la technique statistique peut éviter, concernent l'énumération des cas, premier but de la statistique.

D'autres erreurs, dont les causes sont parfois obscures, affectent le mesurage des phénomènes collectifs, autre but de la statistique.

Pourquoi les estimations des experts, en particulier celles concernant les récoltes, qui sont dignes de confiance pour les valeurs normales et sous-normales, restent-elles en général en dessous de la réalité pour les valeurs excédant la normale, et ceci d'autant plus que la valeur réelle s'élève au-dessus de la normale? ([3]). Les vérifications statistiques paraissent avoir mis ce résultat hors de doute, mais l'explication nous échappe.

Et pourquoi encore les prévisions des variations de prix faites par les experts sont-elles erronées le plus souvent par excès que par défaut?

Des valeurs excessives, paraît-il, s'obtiennent tout au moins en Amérique pour le coton, en attribuant aux prévisions des « bears » le même poids qu'à celles des « bulls » ; des statisticiens avisés, après avoir mesuré de la sorte la tendancialité de leurs prévisions, ont su la neutraliser avec la pondération appropriée et régler avec succès leur conduite en bourse ([4]). Dire que cela arrive parce que l'homme est optimiste est une explication toute pareille à celle qui attribue à la *virtus dormitiva* de l'opium ses qualités narcotiques.

([3]) *Voir* la communication de F. Yates, ***Estimation of the yield of potatoes in England and Wales by sampling methods***, présentée à la XII[e] Réunion scientifique de la Société italienne de Statistique, Rome, janvier 1953.

([4]) Cf. *Patologia economica*, 5[e] édition, Utet, Turin, 1952, p. 298.

D'autres fois, les causes de nos impressions erronées peuvent être indiquées avec certitude.

Les anciennes publications des biologistes donnaient une prépondérance, parfois énorme, aux mâles par rapport aux femelles dans la population de maintes espèces animales et particulièrement des papillons ; les plus récentes ont beaucoup réduit et quelquefois même contredit cette prédominance. Les premières se basaient généralement sur les observations courantes, les dernières se fondent sur des statistiques systématiques. La différence entre les résultats des unes et des autres tient à ce que dans maintes espèces animales, et particulièrement chez les papillons, il y a un dimorphisme sexuel très marqué, dans les caractères morphologiques et psychiques, qui a pour effet que les mâles sont beaucoup plus mobiles et d'un aspect beaucoup plus frappant que les femelles, celles-ci étant relativement immobiles et se confondant souvent par mimétisme avec le milieu environnant. A défaut de technique statistique, les mâles apparaissent par conséquent aux yeux de l'observateur avec une fréquence beaucoup plus élevée que les femelles. Dans les espèces qui sont l'objet de la poursuite de l'homme (chasse, pêche), une impression erronée peut provenir de la plus grande facilité d'échapper qu'ont les représentants d'un sexe en comparaison avec ceux de l'autre. Il paraît que, dans certaines espèces de poissons, ce sont les mâles que l'on pêche plus facilement, tandis que les femelles des lièvres tomberaient plus souvent sous le plomb des chasseurs (5). Les pionniers anglais de la Compagnie des Indes commirent semblable erreur lorsqu'ils rapportèrent que dans ce pays il y avait je ne sais combien de femmes pour chaque homme. Les recensements qui suivirent démontrèrent que, non seulement ce n'était pas vrai, mais que, dans les Indes, contrairement à ce que l'on constate en Angleterre et dans d'autres pays européens, ce sont les hommes qui l'emportent au point de vue numérique. L'erreur provenait du fait que lesdits pionniers s'étaient surtout trouvés en contact avec les familles riches du pays, où la polygamie était admise, ces familles pouvant se permettre le luxe d'un

(5) Dans l'ouvrage cité (*Il sesso*..., chap. VI : *La distribuzione dei sessi delle specie animali e vegetali*, p. 183-229), on peut trouver des informations à ce sujet pour les espèces mentionnées dans le texte, ainsi que pour quelques autres (rats, oiseaux, araignées, etc.).

harem bien fourni. Ils généralisaient de la sorte des observations qui n'étaient pas du tout représentatives (6).

Même, lorsque tous les cas ont une chance égale de tomber sous notre observation, il se peut que quelques-uns frappent systématiquement notre attention plus que d'autres. Ils s'inscrivent, par conséquent, plus profondément dans notre mémoire et nous laissent l'impression d'être plus fréquents qu'ils ne le sont en réalité.

Les contrastes nous frappent plus que les concordances. Si, dans un couple, l'homme et la femme sont bien assortis par la taille ou l'âge, s'ils se comportent d'une façon analogue, si tous deux sont pourvus d'un degré de culture à peu près semblable, si l'un et l'autre sont bruns, ou blonds, nous n'y prêtons pas spécialement attention; mais si l'un est brun foncé et l'autre tout à fait blond; si l'un est remuant et bavard, l'autre silencieux et timide; si l'un est de grande taille et l'autre très petit; si l'un est jeune et l'autre âgé; si l'un est manifestement très cultivé tandis que l'autre fait preuve d'une ignorance exceptionnelle, nous le remarquons et ne l'oublierons pas facilement. On comprend dès lors comment l'impression peut être si répandue — même parmi les hommes de sciences — que dans les mariages il y a hétérogamie, c'est-à-dire tendance à s'accoupler entre personnes possédant des caractères contraires. Par contre, les recherches statistiques, vastes et approfondies, que l'on a faites sur ce sujet mettent hors de doute l'existence d'une homogamie, c'est-à-dire d'une tendance à s'accoupler entre semblables, quels que soient les caractères examinés : âge, culture, nationalité, religion, lieu de résidence, taille, pigmentation, ou n'importe quel autre caractère, physique ou psychique (7).

Si un bon élève remporte au cours de sa carrière des succès professionnels, cela nous paraît tout à fait naturel; mais, si un garçon qui était brillant à l'école fait faillite dans la vie, ou si, au contraire, un mauvais étudiant se comporte ensuite très honorablement, nous en restons saisis. Cela explique l'opinion assez répandue, qu'il n'y a aucune relation positive entre le succès ou l'insuccès de la carrière scolaire et celui de la carrière professionnelle. Cette croyance n'a

(6) Voir *Corso di statistica*, cité, p. 22.

(7) Des recherches étendues sur l'homogamie ont été faites en Italie surtout par R. Benini (*Principii di Demografia*. Barbera, Firenze, 1901) et par F. Savorgnan (*La scelta matrimoniale*, *Studi statistici*, Biblioteca di *Metron*, Roma, 1921).

pourtant pas de fondement dans les faits. A. L. Lowell, l'ancien recteur de Harvard, raconte dans son ouvrage bien connu *Public opinion in War and Peace* (8) comment, un de ses collègues lui ayant exprimé une telle opinion, il l'invita à citer des cas à l'appui. Son interlocuteur ne put en mentionner qu'un seul cas probant et un autre douteux; combien d'autres élèves n'avait-il pas connus au cours de sa carrière professorale, dont le comportement aurait témoigné en faveur d'un accord entre le succès scolaire et le succès professionnel ! Ils s'étaient tous effacés de sa mémoire, tandis que le seul cas exceptionnel s'y était imprimé d'une manière inébranlable. Des recherches statistiques ont été faites et il va sans dire qu'elles ont démontré une corrélation fortement positive entre les notes de la carrière académique et le succès de la vie professionnelle.

Peut-être est-ce de la même circonstance — l'impression plus forte que nous recevons des contrastes — que dérive l'affirmation courante que dans certaines espèces les accouchements doubles donnent toujours lieu à un mâle et à une femelle. Les jumeaux seraient la règle pour le chevreuil et, d'après Bellingeri, leur sexe seraient toujours différent. Pour le pigeon, une telle croyance ne remonte pas moins qu'à Aristote, mais les contrôles faits au commencement de ce siècle l'ont démentie : il paraît que les proportions des deux sexes se conforment au contraire au calcul des probabilités (9).

Si, laissée à elle-même, sans coercition, l'attention de l'observateur se porte de préférence sur les contrastes que sur les accords, cela ne signifie pas qu'elle ne puisse être dirigée artificiellement vers ces derniers et détournée des premiers; mais il est d'ailleurs possible qu'elle soit impressionnée par les contrastes plus qu'elle ne le serait spontanément. En quoi consiste en effet la propagande — je parle de la propagande honnête — si ce n'est dans l'art d'attirer l'attention sur les faits qui sont en accord avec la thèse défendue et sur ceux qui sont en contraste avec la thèse adverse ?

Toute propagande mise à part, les événements qui nous choquent, ou tout au moins certaines catégories de ces événements, frappent notre attention et s'inscrivent dans notre mémoire plus profondément que

(8) Abbot Lawrence Lowell, *Public opinion in War and Peace*, Cambridge, Massachusetts, 1923.

(9) Voir *Il sesso*.... cité, p. 191-192.

ceux qui se déroulent en contraste avec notre attente ou tout au moins avec notre désir.

Dans notre société, les parents préfèrent souvent — et plus souvent ils le préféraient autrefois — un garçon à une fille. La statistique a démontré que dans les naissances la fréquence des garçons l'emporte d'environ 5 % sur celle des filles. C'est la première loi statistique que l'on ait établi (10). Avant qu'elle ne le fût, certaines gens prétendaient qu'au contraire six ou sept filles naissaient pour chaque garçon (11). D'une manière analogue, les éleveurs de chiens, interrogés, répondaient que le nombre de femelles dans les portées se trouvait dépasser celui des mâles; même impression pour les chats. Pour les chiens, tout au moins, la statistique a établi le contraire. Les éleveurs de chiens et de chats préfèrent également les mâles et restent — comme les parents dans nos familles — déçus par les naissances de femelles.

Il y a lieu de rappeler encore que, lorsque les suffragettes bouleversèrent la tranquille routine de la vie politique anglaise de façon qu'il fut question de faire droit à leurs revendications, les députés qui les combattaient avançaient, comme un des arguments décisifs, le fait que de cette façon on aurait eu trois fois plus de voix féminines que de voix masculines à cause de l'énorme excès de femmes dans la population adulte. Or, si un certain excès de femmes peut être effectivement constaté dans les pays les plus civilisés d'Europe à cause du traitement particulièrement favorable que les femmes se sont assuré dans nos sociétés — et ce déséquilibre est particulièrement marqué dans les pays d'émigration tels que la Grande-Bretagne — on est, malgré tout, très loin de l'énorme disproportion avancée par les antisuffragistes, qui auraient d'ailleurs pu s'en convaincre facilement en consultant les chiffres des recensements (12).

Plusieurs causes d'erreurs entrent ici en ligne de compte.

D'abord, il est certain que l'activité tapageuse des suffragettes devait donner l'impression qu'elles — et, par conséquent les femmes, dont elles faisaient partie — étaient plus nombreuses qu'en réalité. En second lieu, pour les mysogines, parmi lesquels se recrutaient évidem-

---

(10) Voir *Il sesso...*, cité, p. 27-29 et 33.
(11) Voir *Corso di statistica*, cité, p. 23.
(12) Voir *Corso di statistica*. cité. p. 23.

ment les antisuffragistes, la rencontre d'une femme devait être plus désagréable que celle d'un homme et, en frappant davantage leur attention, leur faisait exagérer le nombre des femmes. Enfin, et c'est ici une nouvelle circonstance perturbatrice à signaler, lorsque l'on a embrassé une thèse, on est inconsciemment porté à exagérer tous les arguments qui militent en sa faveur, de sorte qu'il n'est pas surprenant que les antisuffragistes aient exagéré la disproportion des sexes en combattant la réforme, et le fait qu'ils ne consultaient pas les données des recensements montre bien qu'ils ne gardaient pas une objectivité parfaite dans le feu des discussions.

Sur ce point une remarque s'impose. Lorsque nous nous laissons guider impartialement par notre conscience, les événements contraires à notre attente nous frappent davantage, de sorte que nous en exagérons l'importance. Lorsque, au contraire, nous avons embrassé une thèse, notre subconscient entre en action impérieusement et nous tendons à négliger les cas contraires à notre attente. Le conscient et le subconscient se comportent donc différemment et nous pouvons accepter, d'une part, l'explication avancée par les statisticiens relativement aux affirmations erronées des pré-statisticiens quant au rapport des sexes dans les familles humaines et les portées de chiots, et d'autre part l'enseignement des psychanalistes, d'après lesquels notre subconscient refoule dans l'inconscient les événements désagréables ([13]).

C'est la raison pour laquelle les casuistiques des médecins — si sérieux que ceux-ci puissent être, et ils le sont quelquefois — n'ont aucune valeur scientifique et restent souvent démenties par les statistiques.

Et c'est encore pour cette raison que tout statisticien expérimenté sait combien il faut se méfier de soi-même lorsqu'il s'agit de vérifier les chiffres suspects. Le bon système consiste à faire derechef le relevé ou les calculs, car, si l'on se borne à ne vérifier que les chiffres que l'on soupçonne, on est porté inconsciemment à ne soupçonner et ne vérifier que ceux qui vont à l'encontre de notre attente et à accepter sans contrôle ceux qui la confirment.

C'est de cette manière que les préjugés s'alimentent et prennent pied.

([13]) *Voir*, par exemple, H. Wolterek. *La porta dell'anima*. G. Casini. Roma, 1951, p. 68.

Comme nous l'avons relevé, un contraste nous frappe plus souvent qu'un accord et en effet les préjugés défavorables ou portant sur des événements défavorables sont beaucoup plus fréquents que ceux de la bonne chance. Par une application risquée du principe de la concordance ou de celui de la différence, nous établissons une connexion causale entre les phénomènes en question. Cette croyance établie, nous arrêtons notre attention sur tous les événements qui la confirment et négligeons tous ceux qui la contredisent. Ainsi la croyance prend pied et devient un préjugé.

Le seul moyen de le déraciner ou, le cas échéant, de le confirmer, serait de le soumettre à l'épreuve du feu de la statistique.

J. Graunt en Angleterre et G. Neumann en Allemagne l'on fait avec succès pour certains préjugés de leur époque, d'après lesquels les phases de la lune et les conjonctions des astres auraient affecté la santé publique et les changements du gouvernement auraient été suivis d'épidémies de peste, sans parler de l'influence des nombres cabalistiques ([14]). Leurs résultats dans ce domaine ont certainement contribué à gagner à la science statistique nouvelle apparue la confiance du public. Dès lors, ces recherches n'ont plus été poursuivies, et cependant il y aurait un vaste domaine ouvert à leurs applications en météorologie et en agriculture, où de multiples proverbes sont couramment répétés et reçoivent pleine créance de la part des agriculteurs qui règlent d'après eux les semailles et les travaux des champs.

Il se peut que la statistique les démente, mais il se peut aussi qu'elle les confirme et les précise.

En effet, l'observation courante ne nous trompe pas toujours; elle est souvent fondée.

Quelquefois même, par un processus qui nous échappe, son résultat concorde d'une façon surprenante avec les mensurations que l'on tire des statistiques. C'est ainsi que l'école clinique de De Giovanni, s'occupant de la constitution humaine, a mis en lumière que les individus que nous jugeons normaux présentent avec une très bonne approximation les dimensions de l'homme moyen ([15]); c'est ainsi que, comme nous le

([14]) *Voir* la conférence déjà mentionnée : *The first steps of Statistics*, p. 40.

([15]) *Voir* en particulier, G. Viola, *Le dimensioni dell'uomo medio normale* (*Lavori dell'Istituto della Clinica Medica di Padova*, t. II, Hoepli, Milano. 1904-1905; reproduit en *La costituzione individuale*, t. I. Cappelli, Bologna, 1932).

verrons ensuite, l'âge qui, dans certaines régions au moins de l'Italie, est jugé le plus propre pour l'épouse quand l'époux a un âge donné, correspond assez bien à la moyenne des âges effectivement présentés par les épouses. Mais, en dehors de ces cas isolés, l'observation courante ne nous donne que de vagues impressions qualitatives : c'est à la statistique que nous devons recourir pour obtenir des mesures précises.

Point n'est besoin de statistique pour nous apprendre que les tailles moyennes sont les plus fréquentes et que les tailles élevées ou, au contraire, basses sont plus rares, et même d'autant plus rares qu'elles s'éloignent le plus des moyennes; mais, si nous voulons connaître la courbe de distribution des tailles, le recours à la statistique devient nécessaire.

Nous sommes presque toujours en mesure de dire avec certitude, sans l'aide de la statistique, que le coût de la vie augmente ou au contraire, diminue. Mais, si nous voulons préciser la portée de la hausse ou de la baisse, nous ne pouvons nous passer de la statistique.

Enfin il est des domaines dans lesquels, sans l'aide de la statistique, une seule observation fournit la mesure d'un phénomène avec une certaine approximation, mais, si celle-ci peut être suffisante dans certaines recherches et en vue de certains buts, elle ne l'est plus pour d'autres. La mesure d'une longueur, d'une hauteur, d'un poids, d'un intervalle de temps, est affectée par des perturbations dues soit aux agents extérieurs, soit aux instruments employés, soit encore à l'observateur lui-même, perturbations qui ont pour effet, lorsque l'on effectue deux fois de suite la même mesure sur le même objet, que l'on n'obtient qu'exceptionnellement deux valeurs exactement les mêmes. Et, si l'on veut un résultat plus exact, il faut procéder à une masse d'observations d'après la technique statistique et recourir encore à celle-ci pour en déduire la valeur la plus plausible à l'aide d'une moyenne.

En résumé, nous pouvons dire qu'il existe des phénomènes dont, sans l'aide de la statistique, nous ne pouvons pas nous rendre compte ; d'autres qui, sans cette aide, ne nous laissent qu'une impression erronée de leur dimension ; d'autres encore qui nous fournissent des informations exactes, bien que seulement qualitatives et pour lesquels une évaluation quantitative nécessite l'emploi de la statistique ; d'autres enfin pour lesquels nous pouvons obtenir une évaluation quantitative sans l'emploi de la statistique, mais alors seulement avec une certaine approximation, tandis que nous devons recourir à la statistique pour obtenir une précision plus grande. Dans une condition idéale, où chaque phénomène serait énuméré et mesuré avec toute l'exactitude possible, il n'y en aurait donc aucun qui serait observé sans l'aide de la statistique. Et si, dans la pratique, on n'y recourt pas toujours, c'est que l'on s'accommode de notions qualitatives ou de mesures approchées par économie de temps et d'argent ([16]).

*
* *

Si la statistique est nécessaire en théorie pour la mesure exacte de tout phénomène, et en pratique tout au moins pour le traitement quantitatif des phénomènes collectifs, elle est pourtant dangereuse.

« Avec la statistique, on démontre tout ce que l'on veut », entend-on dire souvent et pour le démontrer on n'a pas disette d'exemples de statisticiens qui arrivent, sur les mêmes phénomènes, à des conclusions opposées. « Il y a trois espèces de mensonges — disent les Anglais — dont la pire est représentée par les statistiques ».

Si l'on entend par là que les statistiques mensongères constituent le pire des mensonges, je suis tout prêt à souscrire. C'est le pire des mensonges, car c'est le plus difficile à dévoiler. Cela est dû au caractère extrêmement synthétique des statistiques. En effet, on ne réalise pas généralement quel énorme travail est condensé dans un chiffre et particulièrement dans un de ces grands nombres qui sont caractéristiques de la statistique. Condensé et caché! Pour l'analyser et le contrôler, des recherches laborieuses seraient nécessaires. Sauf un intérêt tout spécial, on ne les fait pas et l'on accepte le chiffre tel quel!

Non seulement le chiffre est synthétique, mais les qualifications avec lesquelles on l'accompagne sont, dans bien des cas, elliptiques. Et c'est d'ici que proviennent certaines contradictions — apparentes plutôt que réelles — des statisticiens. Voyons-en un exemple.

Soient deux pays A et B. Dans lequel des deux la mortalité est-elle plus élevée? Le rapport des morts à la population est plus élevé dans A que dans B (supposons que la différence relative soit de 10 %), et alors de nombreux statisticiens en concluent que la mortalité est plus

---

([16]) Cf. *Corso di statistica*, cité. p. 24.

élevée dans A. Mais d'autres disent : — Non, cette conclusion n'est pas autorisée ; elle est même erronée. Il y a dans A beaucoup plus d'enfants que dans B, et la mortalité est, comme on sait, particulièrement élevée chez les enfants ; il faut éliminer l'influence de l'âge par la méthode, par exemple, de la population type. On trouve alors que la mortalité est au contraire, plus élevée dans B avec une différence relative, disons de 8 %. — Un instant ! objecte un troisième groupe de statisticiens, A est un pays de forte émigration et le nombre des hommes y reste beaucoup inférieur à celui des femmes, ce qui n'a pas lieu, par contre, dans B ; nous devons éliminer aussi l'influence du sexe, du moment qu'il est certain que la mortalité, à égalité d'autres circonstances, est plus élevée pour le sexe masculin. Cette influence étant aussi éliminée, les coefficients de mortalité apparaissent égaux pour les deux pays. — Mais cela n'est pas suffisant, observe un quatrième groupe; la disproportion des sexes rend plus rares les mariages et, par conséquent, la proportion des mariés dans A est moins élevée : il faut éliminer aussi l'influence de l'état civil. Après cette troisième élimination, la mortalité apparaît de nouveau plus élevée dans B, mais en faible proportion (différence relative de 4 %). Et pourquoi ne pas éliminer aussi l'influence de la profession, qui est si forte, sur la mortalité? Ce facteur encore étant éliminé, la différence relative de la mortalité au dommage de B monte à 15 %. Mais pourquoi ne pas éliminer aussi l'influence de la richesse ? Après cette autre élimination, la mortalité apparaît de nouveau plus élevée dans A, mais en mesure à peine sensible (différence relative de 1 %). Cependant le climat est très différent dans les deux pays. Admettons qu'il soit possible d'éliminer aussi l'influence de ce facteur et qu'à la fin la mortalité apparaisse presque égale dans les deux pays avec une différence relative de 2 % au dommage de B. Ainsi, dans une question aussi simple que celle de savoir si la mortalité est plus élevée dans A ou dans B, nous sommes arrivés, par les méthodes statistiques, à sept conclusions plus ou moins différentes. La différence relative à l'avantage de B est, en effet, apparue successivement : 10 %, — 8 %, 0 %, — 4 %, — 15 %, + 1 %, — 2 %. Et nous ne pouvons même pas dire qu'il s'agisse d'approximations successives vers le résultat final ; la valeur la plus approchée (0 %) du résultat final (— 2 %) a été obtenue, en effet, après la deuxième élimination et le résultat le plus divergent (— 15 %) après la quatrième élimination. En réalité, les valeurs successivement obtenues se disposent irrégulièrement autour du résultat final.

Quelle est donc la conclusion exacte?

En réalité, elles sont toutes exactes en un certain sens et toutes peuvent être erronées en un certain sens. La divergence dépend du fait que le terme « mortalité » s'entend en des sens différents. En effet, nous pouvons vouloir mesurer, au moyen de la mortalité, la résistance vitale de la population considérée dans sa composition effective par âge, sexe, état civil, profession, richesse et dans le milieu où elle vit effectivement; et, pour certains buts, c'est justement ce qu'il nous faut. Ou bien nous pouvons vouloir mesurer sa résistance vitale en abstrayant de la circonstance passagère pour tout individu, de l'âge; et cela peut aussi répondre à certains buts. Ou bien, encore, nous pouvons vouloir mesurer la résistance vitale en abstrayant aussi du sexe et de l'état civil, ce qui peut être pleinement justifié lorsque, par exemple, on veuille faire une prévision sur la résistance vitale des générations futures. Pour d'autres buts encore, on peut avoir intérêt à comparer la résistance vitale des deux populations en abstrayant de toute influence sociale et éliminer pour cela aussi l'influence de la profession et de la richesse; et il se peut, enfin, qu'il soit opportun, eu égard au but de la recherche, d'éliminer aussi l'influence du milieu physique : par exemple, lorsqu'on aurait intérêt à déterminer laquelle des deux populations serait mieux apte à coloniser un nouveau territoire dont le climat serait intermédiaire entre ceux des deux pays considérés.

Définir avec précision l'objet de la recherche est une nécessité qui s'impose non seulement pour les investigations statistiques, mais pour les investigations scientifiques de tout genre et nous pouvons donc dire que de semblables inconvénients se présentent dans toutes les sciences. — Oui, mais se présentent-ils dans la même mesure? Il nous faut admettre qu'il serait bien difficile de trouver, dans la définition d'un phénomène non statistique, de si nombreuses significations que celles que nous avons vu qu'il est possible d'attribuer au terme « mortalité » et qui pourraient de même, être attribuées aux termes « natalité », « nuptialité », « morbidité » et ainsi de suite. Le fait est que les phénomènes collectifs dont la statistique s'occupe dépendent d'une multiplicité de facteurs, et que, suivant les différents cas, on peut avoir intérêt à éliminer l'un ou l'autre d'entre ces facteurs, ou à en éliminer plusieurs, ou à les éliminer tous sauf un, ou bien à n'en éliminer aucun. La possibilité de conclusions contradictoires, qui dérive de conceptions divergentes du phénomène examiné, existe donc dans toutes les recherches scientifiques, mais elle a une importance particulière dans les recherches statistiques.

D'autre part, nous avons supposé jusqu'ici que nous disposions de toutes les données nécessaires pour éliminer l'influence des facteurs qu'il nous importait d'exclure. Mais, pratiquement, c'est tout autre chose.

Parfois c'est la nature même du phénomène qui fait que les données sont insuffisantes dans ce but, comme c'est le cas lorsqu'il s'agit d'éliminer l'influence du climat. Pour cela nous devrions disposer de données sur la mortalité du pays A relatives à des années dans lesquelles le climat y a été exceptionnellement semblable à celui qui est normal dans le pays B ; et, réciproquement, de données sur la mortalité du pays B dans des années dans lesquelles le climat y a été exceptionnellement semblable à celui qui est normal dans le pays A. Mais le seul fait que cela ne peut avoir lieu qu'exceptionnellement rend difficile d'obtenir une masse de données suffisante, car il faut se rappeler que, s'agissant de phénomènes statistiques, non seulement il est nécessaire d'avoir des données, mais il faut avoir des données assez nombreuses pour que l'on soit autorisé à considérer comme neutralisées les circonstances perturbatrices de caractère accidentel.

Pour l'élimination d'autres facteurs, la nature peut bien nous offrir une expérience assez large, mais les relevés statistiques peuvent être incomplets. Quel est le pays qui permet de classifier les morts et les vivants suivant les combinaisons d'âge, de sexe, d'état civil, de profession et de richesse, de sorte qu'il soit possible d'éliminer à la fois l'influence de tous ces facteurs des coefficients de mortalité ? Pas un que je sache. Le statisticien doit donc se contenter, dans ce cas, d'éliminer l'influence de quelques-uns des facteurs qu'il voudrait exclure, en parvenant ainsi à des conclusions approchées. C'est même là la règle.

D'autres fois il arrive que le matériel statistique permet bien l'élimination de facteurs déterminés mais seulement sur le fondement de certaines hypothèses, valables rien qu'approximativement ; par exemple, sur le fondement de l'hypothèse de la superposition des effets. C'est l'hypothèse qui est à la base de la méthode des corrélations partielles, à laquelle le matériel statistique permet d'avoir recours dans certains cas dans lesquels il n'est pas possible, par contre, d'appliquer les procédés de la population type, dont nous avons parlé, ou le procédé symétrique des coefficients types.

D'autres fois, enfin, le matériel statistique ne permet pas d'éliminer tous les facteurs γ, δ, ε, mais seulement les deux facteurs γ et δ, ou bien δ et ε, ou bien γ et ε. Il est loisible d'éliminer un quelconque de ces couples de facteurs. Alors, il y a celui qui préfère en éliminer un et celui qui préfère en éliminer un autre. La conclusion sera toujours approchée, mais l'approximation pourra se réaliser en un sens divers selon que l'un ou l'autre couple aura été éliminé. Dans l'exemple cidessus, si l'on élimine l'âge, le sexe, l'état civil et la profession, la mortalité apparaît plus élevée dans le pays B que dans le pays A ; si l'on élimine l'âge, le sexe, l'état civil, la profession et la richesse, elle apparaît plus élevée dans le pays A que dans le pays B. Or, supposons que la classification ne dispose pas des données nécessaires pour permettre d'éliminer à la fois, outre l'âge, le sexe et l'état civil, également la richesse et la profession. Alors, celui qui élimine la richesse arrivera à des conclusions différentes de celles auxquelles parviendra celui qui élimine la profession. Dans l'impossibilité d'éliminer tous les facteurs de perturbation, c'est à l'intuition statistique du chercheur qu'est remise l'élimination de l'un ou de l'autre facteur. Ce cas se présente dans un très grand nombre de recherches statistiques, je dirais même dans toutes ces recherches et c'est là que réside la raison fondamentale pour laquelle l'intuition a une si grande importance dans la statistique appliquée.

Il est bien vrai que la possibilité existe d'éliminer, l'un après l'autre, tous les groupes de facteurs éliminables : par exemple, d'abord les deux facteurs γ et δ et puis les deux facteurs δ et ε, et enfin les deux facteurs γ et ε. Cela montrerait que les données disponibles ne permettent pas d'arriver à une conclusion certaine, mais cela éviterait qu'une conclusion considérée certaine fût contredite ultérieurement. Cependant, il faut tenir compte des calculs excessivement laborieux que cette recherche exige et qui seraient souvent en contraste avec le temps et les moyens disponibles.

En somme, l'insuffisance des données que la nature elle-même nous fournit et, non moins souvent, l'insuffisance des relevés statistiques, empêchent ordinairement que la statistique arrive à des conclusions certaines et, en s'ajoutant au peu de temps et de moyens disponibles, elles laissent un vaste champ à l'intuition personnelle qui conduit inévitablement les divers chercheurs à des conclusions divergentes. C'est ce qui arrive dans tous les champs de recherche où le résultat envisagé dépend d'une multiplicité complexe de facteurs impossibles à encadrer objectivement : il en est ainsi dans la médecine, dans les beaux-arts, dans l'art militaire, et — pouvons-nous dire — dans tous les arts. La large part faite à l'intuition est justement ce qui distingue les arts des sciences. La statistique appliquée est encore en grande partie, et sera toujours en une certaine mesure, un art, pour le succès duquel l'intuition, ou si l'on préfère le génie, du chercheur joue et jouera toujours un rôle important. Certains pourront considérer ce caractère de la statistique comme un avantage, d'autres comme un désavantage, par rapport aux autres sciences. Il nous explique, en tous cas, une grande partie des divergences et des contradictions que l'on reproche à ses spécialistes. Celui qui, d'un point de vue exclusivement scientifique, juge ce caractère désavantageux, ne doit pas perdre de vue, d'autre part, que les résultats obtenus par la statistique, bien qu'ils soient souvent approchés et incertains, ne pourraient généralement pas être atteints par d'autres méthodes. La statistique est une science marginale qui, avec les données et les moyens disponibles, arrive là où les autres sciences ne peuvent aboutir. Parfois les autres sciences y arrivent plus tard, alors que les données expérimentales se sont accrues et les moyens techniques ont été développés. En ce cas, la statistique fait principalement fonction d'éclaireur, fonction toujours dangereuse mais précieuse. Telle a été, par exemple, la fonction de la statistique pour l'étude de la variabilité et de l'hérédité des caractères. Dans d'autres domaines, les autres sciences n'arrivent jamais. La statistique exploite alors des terrains qui ne seraient autrement jamais exploités. Parfois, il s'agit — il est vrai — de terrains d'un rendement inférieur, et l'on comprend ainsi que, à l'égard de certains sujets que l'on rencontre, par exemple, dans l'économie et dans la biologie, les résultats de la statistique paraissent d'importance secondaire en comparaison de ceux que l'on peut obtenir par d'autres méthodes. Mais, d'autres fois, c'est à la statistique que les constructions scientifiques sont redevables des finissages qui les font apprécier, ainsi qu'il est advenu en ce qui concerne de nombreuses branches des sciences physiques, biologiques et sociales. Sans la statistique, elles seraient souvent obligées de contenir leur conclusions dans le champ qualitatif ou de se contenter d'approximations insuffisantes.

*
* *

Voyons un autre exemple où la statistique semble avoir conduit à des résultats contradictoires.

En classifiant dans un même pays les diverses circonscriptions territoriales d'après les indices de la prospérité économique (revenu, richesse, épargne), on constate que la natalité apparaît plus basse dans les parties les plus prospères du territoire. Aussi la comparaison entre pays riches et pays pauvres montre généralement une natalité plus basse dans les pays les plus riches. Et une même population présente ordinairement une natalité plus élevée au début de son histoire que dans les périodes suivantes, quand son organisation économique s'est développée et la richesse accumulée s'est accrue. Enfin, dans de nombreux pays les classes sociales supérieures présentent une natalité plus basse que celle de l'ensemble de la population. On en tire la conclusion que le bien-être économique est un facteur défavorable à la reproduction.

La comparaison entre le nombre des naissances dans des mois et des années successifs a montré, d'autre part, que la natalité baisse au cours des périodes de crise économique et croît au cours des périodes de prospérité. Dans une même classe sociale, les familles disposant de moyens économiques plus larges ont, en général, plus d'enfants que les autres. Cela a amené d'autres chercheurs à la conclusion contraire que le bien-être économique favorise la reproduction.

Enfin en comparant le nombre des enfants des familles des contribuables classifiés d'après leur revenu, il est apparu que ce nombre croît à mesure que le revenu croît jusqu'à un certain point, après quoi il diminue, ce qui a donné lieu à la conclusion intermédiaire que le bien-être économique favorise la reproduction jusqu'à un certain point, au-dessus duquel il l'entrave.

De semblables divergences donnent beau jeu aux détracteurs de la statistique.

Un examen soigneux montre cependant que, en réalité, elles ne sont, en partie tout au moins, qu'apparentes.

Pour les savants qui s'occupent de faits biologiques, la constatation n'est certes par neuve que les réactions de l'organisme humain à l'action momentanée d'un certain facteur peuvent être l'opposé de celles que provoque une action prolongée de celui-ci. C'est même la règle en ce qui concerne les réactions provoquées par les excitants. Cela peut expliquer, du moins jusqu'à un certain point, le fait que, lorsqu'a lieu une variation de conditions économiques, la natalité réagit de façon différente selon qu'il s'agit d'oscillations mensuelles ou annuelles ou, par contre, de modifications évolutives à longue échéance ou de différences permanentes qui se manifestent selon le territoire ou la classe sociale.

C'est aussi une observation générale que, en ce qui concerne de très nombreux phénomènes physiques, biologiques et sociaux, les effets ne sont pas proportionnels à leur intensité et parfois, selon le niveau initial ou selon l'intensité des variations, ils diffèrent aussi en ce qui concerne leur sens. Pour la température, l'alimentation, les exercices physiques, nous en faisons l'expérience chaque jour. Le principe de la superposition des effets — largement admis dans les sciences physiques et qui dans la statistique, est à la base, ainsi que je l'ai rappelé, de la méthode des corrélations partielles et, plus généralement, de la méthode des résidus — est généralement admissible en tant que première approximation, mais seulement dans de brefs intervalles de la variable. Il n'y a donc pas lieu de s'étonner si les variations de la reproduction humaine sont différentes quant à leur intensité et à leur sens, alors que varie le bien-être économique. Cela peut expliquer, par exemple, le fait que, tandis que la natalité baisse d'autant plus que la classe sociale est plus élevée, on constate que parmi les classes inférieures le nombre des enfants est moins élevé chez les catégories économiques les plus basses.

Dans ces cas aussi, l'observateur attentif apercevra que les contrastes dérivent du fait que, en réalité, les expressions « bien-être économique » et « amélioration économique » sont des expressions vagues qu'il est nécessaire de préciser pour ce qui concerne soit la durée du bien-être, soit son niveau initial et l'intensité de ses variations. Des inconvénients semblables peuvent évidemment avoir lieu, et ont lieu en effet, dans tous les champs de la recherche scientifique et ne sont pas particuliers à la statistique.

*
* *

D'autre part, cependant, les considérations ci-dessus n'expliquent pas complètement les contradictions susmentionnées. En ce qui les concerne, il y a évidemment un autre facteur qui intervient.

Lorsque nous comparons des données relatives à des mois, à des années, à des circonscriptions territoriales, à des périodes évolutives, à des catégories de revenu, à des classes sociales et, dans la même classe sociale, à des rangs hiérarchiques qui diffèrent quant à leur bien-être économique, nous considérons des modalités qui, en réalité, ne diffèrent pas par ce seul caractère. Des mois et des années de crise diffèrent souvent de mois et d'années prospères aussi en ce qui concerne les conditions climatiques et, généralement, en ce qui concerne la mortalité. Les contribuables des différentes catégories de revenu, ainsi que les contribuables appartenant aux différents rangs hiérarchiques d'une même classe sociale, diffèrent quant à l'âge. Les circonscriptions territoriales, riches et pauvres, diffèrent en ce qui concerne la densité de la population, le degré d'urbanisation, les occupations professionnelles. Les différentes classes sociales et les populations qui se trouvent dans des stades d'évolution différents diffèrent par de multiples aspects psychologiques : par exemple, en ce qui concerne la connaissance des moyens anticonceptionnels. Jusqu'à quel point les relations observées entre les variations de la natalité et les variations du bien-être économique selon les mois, les années, les circonscriptions territoriales, les classes sociales, les rangs hiérarchiques, les catégories de revenu, les stades évolutifs, dépendent-elles effectivement du bien-être économique et non d'autres facteurs différentiels qui parfois accentuent, parfois diminuent ou neutralisent et parfois suppriment l'influence des variations du bien-être économique?

Des difficultés semblables dérivent du fait que les phénomènes que l'on examine dépendent d'un ensemble de variables nombreuses et difficiles à isoler complètement; elles se manifestent dans tout champ de la recherche scientifique, mais elles sont plus fréquentes, plus graves et moins facilement éliminables dans le champ de la recherche statistique, surtout lorsque celle-ci s'occupe de phénomènes collectifs, sociaux ou biologiques, relatifs à l'espèce humaine et à l'égard desquels l'expérimentation n'est pas possible.

* * *

En conclusion, il n'y a rien de surprenant si avec la statistique on arrive parfois à des conclusions divergentes. Mais cela ne met pas en cause le bien-fondé de ses méthodes, pas plus que le bien-fondé des procédés de la logique n'est mis en cause du fait que chaque jour avec la logique, on arrive à des conclusions discordantes.

Il est vrai aussi que les qualifications elliptiques des données statistiques, l'insuffisance de celles-ci, la complexité des phénomènes collectifs, la multiplicité et la variété de leurs effets, l'impossibilité pratique de pousser le calcul laborieux et coûteux au-delà de certaines limites, la place qui, par conséquent, est laissée à l'intuition individuelle rendent les divergences entre les conclusions des différents chercheurs plus faciles dans le domaine des applications statistiques que dans celui d'autres recherches scientifiques. Cela prouve seulement que la statistique doit faire face à des difficultés spéciales, mais non pas qu'elle est moins puissante que les autres moyens d'investigation. Au contraire, aucun autre moyen ne saurait pénétrer où elle pénètre. Les dangers qui intéressent son chemin pourraient décourager plusieurs ; ils pourront représenter pour d'autres une attraction. *Militia est vita hominum super terram :* « La vie des hommes sur la terre est un combat », disait saint Paul et Nietzsche proclamait que le secret de la plus grande productivité et des plus grandes satisfactions de la vie s'appelle *gefährlich leben :* « vivre dangereusement ».

CHAPITRE II.

DE L'AUGMENTATION DU NOMBRE DES OBSERVATIONS ET DE SES EFFETS SUR LES CONSTANTES STATISTIQUES.

Le fait que la statistique implique une masse d'observations et qu'en augmentant cette masse, on réduit l'influence probable des erreurs accidentelles sur la moyenne des valeurs observées, a entraîné la conclusion que l'élimination des erreurs accidentelles par l'augmentation du nombre des observations est précisément le procédé qui caractérise la statistique. Cette conclusion est tout à fait injustifiée. Les recensements, les statistiques de l'état civil, les évaluations des récoltes, des échanges, des revenus, des fortunes, qui sont certainement des opérations statistiques de la plus grande importance et bien d'autres, n'ont nullement en vue d'éliminer des erreurs accidentelles. Dans bien des recherches, l'élimination des erreurs systématiques des résultats de la statistique est un élément bien plus important que l'élimination des erreurs accidentelles. Mais il convient surtout de remarquer que l'élimination des erreurs accidentelles ainsi que celle des erreurs systématiques, n'est qu'une fonction réflexe de la recherche statistique; c'est une espèce de garantie, un moyen d'assurer la précision désirable des résultats — fonction sans doute importante, mais secondaire en comparaison de son but principal. Son but principal est l'élaboration des données propres à mesurer quantitativement certains aspects des phénomènes collectifs : intensité, variabilité, distribution, relations mutuelles. Prendre un train qui soit exact à l'heure est certes bien important, mais plus important encore est le fait de prendre un train qui conduise à la destination voulue et par la route désirée.

* * *

Si l'élimination des erreurs accidentelles par l'augmentation du nombre des observations n'est pas le seul but, ni même le but principal, de la statistique, elle n'en est pas moins un but très important sur lequel il vaut bien la peine de se pencher attentivement.

A ce point de vue, la méhode statistique est généralement opposée à la méhode expérimentale [17]. Dans celle-ci, on se propose de mettre en évidence les effets des variations d'une circonstance (ou d'un groupe de circonstances), en éliminant les effets des variations de toutes les autres, qui sont maintenues constantes. Avec la méthode statistique, au contraire, étant donnée l'impossibilité de maintenir constantes les autres circonstances, on examine les effets de la variation de l'une d'entre elles, tandis qu'en même temps une foule d'autres agissent. On suppose

---

[17] Cf. *Alle basi del metodo statistico, Il principio della compensazione degli errori accidentali e la legge dei grandi numeri*, dans *Metron*, t. 14, nos 2-4, 1941, p. 173-175.

alors que ces dernières peuvent être considérées comme accidentelles vis-à-vis de la circonstance prise en considération et que, par conséquent, leurs effets sont neutralisés par leur grand nombre.

Je traiterai en détail des rapports entre la méthode statistique et la méthode expérimentale dans le chapitre III. Ici je me bornerai à remarquer que la nette opposition que l'on envisageait autrefois entre les deux méthodes a été considérablement réduite. D'un côté, on a reconnu qu'il n'est pas facile de maintenir absolument constantes, dans toute expérience, toutes les circonstances (hormis celle que l'on veut isoler); cela est vrai en particulier pour les expériences agricoles faites au grand air, où les conditions atmosphériques ainsi que celles du sol sont difficilement égales pour toutes les parcelles de terrain. Il faut alors choisir des parcelles qui se compensent — ce qui ne peut avoir lieu qu'approximativement — ou bien s'en remettre à la compensation statistique, ce qui est d'ailleurs inévitable pour les conditions atmosphériques. D'un autre côté, il arrive fréquemment que dans les observations successives, les effets de la circonstance isolée ne présentent pas de différences marquées, et pour juger si celles-ci sont significatives on a recours à certaines méthodes statistiques dont nous parlerons par la suite.

Somme toute, on devrait parler aujourd'hui d'une méthode statistico-expérimentale par opposition à une méthode statistico-observationnelle, plutôt que d'une méthode expérimentale opposée à une méthode statistique.

La différence entre les deux méthodes se réduit notablement dans certains cas où les données observées permettent d'isoler une circonstance, sinon de toutes les autres, du moins des plus importantes d'entre elles, avec une rigueur qui ne reste pas trop au-dessous de celle de l'expérimentation. Cela se produit par exemple, dans l'étude de l'importance de l'hérédité et du milieu à partir des données concernant les jumeaux monozygotes ou dizygotes.

On en conclut qu'aujourd'hui le principe de l'élimination des erreurs accidentelles n'est pas seulement à la base de maintes applications de la statistique aux faits observés, tels qu'ils se présentent dans la nature ou dans la société, mais encore qu'il est souvent à la base de ses applications aux données expérimentales. Cette constatation ne peut évidemment qu'augmenter l'intérêt d'un examen détaillé de son fondement.

*
* *

Nous devons d'abord chercher ce qui justifie notre confiance dans le fait que les erreurs accidentelles sont éliminées, ou tout au moins tendent à être éliminées, lorsque l'on augmente le nombre des observations.

A ce sujet, il est utile de distinguer deux catégories d'erreurs accidentelles : celles qui proviennent du nombre limité des observations — nous les appellerons *erreurs de fréquence* — et celles qui proviennent d'inexactitudes dans la mesure, l'estimation ou l'énumération, que nous appellerons *erreurs de dimensions*. Les premières se produisent seulement dans les collectifs, alors que les dernières affectent d'abord les grandeurs individuelles.

Or, la justification en question est différente pour les deux catégories d'erreurs. Pour les erreurs de dimension, on fait appel au *principe de la compensation des erreurs accidentelles;* pour les erreurs de fréquence, *à la loi des grands nombres.*

*
* *

Commençons par la loi des grands nombres ([18]). Elle a été formulée par Poisson, qui l'a déduite de l'extension du théorème bien connu de Jacques Bernoulli d'après lequel la fréquence relative d'un événement tend, avec l'augmentation du nombre des observations vers sa probabilité, et l'atteindrait même, si le nombre des observations était infini. Bernoulli supposait cette probabilité constante; Poisson démontra que, si elle varie irrégulièrement d'une observation à l'autre, le théorème garde sa validité pour la probabilité moyenne.

C'est un théorème remarquable. Mais un théorème est toujours un théorème, c'est-à-dire une proposition théorique, basée sur des hypothèses. La loi, au contraire, exprime un rapport constant, c'est-à-dire constaté maintes fois par des expériences. Ce sont des choses bien différentes. Il y a des hommes de science — tel le Professeur Marbe — qui, sur la base d'observations étendues, soutiennent que, même pour les jeux de hasard, les prévisions du Calcul des Probabilités

([18]) *Voir*, pour ce qui suit à ce sujet, *Alle basi del metodo statistico*, art. cité, p. 222-231.

ne se vérifient pas dans la pratique. Et c'est là une question que l'on ne peut résoudre que sur la base des faits. D'ailleurs, comment établir rigoureusement une loi dont la vérification nécessite un nombre infini d'observations? Il est impossible de la vérifier même une seule fois, tandis que, pour l'admettre, il faudrait l'avoir vérifiée maintes fois. On doit donc conclure à l'impossibilité théorique et pratique de démontrer la loi des grands nombres.

Malgré cela, il est certain que tout le monde a confiance en elle. Comment expliquer ce paradoxe? Je crois l'avoir expliqué par le fait que en réalité, nous sommes intéressés toujours, ou presque toujours, par un nombre fini d'événements et non par des univers que l'on dit infinis. La probabilité d'un événement dans une certaine masse de ces événements n'est que sa *fréquence totalitaire* — ainsi que je l'ai appelée — c'est-à-dire sa fréquence dans le total des cas qui nous intéressent, total qui peut être un nombre très grand, et aussi un nombre indéfini, dans le sens que nous ne connaissons pas le nombre de ces cas, qui reste pourtant toujours un nombre fini. Or, il est évident que lorsque l'on augmente le nombre des cas considérés, la fréquence observée s'approche, dans la moyenne des cas, de la probabilité ainsi conçue jusqu'à coïncider avec elle lorsque l'on considère le total des cas qui entrent dans la masse. C'est dans ce sens que la confiance en la loi des grands nombres est justifiée.

Il y a un autre point, dans la formulation de la loi des grands nombres, qu'il faut préciser : Poisson lui-même n'a pas été toujours précis à ce sujet, tandis qu'il est très important de l'être, surtout en vue des applications pratiques de la loi.

La loi des grands nombres — ainsi que nous l'avons dit — est une généralisation du théorème de Bernoulli aux cas où les probabilités ne sont pas constantes, mais varient *irrégulièrement*, voire *accidentellement*. Celle-ci paraît bien avoir été la pensée de Poisson, formulée dans le préambule de son ouvrage ([18 bis]) où il parle de « causes qui varient irrégulièrement, tantôt dans un sens, tantôt dans l'autre, c'est-à-dire sans que leur variation soit progressive dans aucun sens déterminé » (p. 7). Elle correspond aussi aux énoncés des théorèmes donnés ensuite aux pages 138 et suivantes. Mais, au contraire, dans la Table des Matières,

([18 bis]) D. S. POISSON, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, Paris, Bachelier, 1857.

très détaillée, soit à propos de ces énoncés (p. III), soit ensuite dans le titre du Chapitre IV (p. V), on parle de « chances variables d'une manière quelconque », ce qui serait bien différent. Peut-être, la Table des Matières n'a-t-elle pas été rédigée par Poisson lui-même? En tout cas il est important d'établir que l'on ne peut parler de la loi des grands nombres que lorsque la probabilité de l'événement varie irrégulièrement, c'est-à-dire accidentellement au cours des observations.

Or, on peut penser que cette condition se réalise dans les jeux de hasard aussi bien que dans les expériences où l'on s'efforce de maintenir constantes toutes circonstances en dehors de celle dont on veut établir l'influence; mais, dans l'observation des faits tels qu'ils se présentent dans la pratique, on a bien raison de douter que la condition soit réalisée. On peut tirer plusieurs fois d'une urne la même boule ou des boules équivalentes aux fins du résultat et les tirer dans des conditions qui ne varient pas régulièrement d'un cas à l'autre, mais on ne peut pas faire naître ou mourir une personne plus d'une fois, ni être certain que les conditions dans lesquelles une autre personne vient à naître ou à mourir ne diffèrent des conditions précédentes que par des circonstances accidentelles. Lorsque, dans l'observation des faits d'un phénomène qui fait l'objet de notre recherche, nous disposons d'un nombre de cas que nous jugeons restreint et nous nous proposons en conséquence d'augmenter ce nombre, il nous faut toujours nous rendre compte qu'il est probable — j'inclinerais même à dire qu'il est certain — que les conditions dans les observations à faire seront plus ou moins différentes — et systématiquement différentes — de celles des observations déjà faites, de façon que le résultat basé sur le nombre des cas augmenté concernerait en réalité un collectif systématiquement différent du collectif qui a formé l'objet de notre recherche.

Nous sommes de la sorte dans l'alternative de nous contenter des résultats obtenus, affectés d'erreurs accidentelles de fréquence, à cause du nombre trop limité de cas qui se présentent, ou bien d'éliminer cet inconvénient en tombant dans l'autre erreur qui consiste à comprendre dans nos observations des cas différant systématiquement de ceux qui concernent le phénomène qui nous intéresse.

Quelle est la probabilité que M^me^ X donne le jour à un garçon? Elle a eu jusqu'à présent deux garçons et une fille. Trois est évidemment un nombre trop limité pour être la base d'une prévision raisonnable; mais, si, pour augmenter le nombre des cas, nous prenons en considération les enfantements d'autres femmes, nous devons compter sur le fait que, sauf un hasard, celles-ci ne peuvent avoir exactement la même probabilité que M^{me} X d'enfanter un garçon ou une fille. Il faut choisir entre les deux inconvénients. Tout cela est évident. Mais ce qui est important, pour se rendre compte de la portée de notre considération, est, qu'en réalité, les conditions changent même lorsque nous ne nous en apercevons pas. Les résultats de certaines recherches expérimentales, poursuivies pendant plus de 30 ans sur l'interpolation et l'extrapolation d'abord à l'Université de Padoue et après à celle de Rome, sont instructifs à ce sujet. Permettez-moi d'en rappeler les résultats.

Si nous voulons interpoler ou extrapoler un terme d'une série statistique, nous pouvons naturellement le faire en nous basant sur un nombre variable de termes; le procédé le plus simple serait d'interpoler une droite sur les deux termes contigus, dans le cas d'interpolation, ou sur les deux termes précédents dans le cas d'extrapolation. Mais qui songerait jamais à se baser sur un nombre de termes si restreint? Bon, mais, si à la place d'une droite vous interpolez une parabole de deuxième ordre sur trois termes, ou de troisième ordre sur quatre termes, ou d'un nombre plus élevé sur un nombre de termes plus grand, loin d'obtenir des prévisions plus exactes, vous vous exposez à des inexactitudes plus fortes, qui tendent à augmenter avec le nombre des termes considérés. Maintes applications à des phénomènes variés dans les domaines démographiques et économiques ont été faites par divers chercheurs et toujours avec le même résultat. C'est qu'avec le temps les conditions dans lesquelles les phénomènes se déroulent changent systématiquement sans que nous nous en apercevions et cette influence perturbatrice systématique s'avère plus importante que l'influence perturbatrice accidentelle du nombre restreint des termes pris en considération. Que l'explication que nous venons de donner soit fondée est confirmée par des recherches de contrôle. Supposons, par exemple, avoir sept termes et vouloir interpoler le terme central, c'est-à-dire le 4^e que nous supposons inconnu. Si, ainsi que nous le disions, nous interpolons une droite sur les 3^e et 5^e termes, nous obtenons en général, une approximation meilleure qu'en interpolant une parabole de troisième ordre sur les 2^e et 3^e, 5^e et 6^e termes. C'est — nous l'expliquons — que les conditions dans lesquelles le phénomène se déroule, ont varié du 3^e au 5^e terme moins que du 2^e au 6^e. Mais si nous interpolons une droite sur les 2^e et 6^e termes et une parabole de troisième ordre sur les 1er, 2^e, 5^e et 7^e, l'approximation obtenue n'est pas, en général, meilleure avec la droite, mais plutôt avec la parabole. On peut l'expliquer en considérant que, dans ce cas, on ne peut pas penser que du 2^e au 6^e terme les conditions aient varié moins que du 1er et 3^e au 5^e et 7^e termes ([19]).

* * *

Passons au principe de la compensation des erreurs accidentelles de dimension ([20]). Sa portée dépend de la définition de l'expression « erreurs accidentelles ». Si, comme quelqu'un l'a fait, on définit les erreurs accidentelles comme les erreurs qui se compensent, ce principe est sans valeur, car il se réduit à une pure tautologie. Il a, au contraire, un sens si l'expression « erreurs accidentelles » est entendue comme une abréviation de l'expression « erreurs dépendant de circonstances fortuites » ou, si vous préférez, de circonstances « accidentelles » en entendant par là des circonstancès qui tendent à se compenser. Il faut rappeler à ce sujet la distinction de Quetelet entre causes constantes, variables (c'est-à-dire systématiquement variables) et fortuites ou accidentelles, ces dernières pouvant apparaître ou non et lorsqu'elles apparaissent le faisant avec une intensité indépendante des causes systématiques (constantes ou variables) de telle façon que la valeur probable de leurs intensités soit nulle.

Or, peut-on déduire que la valeur probable de leurs effets, c'est-à-dire des erreurs, est également nulle ? Voilà la question.

La réponse, sauf dans des cas particuliers, est évidemment négative, car la valeur probable des effets n'est pas nécessairement nulle. Elle peut l'être sous certaines hypothèses, par exemple :

a. si les effets sont proportionnels aux causes fortuites ou, plus généralement, si la relation entre causes et effets est linéaire;

([19]) Les dernières recherches sur le sujet sont exposées dans les publications suivantes du Professeur E. Pizzetti qui a aussi résumé les résultats de recherches précédentes : *Verifiche sperimentali sull'interpolazione* (Communication présentée à la IX^e Réunion scientifique, janvier 1950, de la Société italienne de Statistique, publiée dans les *Actes* de ladite Société, Rome, 1952); *Indagini sperimentali sull'interpolazione* (Communication présentée àl a XXVIII^e Session de l'Institut Int. de Statistique, Rome, 1954).

([20]) *Voir*, pour ce qui suit à ce sujet, l'article cité, *Alle basi del metodo statistico*, p. 186-210.

b. si la courbe de fréquence des causes est symétrique et si la relation entre causes et effets, bien que non linéaire, est égale pour des valeurs symétriques des causes;

c. si la relation entre causes et effets est monotone et si la médiane coïncide avec la moyenne arithmétique tant pour les causes que pour les effets;

d. si la moyenne arithmétique des effets coïncide avec leur mode et si l'intensité des effets est en relation inverse avec leur fréquence.

Il n'y a évidemment aucune raison de penser que l'une ou l'autre de ces hypothèses doive nécessairement se réaliser. On doit conclure que le principe de la compensation, ou mieux de la tendance à la compensation des erreurs accidentelles ne peut pas être admis comme un principe ayant une validité générale; il faudra examiner dans chaque cas si des hypothèses suffisantes justifient un tel principe même si elles ne se réalisaient que d'une façon approximative.

Si les circonstances fortuites tendent à se compenser, elles ne sont pas les seules à jouir de cette propriété. C'est là une propriété plus générale, dont jouissent les déviations de la moyenne arithmétique. Le principe de la compensation des erreurs accidentelles de dimension n'est donc qu'un cas particulier du principe de la correspondance entre la moyenne arithmétique de la variable et la moyenne arithmétique de la fonction. C'est généralement le cas lorsque la variable est représentée par une circonstance fortuite qui affecte la mensuration d'une grandeur objective par les erreurs d'observations qui s'ensuivent. Mais il est évident qu'il n'est pas nécessairement vrai que la moyenne d'une fonction corresponde toujours à la moyenne de la variable. Il y a, par exemple, des phénomènes qui, pour ce qui concerne certains effets, présentent un optimum s'approchant parfois de la moyenne, telle dans un pays qui jouit d'un bon climat, la température pour ce qui concerne son influence sur la santé : les déviations positives de la température moyenne ne correspondent pas à des déviations positives, mais à des déviations négatives de la santé moyenne; la moyenne de la température ne correspond pas à la moyenne de la santé, mais correspond au maximum de la santé, ou tout au moins s'y approche.

Voyons la portée pratique de ces considérations.

Si le principe de la correspondance entre les moyennes arithmétiques de la variable et de la fonction était valable, on pourrait dire par exemple que la taille moyenne de la population adulte française serait la taille moyenne que l'on obtiendrait si tous les Français jouissaient de la même richesse, de la même nutrition, du même exercice, etc. correspondant à la richesse, à la nutrition, à l'exercice moyen de la population. En réalité, nous ne pouvons pas l'affirmer; nous pouvons seulement dire que richesse, nutrition, exercice, sont des circonstances qui exercent une influence sur la taille des Français. Il nous est impossible de l'affirmer, précisément parce que nous ne pouvons pas affirmer que les déviations de la moyenne arithmétique de la richesse, et de même de la nutrition, de l'exercice, etc. — déviations qui se compensent — déterminent sur la taille des effets qui se compensent eux aussi. Cela serait vrai si l'une des conditions a, b, c, d se réalisait, mais il peut n'être pas vrai dans le cas contraire.

Il y a une autre réserve à faire. Même si les déviations de la moyenne arithmétique se compensaient pour chacune des circonstances considérée séparément, il n'est pas dit que, lorsque lesdites circonstances agissent conjointement, l'ensemble de leurs déviations se compense aussi. Ce serait le cas si les effets desdites circonstances jouissaient de la propriété additive, c'est-à-dire si l'effet global de toutes les circonstances agissant conjointement était égal à la somme des effets desdites circonstances agissant séparément, mais on ne pourrait plus l'affirmer si la composition des effets suivait une loi différente. Ce ne serait pas du tout le cas, par exemple, si les effets défavorables, correspondant à des déviations négatives prédominaient sur les effets favorables, correspondant à des déviations positives. C'est ce qui s'avère pour les produits agricoles. En effet, la récolte dépend, outre que de la nature du sol, de maintes circonstances : température, précipitations atmosphériques et en particulier orages, fumage, main-d'œuvre, parasites, etc. Si une grêle dévastatrice ou une invasion de sauterelles survient, toutes les autres circonstances peuvent avoir été favorables au maximum, mais cela ne compense pas le dommage. D'une façon analogue, si, au moment de la récolte, les travailleurs chôment, le produit peut être perdu complètement sans que les autres circonstances, pour favorables qu'elles eussent été, puissent avoir aucune influence compensatrice. Il s'ensuit que le produit moyen de l'agriculture est sensiblement inférieur au produit que les agriculteurs considèrent comme normal, dans le sens du produit qui s'obtiendrait si toutes les dites circonstances se présentaient avec leur intensité moyenne : c'est que les déviations positives, causées par l'action de certaines circonstances favorables, ne compensent pas les déviations négatives causées par d'autres circonstances défavorables.

*
* *

Nous avons considéré jusqu'à présent la question de l'élimination, moyennant l'augmentation du nombre des observations, des effets que les erreurs accidentelles de dimension et de fréquence exercent sur l'intensité moyenne des phénomènes statistiques. C'est de cette façon que la question est traditionnellement abordée : c'est une façon de l'envisager qui se ressent d'un stade dépassé dans l'évolution des méthodes statistiques, alors qu'elles étaient justement limitées à l'étude de l'intensité des phénomènes collectifs. Depuis lors la méthodologie statistique s'est développée et à l'étude de l'intensité se sont ajoutées celles de la variabilité et des relations mutuelles des phénomènes collectifs.

Or, quels sont les effets des erreurs accidentelles de dimension et de fréquence sur ces autres aspects des phénomènes statistiques ? et, à ce point de vue, quelles sont les conséquences d'une augmentation des observations ?

Il est nécessaire ici de distinguer les différents collectifs pour lesquels le problème peut se poser.

Je rappellerai à ce sujet la distinction courante entre *moyenne objective*, lorsque les termes sur lesquels elle est établie correspondent à la même grandeur, dont nous avons des expressions différentes à cause d'erreurs de mensuration, d'évaluation ou d'énumération (telles les diverses évaluations de la richesse d'un pays données par des auteurs différents) et *moyenne subjective*, lorsque les termes correspondent à des grandeurs réellement différentes (telles les tailles des différentes personnes). En harmonie avec cette distinction, nous appellerons *collectifs objectifs* ceux qui donnent lieu à une moyenne objective et *collectifs subjectifs* ceux qui donnent lieu à une moyenne subjective.

Considérons d'abord quatre types de collectifs que l'on pourrait appeler des *collectifs de premier ordre*.

a. Ensemble des observations répétées de la même grandeur et affectées d'erreurs de dimension.

Par exemple, déterminations répétées de la hauteur d'une tour ou de la force musculaire de la main d'une personne. Nous les appellerons *collectifs objectifs* (car ils donnent lieu à des moyennes objectives) *affectés d'erreurs de dimension.*

b. Ensemble d'observations non répétées, affectées d'erreurs de dimension, concernant les grandeurs individuelles qui composent la masse.

Par exemple, taille mesurée par les médecins militaires, des conscrits italiens, nés dans une certaine année. Nous les appellerons *collectifs subjectifs* (car ils donnent lieu à des moyennes subjectives) *affectés d'erreurs de dimension.*

c. Ensemble d'observations exactes non répétées concernant une partie des grandeurs qui composent la masse ou univers.

Par exemple, années de naissance inscrites sur les registres de l'état civil d'un certain nombre de personnes choisies au hasard dans la population, c'est-à-dire d'après un caractère indépendant de l'année de naissance de la personne (par exemple, personnes nées le jeudi). Nous les appellerons *collectifs subjectifs affectés d'erreurs de fréquence.*

d. Ensemble d'observations non répétées, affectées d'erreurs de dimension, concernant une partie des grandeurs qui composent la masse ou univers.

Par exemple, les mesures prises par les médecins militaires, de la taille des conscrits italiens qui sont nés le jeudi. Nous les appellerons *collectifs subjectifs affectés d'erreurs de fréquence et de dimension.*

Dans tous ces cas, les erreurs de fréquence, provenant du fait que l'on considère seulement un échantillon, tendent à réduire la *variance* du collectif — mesurée à partir de la moyenne des valeurs observées — tandis que les erreurs de dimension tendent à l'augmenter. Avec le mot « tendent », j'entends qu'ils l'augmentent ou la diminuent dans la moyenne des échantillons, non nécessairement dans un échantillon individuel.

Il est en effet bien connu que, si v_u est la variance de l'univers,

$$\overline{V}_e = \frac{n-1}{n} v_u < v_u$$

sera la valeur probable de la variance d'un échantillon comprenant n observations. Il s'ensuit qu'en augmentant le nombre n des observations, la variance des valeurs observées à partir de la moyenne arithmétique desdites valeurs, tend à augmenter et l'influence des erreurs de fréquence sur la-dite variance, tend à se réduire (à la limite à s'effacer).

Au contraire, l'influence des erreurs de dimension tend à augmenter. En effet, quel que soit le nombre des observations, la valeur probable de la variance $\bar{V}_{a'}$ des valeurs affectées par des erreurs de dimension peut être subdivisée en deux parties qui s'additionnent : la partie systématique V_a qui se réaliserait s'il n'y avait pas d'erreurs de dimension et la partie accidentelle $\bar{V}_e$ qui est due eux erreurs en question. On a donc $\bar{V}_{a'} = V_a + \bar{V}_e$, où $\bar{V}_e$ représente la valeur probable de la variance des erreurs de dimension à partir de leur moyenne, et par conséquent toujours $\bar{V}_{a'} > V_a$.

Les erreurs accidentelles de dimension qui affectent les n observations représentent un échantillon tiré de l'univers de toutes les erreurs de dimension possibles et, par conséquent, la valeur probable de leur variance $\bar{V}_e$ à partir de leur moyenne augmente avec n. Par conséquent aussi, la variance $\bar{V}_{a'}$ et de même la différence $\bar{V}_{a'} - V_a$, augmente avec le nombre des observations.

*
* *

Passons maintenant au cas moins simple d'un collectif, que l'on pourrait appeler *collectif de deuxième ordre*, constitué par l'ensemble des moyennes concernant des grandeurs différentes, chaque moyenne se référant aux observations répétées d'une même grandeur. Par exemple, pour établir au manomètre la force musculaire de la main d'une personne, on a la bonne habitude de faire plusieurs mensurations, en général trois, et de prendre leur moyenne. De même la tendance d'un couple à engendrer des garçons est mesurée en observant le sexe des enfants et en prenant la fréquence des garçons dans ceux-ci. Dans ces cas, l'ensemble des moyennes pourrait être regardé comme un collectif subjectif, mais chaque moyenne, à son tour, représente un collectif objectif.

Nous pourrons donc appeler le collectif de deuxième ordre en question un *collectif objectif-subjectif*. Il est toujours affecté par les erreurs de fréquence concernant chaque moyenne. Si les grandeurs observées ne constituent qu'un échantillon de la masse ou de l'univers, il sera également affecté par des erreurs de fréquence concernant directement le collectif des moyennes. Enfin, il peut être ou ne pas être affecté par des erreurs de dimension concernant les observations individuelles et par conséquent leurs moyennes. Dans le collectif des rapports des garçons aux naissances, ces erreurs de dimension n'interviennent pas, la discrimination entre un garçon et une fille étant sûre ; elles interviennent au contraire dans le collectif des moyennes de la force musculaire de la main, les mesures de celle-ci étant exposées aux influences perturbatrices des facteurs psychiques, de la préhension plus ou moins correcte, de la lecture du manomètre, etc.

Dans ces *collectifs de deuxième ordre*, il est nécessaire de bien préciser la signification de l'expression : *augmentation des observations*. D'une part, on peut en effet entendre la *répétition des observations sur la même grandeur;* d'autre part, lorsqu'il s'agit d'un échantillon, *l'extension de l'observation à d'autres grandeurs*.

La répétition des observations sur la même grandeur tend à réduire les erreurs de fréquence et de dimension concernant les moyennes, et par conséquent à produire une uniformisation des moyennes qui réduit la variance du collectif; l'extension de l'observation à d'autres grandeurs tend à réduire l'influence des erreurs de fréquence concernant directement le collectif et à accentuer celle des erreurs de dimension, en augmentant d'un côté et de l'autre la variance du collectif, mesurée à partir de la moyenne des valeurs observées.

*
* *

Considérons maintenant l'influence des erreurs de dimension et de fréquence et celle de l'augmentation du nombre des observations sur les relations entre les phénomènes.

Dans un des collectifs objectifs ou subjectifs considérés sous les lettres a, b, c, d, supposons avoir classé les valeurs observées d'après les modalités d'un autre caractère; par exemple, supposons avoir classé les valeurs obtenues pour la force musculaire ou la taille d'après la profession : nous disons qu'il y a connexion entre la force musculaire ou la taille et la profession si la distribution des valeurs obtenues pour la force musculaire ou la taille n'est pas la même pour les différentes professions.

La connexion atteindrait son maximum si, à chaque profession, correspondait une seule valeur de la taille ou de la force musculaire ; elle diminue avec l'augmentation de la variabilité des valeurs de la taille ou de la force musculaire correspondant à la même profession. Nous avons vu que les erreurs de fréquence tendent à réduire cette variabilité tandis que les erreurs de dimension tendent à l'augmenter. Il s'ensuit que les erreurs de fréquence tendent à augmenter la connexion tandis que les erreurs de dimension tendent à la réduire. Si, par exemple, l'on prend seulement un échantillon de taille, il est bien probable que des différences dues au nombre restreint des observations se vérifieront entre les distributions des différentes professions, même si en réalité la profession n'a aucune influence sur la taille, tandis que, si les tailles sont affectées par des erreurs accidentelles de mensuration, la connexion réelle entre profession et taille s'en trouve atténuée.

L'augmentation du nombre des observations, qui tend à réduire et, à la limite, à éliminer les erreurs de fréquence, tend par conséquent à atténuer la connexion observée en la portant vers sa valeur réelle, tandis qu'elle ne tend pas à éliminer, mais plutôt à augmenter l'influence des erreurs de dimension et partant ne tend pas à éliminer, mais plutôt à accentuer la réduction dans la connexion qui en résulte. De tous côtés, l'augmentation du nombre des observations tend donc à atténuer la mesure de la connexion.

Cela dans les collectifs de premier ordre, objectifs ou subjectifs.

Dans les collectifs de deuxième ordre, objectifs-subjectifs, les choses sont un peu plus compliquées.

L'extension de l'observation à d'autres grandeurs a des effets analogues à ceux que nous avons illustrés pour les collectifs de premier ordre, c'est-à-dire elle tend à atténuer la mesure de la connexion. Au contraire, la répétition des observations sur la même grandeur tend à réduire et, à la limite, à éliminer les erreurs de fréquence et — sous les conditions que nous avons mises en évidence — les erreurs de dimension concernant les moyennes et, par conséquent, à éliminer leur influence perturbatrice en faisant augmenter de la sorte la mesure de la connexion.

*
* *

Lorsque les modalités d'après lesquelles on a partagé le collectif des observations sont comparables avec les résultats des observations elles-mêmes, on peut se demander non seulement si les résultats des observations montrent une certaine connexion avec les modalités en question, mais également si cette connexion se manifeste dans le sens positif ou négatif, c'est-à-dire s'il y a concordance ou discordance entre les deux phénomènes ou caractères.

Par exemple, si l'on étudie le collectif des tailles des épouses classées d'après la taille des époux, on peut se demander, non seulement si la distribution des tailles des épouses est différente pour les diverses tailles des époux, mais aussi si les variations des tailles se vérifient dans le même sens ou bien dans le sens contraire pour l'une et pour l'autre.

Les erreurs de dimension exercent sur la concordance une influence analogue à celle qu'elles exercent sur la connexion, c'est-à-dire elles l'atténuent, tandis que les erreurs de fréquence n'exercent pas une influence dans un sens déterminé puisqu'elles peuvent aussi bien augmenter que réduire la concordance.

Ici aussi, comme dans la connexion, l'extension de l'observation à d'autres grandeurs tend à réduire dans les collectifs de deuxième ordre et, à la limite, à éliminer l'influence des erreurs de fréquence, mais non celle des erreurs de dimension, tandis que la répétition des observations sur la même grandeur tend à réduire et, à la limite, à éliminer l'influence des erreurs de fréquence ainsi que (sous certaines conditions) celle des erreurs de dimension concernant les moyennes.

On ne se rend pas toujours bien compte de l'influence des erreurs affectant les moyennes : par exemple, on croyait avoir démontré que la tendance à engendrer des enfants d'un sexe plutôt que de l'autre ne s'hérite pratiquement pas, en se basant sur la concordance très faible entre les rapports des sexes dans les familles des parents et dans celles des enfants. Mais on ne réalisait pas que les mesures de la tendance à produire un sexe, déduites du rapport des sexes dans les familles, étaient affectées d'erreurs de fréquence qui réduisaient la concordance. Après avoir corrigé le coefficient de corrélation, celui-ci prit une valeur analogue à celles que l'on avait obtenues pour les autres caractères physiques ([21]).

Les erreurs ont un effet spécial lorsqu'il s'agit de mesurer des segments proches l'un de l'autre ou la fréquence de deux événements dans des intervalles qui se suivent de près.

([21]) *Voir* le chapitre IX, *L'eredità del sesso* dans le volume cité *Il sesso dal punto di vista statistico*, p. 394 et suiv.

J'ai fait récemment deux expériences à ce sujet ([22]).

Je voulais établir le degré de concordance entre les parties du corps humain, par exemple bras et cuisse, bras et avant-bras, cuisse et jambe, avant-bras et main, hauteur du thorax et hauteur de l'abdomen, etc. Il fallait s'attendre à ce qu'une concordance se réalisât entre les deux segments et je me demandais si elle ne serait pas d'autant plus prononcée que ceux-ci étaient proches l'un de l'autre. En réalité, les coefficients de corrélation prenaient en général des valeurs positives ; mais des exceptions fréquentes se produisaient justement entre deux segments successifs (par exemple hauteur du thorax et de l'abdomen), les corrélations étant souvent négatives. Comment expliquer ce dernier résultat tout à fait inattendu? Les mensurations étaient faites sur le vivant, où les points de repère ne sont pas très faciles à déterminer, surtout lorsqu'on est obligé de travailler un peu à la hâte. Si vous prenez le point qui sépare le thorax de l'abdomen un peu trop haut ou un peu trop bas, la mensuration du thorax se trouve de ce fait erronée par défaut ou par excès, et au contraire celle de l'abdomen est trop élevée ou respectivement trop faible. Dans ce cas, l'erreur de dimension réduit non seulement la concordance, mais a pour effet d'introduire un facteur de discordance qui peut même surpasser la concordance réelle entre les deux segments et porter à un coefficient de corrélation négatif.

La deuxième expérience concerne le rapport des sexes dans les naissances. Il est bien connu que la dispersion du rapport des sexes dans les intervalles successifs de temps est pratiquement normale, c'est-à-dire que ses oscillations peuvent être regardées comme déterminées exclusivement par le hasard. Une méthode plus délicate montre pourtant que les variations se succèdent plus souvent dans le même sens que dans le sens contraire ; j'entends par là qu'une augmentation de ce rapport est suivie d'une autre augmentation et de même une diminution d'une autre diminution plus fréquemment que dans l'hypothèse d'indépendance. C'est dire qu'il y a une certaine continuité dans les variations à travers le temps. Cela peut être vérifié en considérant des intervalles de 20, de 40, de 60, de 80 jours, de plusieurs mois, d'une année, de plu-

---

([22]) *Voir* notre *Corso di Statistica* par les soins de S. GATTI et C. BENEDETTI, (*Nuova edizione aggiornata*). Année académique 1954-1955, Veschi, Roma, p. 509-515. Pour la deuxième expérience, voir un exposé plus détaillé dans l'article *Sulle probabilità che x termini di una serie erratica* ... dans *Metron*, t. 18, nos 3-4, 1954, p. 12-24.

sieurs années. Mais, si l'on considère des intervalles de 10 jours, le résultat est incertain, et, entre les jours successifs, c'est le contraire qui se réalise : dans ce cas, il n'y a pas de continuité, mais une légère compensation entre les rapports des sexes. L'explication la plus plausible me paraît celle qui attribue ce résultat aux erreurs, volontaires ou involontaires, dans les déclarations des dates de naissance : si, par l'effet de ces erreurs, des garçons sont déclarés un jour en excès, les filles seront nécessairement déclarées en surnombre le jour ou les jours suivants. Là également, les erreurs ont pour effet non seulement d'atténuer, mais même d'intervertir la relation effective.

*
* *

Rappelons, pour conclure, notre point de départ.

Nous sommes partis de l'affirmation courante que ce qui caractérise la statistique, c'est l'élimination des erreurs accidentelles par l'augmentation du nombre d'observations.

Nous avons montré d'abord que l'élimination des erreurs accidentelles, non seulement n'est pas la tâche exclusive de la statistique, mais qu'en tous cas, elle n'est qu'un moyen, et non pas le but, de la recherche statistique.

Il n'est pas vrai d'ailleurs que l'augmentation du nombre des observations tend toujours à éliminer les erreurs accidentelles. Cela est vrai, toujours ou sous certaines conditions, pour ce qui concerne leurs effets sur les moyennes, mais n'est pas vrai pour ce qui concerne leurs effets sur les autres constantes statistiques. La question est en réalité plus compliquée et la distinction entre erreurs de fréquence et erreurs de dimensions est, à ce sujet, fondamentale.

CHAPITRE III.

MÉTHODE EXPÉRIMENTALE ET MÉTHODE STATISTIQUE.

On a l'habitude d'opposer la méthode expérimentale à la méthode statistique. La première s'appliquerait surtout dans les sciences physiques : la deuxième surtout dans les sciences sociales. Ces derniers temps pourtant on a établi que les phénomènes dont s'occupe la physique peuvent être considérés comme des phénomènes statistiques.

D'autre part, les statisticiens s'occupent de plus en plus des plans des expériences. Cela suggère qu'entre méthode expérimentale et méthode statistique il n'y a pas l'opposition que l'on pensait et conseille de soumettre la question à un nouvel examen.

Il est bon de rappeler d'abord les concepts de méthode expérimentale et de méthode statistique.

Dans la méthode expérimentale on établit l'influence d'une circonstance sur un certain phénomène en comparant deux cas du phénomène dans l'un desquels la circonstance se présente et dans l'autre non, toutes les autres circonstances ne variant pas.

Dans la méthode statistique, au contraire, on prend en considération tous les cas du phénomène dans lesquels se présente la circonstance en question et tous les cas dans lesquels elle ne se présente pas, les premiers cas différant entre eux — et de même les deuxièmes — par une multitude d'autres circonstances que l'on juge accidentelles vis-à-vis de la circonstance en question. On déduit l'influence de la circonstance en question de la différence entre la moyenne du phénomène dans les cas où la circonstance se présente et la moyenne dans les cas où elle ne se présente pas.

Le but des deux méthodes, expérimentale et statistique, est le même, à savoir mettre en lumière l'influence que la circonstance en question exerce sur le phénomène en faisant abstraction de l'influence de toutes autres circonstances.

La base logique des deux méthodes est pourtant tout à fait différente.

La méthode expérimentale n'est en effet qu'un cas particulier de la méthode d'induction qui s'appelle *méthode des différences* et qui s'exprime de la façon suivante : Si un certain phénomène se produit dans un cas et ne se produit pas dans un autre cas qui ne diffère du premier que par l'absence d'une certaine circonstance, cette circonstance différentielle est la cause ou l'effet ou une partie intégrante de l'effet du phénomène.

La méthode statistique se base au contraire sur le principe qu'en augmentant le nombre des observations, l'influence des circonstances accidentelles tend à s'évanouir, de façon que, dans une grande masse d'observations, elle peut être négligée.

Deux exemples feront bien ressortir les différences entre les deux méthodes.

Si un organisme soumis à un certain régime alimentaire présente certains caractères et, en maintenant toutes autres circonstances égales, nous changeons son régime alimentaire, la variation qui se produira dans ses caractères sera considérée comme l'effet du changement dans le régime alimentaire. C'est là une application de la méthode expérimentale.

D'autre part, si nous considérons les caractères d'un grand nombre d'organismes, soumis, dans les circonstances les plus diverses, à un certain régime alimentaire et ceux d'un grand nombre d'autres organismes soumis à un régime alimentaire différent, également dans les circonstances les plus diverses, la méthode statistique nous permet d'établir un lien entre la différence entre les deux régimes alimentaires et la différence entre les moyennes des caractères des deux groupes d'organismes.

De manière analogue, si deux pièces de terre, qui se trouvent dans les mêmes conditions à tous autres points de vue, reçoivent un fumage différent, la différence qui sera observée dans leurs récoltes pourra être attribuée à la différence des engrais. C'est encore une application de la méthode expérimentale.

Nous aurions recours au contraire à la méthode statistique si nous considérions, d'un côté, les terrains, très variés à d'autres points de vue, dans lesquels un certain engrais a été employé et faisions la moyenne des récoltes qu'ils ont produites, et, d'autre part, les terrains, présentant aussi des conditions bien différentes à bien des égards, qui ont reçu un certain fumage, en faisant aussi la moyenne de leurs récoltes. La différence entre les récoltes moyennes des deux groupes de terrains pourra être attribuée à la différence des fumages.

Si nous examinons de près l'essence de la méthode expérimentale et de la méthode statistique, nous nous rendons compte aisément que les deux méthodes ne s'excluent pas mutuellement.

En réalité il n'y a pas d'opposition entre expérience et statistique, mais entre expérience et simple observation. L'expérience se distingue de la simple observation en ce que, d'un côté, les circonstances dans lesquelles elle se déroule sont disposées à l'avance, et que, de l'autre, les variations de la circonstance dont on veut examiner l'influence sont artificielles, c'est-à-dire contrôlées.

Ces deux conditions sont certainement favorables à l'application de la méthode des différences, dont la méthode expérimentale constitue un cas particulier, mais elles ne sont pas du tout indispensables.

La pré-disposition des circonstances dans lesquelles l'expérience se déroule y est favorable car elle rend plus facile la réalisation de la condition que toutes autres circonstances, en dehors de celle dont on veut examiner l'influence soient égales. Pourtant des conditions se produisent parfois dans la nature qui sont bien analogues à celles d'une expérience. Il y a, pour ainsi dire, des *expériences naturelles*. C'est le cas dans les jumeaux univitellins, dans lesquelles il y a identité du patrimoine héréditaire et différence du milieu, de façon que les différences qui se manifestent entre ces jumeaux puissent être attribuées à l'influence du milieu.

Le fait que dans l'expérience la variation de la circonstance dont on examine l'influence est artificielle permet de juger avec certitude de quel côté est la cause et de quel côté est l'effet. Cela n'est pas toujours possible dans l'observation simple, mais il peut arriver que cela le soit.

Ce fait permet aussi de doser l'intensité de la circonstance dont on examine l'influence.

D'autre part, la multiplicité des observations, qui constitue la condition indispensable pour l'application de la méthode statistique, n'est pas exclue dans l'expérimentation. Il est vrai qu'une seule expérience peut suffire dans bien des cas ; mais il peut être aussi utile de multiplier les expériences, dont les résultats peuvent alors être soumis au procédé statistique.

Une différence que nous avons signalée entre l'expérience et l'observation simple se retrouve entre la méthode expérimentale et la méthode statistique.

Les deux méthodes — expérimentale et statistique — visent toujours à établir un lien entre la présence d'une certaine circonstance et la présence d'une autre. Mais la méthode expérimentale nous précise aussi quelle est celle des deux circonstances qui est la cause et quelle est celle qui est l'effet, tandis que, dans bien des cas, la méthode statistique ne peut qu'établir un lien fonctionnel et, même quand elle permet de conclure que c'est un lien de causalité, il n'est pas dit qu'elle puisse préciser quelle est la cause et quel est l'effet. La différence vient du fait que la méthode expérimentale est dynamique et la méthode statistique est statique. C'est seulement lorsque les deux circonstances liées d'un lien fonctionnel ne sont pas contemporaines, que la méthode statistique elle aussi permet généralement de dire quelle est la cause ; c'est la circonstance antécédente. Par exemple, dans le cas du fumage et de la récolte, la période pendant laquelle agit l'engrais et la période de la récolte sont nettement différenciées dans le temps, de façon que la méthode statistique permet d'affirmer que la différence du régime d'engrais est la cause de la différente quantité de récolte. Au contraire, dans le cas du régime alimentaire, la période pendant laquelle on maintient le régime se superpose à la période pendant laquelle les caractères de l'organisme sont observés de façon que, si nous constatons que tous les organismes qui suivent un certain régime alimentaire présentent certains caractères, en réalité nous ne pouvons pas décider si c'est le régime alimentaire qui détermine la différence dans les dits caractères ou si, au contraire, c'est la différence dans les dits caractères qui provoque la différence dans le régime alimentaire, ou enfin si entre l'un et les autres il y a un lien d'interdépendance. On constate que les alcooliques sont des dégénérés, mais est-ce l'alcoolisme qui détermine la dégénérescence ou bien la dégénérescence qui provoque l'alcoolisme, ou bien les deux causes agissent-elles en même temps? C'est là une différence qui touche à l'efficacité des deux méthodes; supérieure pour la méthode expérimentale.

Une autre différence touche au domaine de leur application.

La méthode statistique comporte une compensation entre les effets qui peuvent se réaliser seulement lorsque ceux-ci ont un caractère quantitatif, tandis que la méthode expérimentale s'applique aussi bien aux cas où les effets sont qualitatifs. Ici encore, la méthode expérimentale a l'avantage.

Une autre différence fondamentale, mais non absolue, s'accuse entre les méthodes expérimentale et statistique; c'est que la méthode expérimentale vise surtout à éliminer les circonstances perturbatrices, tandis que la méthode statistique vise surtout à compenser les effets des dites circonstances. Il y a des cas dans lesquels seulement une des deux éliminations est possible. Cette différence contribue donc à délimiter le domaine d'application propre des deux méthodes.

Au même résultat contribuent les difficultés pratiques d'application qui sont différentes pour les deux méthodes.

Il n'est pas du tout facile, dans les expériences, de rendre égales toutes les circonstances sauf celle dont on veut mesurer l'influence.

Dans bien des cas, pourtant, on peut avoir recours, pour éviter des surprises, à des sujets de contrôle. Par exemple, on soumet un certain animal à un nouveau régime alimentaire, à la suite duquel on constate dans son comportement une certaine variation. Mais celle-ci ne pourrait-elle pas dépendre, par exemple, d'une variation du climat que d'ailleurs il aurait été impossible de rendre constant ? Pour écarter ce doute on peut avoir recours à un animal de contrôle semblable à celui qui a fait l'objet de l'expérience et qui serait placé dans les mêmes conditions que le premier, mais pour lequel le régime alimentaire ne serait pas changé.

Lorsque, comme dans le cas qui précède, la variation de la circonstance se produit sur le même organisme, on est exposé au risque de confondre ce qui n'est que l'effet temporaire avec ce qui est l'effet permanent de la modification. Il peut bien arriver que pour une certaine période, pendant laquelle l'organisme n'est pas encore accoutumé au nouveau régime, un effet se produise qui est tout à fait différent de la condition qui s'établira lorsque l'organisme s'y sera adapté. De cette façon il serait erroné, par exemple, de conclure à un effet désagréable du tabac en se basant sur les sensations qui se produisent lorsqu'on commence à fumer, tandis qu'une comparaison faite suivant la méthode statistique entre les conditions des fumeurs et celles des non fumeurs pourrait porter à un résultat opposé.

D'autre part, la méthode statistique présente elle aussi ses difficultés pratiques d'application.

D'un côté il n'est pas facile, dans certains cas, d'atteindre un nombre d'observations suffisant pour que les effets des circonstances perturbatrices se compensent.

De l'autre, le danger existe que certaines circonstances, dont les effets devraient se compenser en tant qu'accidentels, ne soient pas en réalité indépendantes de la circonstance dont on veut mesurer la variation. Dans ce cas, la différence entre les moyennes observées, qui devrait mesurer l'influence de la circonstance à éprouver, vient à dépendre aussi de l'influence de ces autres circonstances et parfois seulement de celles-ci. Par exemple, si nous considérons toutes les femmes qui ont 20 ans et celles qui en ont 40 et comparons le poids moyen des nouveaux-nés des premières et des deuxièmes, nous trouvons une différence en la faveur des enfants de la deuxième catégorie qui se révèlent d'un poids supérieur. Pourtant une analyse approfondie démontre que la différence ne peut pas être attribuée à la différence d'âge, mais bien à une autre circonstance, le nombre des accouchements précédents, qui est en moyenne supérieur pour les femmes de 40 ans que pour celles de 20, tandis qu'à parité du nombre d'accouchements l'âge de la mère n'exerce aucune influence systématique.

De tout ce que nous venons de dire il ressort qu'entre la méthode expérimentale et la méthode statistique il n'y a pas en réalité d'opposition, mais plutôt qu'elles sont complémentaires, de façon que l'on peut bien considérer l'opportunité de les employer, lorsque cela est possible, toutes les deux.

Ce ne sont pas d'ailleurs les seules méthodes auxquelles on peut avoir recours pour éliminer l'influence des circonstances perturbatrices.

Cela ressortira du plan, que nous allons esquisser, pour épurer les résultats des observations des perturbations auxquelles celles-ci sont exposées.

Nous avons déjà distingué les procédés par lesquels on élimine les circonstances perturbatices, des procédés par lesquels on élimine leurs effets sur les résultats des observations.

On peut encore les distinguer selon que cette élimination se produit avant ou après l'observation.

Notre plan comportera donc quatre paragraphes, suivant que les phénomènes visent à :

1° L'élimination préalable des circonstances perturbatrices;

2° La compensation préalable de leurs effets;

3° L'élimination postérieure des circonstances perturbatrices;

4° La compensation postérieure de leurs effets.

1. Élimination préalable des circonstances perturbatrices. — Nous sommes dans le domaine traditionnel de la méthode expérimentale. L'idéal est de garder invariables toutes les circonstances sauf celle dont on veut étudier l'influence, mais cela n'est pas toujours possible.

Si, parmi ces circonstances, on faisait entrer en ligne de compte le lieu et le temps, cela serait toujours évidemment impossible, car les deux situations que l'on veut considérer doivent différer au moins à l'une de ces deux circonstances : le lieu ou le temps.

Il convient pourtant d'observer qu'il y a, à la base de toute recherche scientifique, l'hypothèse, d'ailleurs fondée par l'expérience, que l'espace et le temps n'ont *per se* aucune influence sur le développement des phénomènes. Sans cela on ne pourrait jamais parler, au sens scientifique, d'une *loi*, c'est-à-dire d'un rapport constant (dans l'espace ou dans le temps) entre deux phénomènes.

Mais espace et temps sont parfois indissolublement liés à des circonstances qui exercent bien une influence sur la circonstance dont on voudrait isoler les effets.

Qui voudrait mesurer le progrès de l'agriculture, en comparant le produit du blé obtenu par hectare en Mésopotamie à l'époque préhistorique avec le produit actuel, arriverait à un résultat bien décevant, car le lieu est bien le même, mais le climat est — selon toute vraisemblance — radicalement changé en pire.

En pratique, le laps de temps pendant lequel les circonstances doivent se maintenir invariables n'est généralement pas si long, mais les changements liés aux saisons sont assez rapides pour ne pas souffrir d'être négligés dans beaucoup d'expériences. Il ne suffirait même pas de considérer les saisons correspondantes de deux années successives, car les conditions de climat pendant la même saison peuvent être bien différentes d'une année à l'autre.

Ne pouvant pas réaliser dans l'expérience l'identité de lieu, et simultanément, de temps, il est naturel que l'on renonce à l'identité de l'un ou de l'autre selon que l'un ou l'autre paraît avoir plus d'influence ou une influence plus irrégulière. Par conséquent, pour l'agriculture, qui est très sensible aux variations saisonnières, on renonce à l'identité de lieu pour garder celle de temps. Au contraire lorsque les objets sur lesquels l'expérience se fait sont rares ou coûteux, on renonce à l'identité de temps en comparant le même objet en des moments divers.

Lorsque l'objet sur lequel l'expérience est faite n'est pas seulement rare, mais aussi difficile à déplacer, il se peut que l'on doive renoncer à l'identité de lieu et aussi à l'identité de temps.

Lorsqu'on opère sur le même objet il faut aussi porter attention sur ce que l'on pourrait nommer l'« auto-influence ». Si l'on soumet successivement un organisme à deux traitements, il se peut que le premier influence l'organisme de façon que le second traitement ne puisse s'effectuer dans les mêmes conditions que le premier. L'identité de lieu est dans ce cas seulement apparente.

Si l'on renonce à l'identité de lieu pour garder celle de temps, il faut maintenir cette identité pour toute la durée de l'expérience, qui est nécessaire et parfois longue. Pour deux lopins de terre, par exemple, les conditions atmosphériques peuvent être tout à fait semblables en été et au contraire bien inégales en hiver; il est donc nécessaire que l'expérience dure au moins une année.

Souvent un agent dont on veut éprouver l'influence ne produit son plein effet qu'après un certain temps; dans ce cas, la durée insuffisante de l'observation devrait être considérée comme une circonstance perturbatrice à éliminer.

Une identité qui a une grande importance dans certaines recherches, est celle de l'expérimentateur ou observateur, non seulement dans le sens qu'il s'agisse de la même personne, mais aussi dans le sens que ses conditions physiques et psychiques soient à peu près les mêmes. Et pourtant il y a des cas où les diverses phases de l'expérience ont lieu en des endroits très distincts l'un de l'autre ou à des dates assez éloignées et où il faut avoir recours à des personnes différentes.

Lorsqu'on parle d'éliminer préalablement les circonstances perturbatrices, on entend se référer aux circonstances perturbatrices systématiques, car il est difficile de prévoir l'action des circonstances perturbatrices accidentelles.

On vise toutefois à les éliminer lorsqu'on concentre l'expérience ou l'observation sur des sujets ou objets typiques pour ce qui concerne la circonstance que l'on a en vue d'influencer.

Si je veux éprouver l'influence d'un engrais sur la récolte et choisis pour l'expérience deux lopins de terre sur lesquels la récolte est normale, je peux penser avoir éliminé préalablement les influences perturbatrices accidentelles. En effet, si l'on admet (hypothèse a) que chaque circonstance perturbatrice soit d'autant plus fréquente qu'elle est plus faible, on peut regarder la valeur typique comme celle qui se serait manifestée en l'absence de toutes circonstances perturbatrices.

Pour faire confiance au résultat de l'expérience, il faut pourtant admettre aussi (hypothèse b) que l'hypothèse a, valable dans les conditions habituelles, garde sa validité lorsque le nouvel engrais est employé.

Au lieu de se baser sur les hypothèses susdites, on pourrait penser à se baser sur le principe que les circonstances perturbatrices à caractère accidentel se compensent. Le principe ne peut pas être révoqué en doute. On pourrait même dire qu'il n'exprime qu'une tautologie, car les circonstances perturbatrices à caractère accidentel par définition se compensent; mais la compensation des circonstances perturbatrices n'apporte pas nécessairement avec elle la compensation de leurs effets. Nous reviendrons sur ce point dans le paragraphe qui suit.

2. Compensation préalable des effets des circonstances perturbatrices. — Les procédés sont différents suivants que les circonstances perturbatrices présentent un caractère systématique ou un caractère accidentel.

Dès le premier cas, il s'agit de trouver un procédé qui fasse tomber équitablement l'influence systématique des dites circonstances sur les deux situations que l'on veut comparer.

Nous voulons décider si un certain excitant, par exemple le café, facilite ou au contraire entrave le succès aux examens scolaires. Avant l'examen, nous donnons à une partie des étudiants du café, aux autres de l'eau pure. Mais n'y-a-t-il pas une circonstance perturbatrice à éviter ayant caractère systématique, représentée par l'ordre de succession des examens?

Les étudiants examinés en second lieu ont l'avantage d'avoir écouté les questions et les réponses des examens précédents. On prétend aussi que les examinateurs sont moins sévères en poursuivant leur tâche.

On pourrait alterner les étudiants des deux groupes : le premier, le troisième, etc., comprenant des étudiants qui ont pris du café; le deuxième, le quatrième, etc., des étudiants qui n'ont bu que de l'eau. Il y aurait de la sorte un petit avantage pour les étudiants des nombres pairs; mais l'avantage serait annulé si le dernier étudiant examiné était de la même catégorie que le premier.

La compensation ainsi réalisée serait exacte si la tendance dont on veut éliminer les effets était absolument linéaire, c'est-à-dire si son intensité était proportionnelle au nombre des examens passés; mais en pratique, dans la plupart des cas, on peut la considérer comme satisfaisante même si cette condition n'est pas réalisée.

En nous basant sur la même hypothèse linéaire, nous pourrions neutraliser l'influence du terrain sur la récolte dans une expérience visant à déterminer l'effet comparatif de deux engrais. On alternera les pièces traitées avec un engrais et celles traitées avec l'autre dans la direction dans laquelle on a raison de croire que le terrain varie progressivement.

Si l'on a des raisons de croire que le terrain varie systématiquement en deux directions normales l'une à l'autre, par exemple du Nord au Sud et de l'Est à l'Ouest, on disposera les pièces de terrain traitées avec les deux engrais comme les carrés blancs et noirs d'un échiquier, la compensation étant ainsi réalisée dans les deux sens.

Lorsque les circonstances perturbatrices ont un caractère accidentel, on vise à la compensation de leurs effets moyennant l'augmentation du nombre des observations.

Nous sommes alors dans le domaine traditionnel de la méthode statistique.

On peut fixer d'abord un nombre d'observations suffisant pour obtenir, avec une certaine probabilité, l'approximation des résultats que la nature et le but de la recherche exigent. C'est le cas lorsqu'on connaît préalablement la variabilité du phénomène. Dans le cas contraire, il est préférable d'établir au préalable l'approximation désirée dans les résultats que l'on obtiendra dans le cours de l'observation et d'arrêter celle-ci lorsque l'approximation sera jugée satisfaisante. Dans ce cas, ce que l'on établit préalablement n'est pas le nombre des observations, mais les critériums selon lesquels on procédera pour apprécier l'approximation obtenue. Il est important à ce sujet de distinguer deux sortes d'erreurs déterminées par les circonstances perturbatrices accidentelles : *erreurs de dimension*, telles que les inexactitudes dues à une technique imparfaite de mensuration et, en général, les écarts de la moyenne dus à l'influence des circonstances indépendantes de celle que l'on veut éprouver; et *erreurs de fréquence*, dues au fait que l'on observe seulement une partie, choisie au hasard, des cas du phénomène.

En augmentant le nombre des observations — dit-on — l'influence des erreurs des deux catégories se réduit de plus en plus et finit par disparaître. On entend par là l'influence sur l'intensité moyenne des phénomènes. En ce sens la proposition est vraie pour ce qui concerne les erreurs de fréquence ; on l'appelle en général *loi des grands nombres*. Pour les erreurs de dimensions, au contraire, elle ne peut à la rigueur être acceptée qu'à certaines conditions (telles que la proportionnalité entre l'intensité de la cause de perturbation et de ses effets, ou la symétrie des erreurs positives et négatives ou d'autres conditions plus ou moins restrictives) qui peuvent être présentes ou ne l'être pas. Il est à ajouter que, même lorsque les effets de chaque circonstance perturbatrice se compensent, il se peut que les effets de la combinaison de plusieurs circonstances perturbatrices ne se compensent pas. La compensation s'avère si les effets des circonstances diverses jouissent de la propriété additive; mais ne se réalise pas nécessairement dans le cas contraire, par exemple si les effets défavorables prédominent sur les effets favorables.

Les circonstances perturbatrices accidentelles exercent une influence non seulement sur l'intensité moyenne des phénomènes, mais aussi sur leur variabilité et sur leurs relations mutuelles. L'augmentation du nombre des observations réduit cette influence lorqu'il s'agit d'erreurs de fréquence. Lorsqu'au contraire, il s'agit d'erreurs de dimensions, les effets sont plus complexes; j'en ai parlé au chapitre II auquel je renvoie ceux qui désirent se renseigner à ce sujet. Il suffit de conclure ici que l'application de la méthode statistique doit être faite avec beaucoup de précaution.

Examinons maintenant l'élimination des circonstances perturbatrices après les observations.

3. Élimination postérieure des circonstances perturbatrices. — On l'obtient en classant les observations selon la modalité de la circonstance ou selon les modalités combinées des circonstances dont on veut éliminer l'influence et en comparant l'influence de la circonstance à examiner, sur les classes d'observations qui présentent la même modalité ou combinaisons de modalités que la circonstance ou les circonstances à éliminer.

Par exemple, si l'on veut étudier l'influence de la profession sur la fécondité du mariage on peut la déduire du nombre d'enfants par mariage, mais il faut évidemment éliminer, dans beaucoup de pays, l'influence des pratiques anticonceptionnelles et, dans tous les pays, l'influence de l'âge au mariage de la femme, mieux encore les âges des deux époux. Ayant classé les couples selon qu'ils ont recours ou non aux pratiques anticonceptionnelles et, dans chacune des deux catégories, d'après l'âge au moment du mariage de la mère ou d'après les combinaisons des âges des deux parents, nous pouvons comparer le nombre d'enfants par mariage pour les couples de différentes professions qui n'ont pas recours aux pratiques anticonceptionnelles et dans lesquels la femme avait de 20 à 25 ans et le mari de 25 à 30 ans au moment où ils se sont mariés. Il n'est pas sûr du tout que le résultat serait le même si l'on considérait les couples qui comme les précédents n'ont pas eu recours aux pratiques anticonceptionnelles, mais qui se sont mariés lorsque la femme avait plus de 25 ans et le mari plus de 30 ans.

Il ne suffit donc pas, dans la plupart des cas, de dire que l'on a rendu égales certaines circonstances perturbatrices; il faut préciser le niveau auquel on les a rendues égales.

L'élimination postérieure des circonstances perturbatrices, dont nous venons de parler, s'applique surtout aux résultats obtenus dans les relevés statistiques, tandis que dans les expériences on a recours surtout à l'élimination préalable. Mais, dans ce deuxième cas aussi, pour donner à l'expérience une signification précise, il faut établir à quel niveau les circonstances éliminées sont rendues égales. A parité de température, par exemple, l'irrigation modifie la récolte d'une certaine culture d'une façon très différente selon que le niveau de la température est bas ou est élevé.

La différence entre l'expérience et l'observation statistique est que dans l'expérience on peut généralement rendre égales beaucoup de circonstances, de façon qu'une seule expérience, ou en tous cas un nombre restreint d'expériences, peut être concluant, tandis que les données statistiques sont d'habitude combinées selon un nombre très limité de circonstances (trois ou quatre au plus), une multitude d'autres circonstances perturbatrices n'étant pas éliminées. Pour celles-ci il faut faire appel à la compensation des effets et par conséquent exiger que les combinaisons considérées comprennent un grand nombre d'observations. C'est là évidemment un désavantage de l'utilisation des données statistiques par comparaison aux résultats de l'expérience.

D'un autre côté l'utilisation des données statistiques présente l'avantage de fournir des données pour des gradations ou combinaisons multiples des circonstances égalisées. Elle équivaut, de la sorte, à une multiplicité d'expériences correspondant aux divers niveaux des circonstances égalisées.

Elle ouvre en outre la possibilité de tenir compte à la fois de tous les résultats, en opérant une synthèse. Dans cette synthèse il faut qu'à chacune des gradations (ou combinaisons des circonstances) égalisées on attribue un poids qui soit constant dans les groupes d'observations à comparer. Pour décider, dans l'exemple que nous avons pris, de l'influence de la profession sur la fécondité des mariages, on tiendra compte, dans la synthèse, non seulement de la combinaison des âges : 20-25 pour la femme et 25-30 pour l'homme, mais de toutes les autres : moins de 20 ans pour la femme et 25-30 pour l'homme, 25-30 ans pour la femme et pour l'homme et ainsi de suite, en attribuant à chaque combinaison un certain poids qui doit être le même pour toutes les professions dont on veut comparer l'influence. On aboutit de la sorte à la *méthode de la population type* et aux autres méthodes analogues appelées en statistique *méthodes d'élimination*, dont le calcul des nombres indices pondérés des prix constitue une application particulièrement importante. Les conditions auxquelles ces méthodes — et en particulier les nombres indices — doivent répondre, et le choix des poids appropriés soulèvent des questions qui sont parmi les plus délicates de la méthodologie statistique.

Cette méthode synthétique a un avantage important sur la méthode analytique que nous avons exposée auparavant. C'est que les combinaisons qui présentent un nombre limité d'observations et n'autorisent aucune conclusion basée sur chacune d'elles, peuvent être utilisées dans la synthèse, car les erreurs de fréquence qui se produisent dans les diverses combinaisons, sont indépendantes les unes des autres et se compensent.

Les méthodes dont nous avons parlé dans ce paragraphe visent à éliminer les circonstances perturbatrices à caractère systématique.

Pour ce qui concerne l'élimination postérieure des circonstances perturbatrices à caractère accidentel, nous rappelons ce que nous avons dit au sujet de leur élimination préalable. La différence est qu'ici on ne choisit pas, pour l'expérience ou l'observation que l'on va faire, les sujets ou objets ayant présenté des valeurs typiques ou moyennes dans des observations ou expériences antécédentes, mais les valeurs typiques ou moyennes qui résultent de la collection des données ou de l'expérience que l'on vient de faire. Les hypothèses, auxquelles leur utilisation est subordonnée, restent les mêmes.

4. Compensation postérieure des effets des circonstances perturba-

trices. — Lorsqu'on n'a pas éliminé les circonstances perturbatrices ni avant ni après la recherche, et que l'on n'a pris aucune mesure préventive pour compenser les erreurs qui en sont les effets, il ne reste qu'à tenter cette compensation sur les données recueillies.

Dans ce but, on a très souvent recours à des coefficients de correction établis sur la base d'autres recherches.

Se basant sur le rapport moyen entre la taille de l'homme et de la femme dans les populations connues, on a l'habitude de multiplier par 1,08 suivant la proposition de Galton, la taille de la femme adulte pour la rendre comparable à celle de l'homme en vue de mesurer la ressemblance entre enfants et parents ou entre époux.

Se basant sur la diminution observée de la température avec l'altitude au-dessus du niveau de la mer, on réduit la température des localités différentes à celle que l'on aurait observée, présume-t-on, si les localités en question avaient été toutes au niveau de la mer. De même, on peut rendre comparables les données sur la température recueillies dans des pays différents à des heures différentes du jour, en les rapportant à la même heure.

Se basant sur les déterminations du métabolisme basal aux différents âges et pour les deux sexes, on réduit le nombre des habitants à leur équivalent en hommes adultes pour mesurer leurs besoins alimentaires.

Ayant établi l'équation personnelle d'un observateur on peut corriger ses mensurations en les rendant comparables à celles obtenues par d'autres observateurs.

Moyennant un coefficient approprié établi par des sondages préalables, on peut rectifier le nombre des morts-nés dans un pays où l'on enregistre comme tels toutes les naissances sans vie déclarées au bureau de l'état civil.

De même, par des coefficients de correction, on vise à éliminer la portée de l'évasion dans les déclarations des revenus, de la richesse, des loyers; à éliminer dans quelques pays les lacunes dans les déclarations des naissances et des décès; dans d'autres, l'effet de la contrebande sur les statistiques du commerce international. On pourrait multiplier les exemples.

Dans tous les cas que nous venons de mentionner, les coefficients de correction sont établis sur des données différentes de celles de la recherche. On peut les appeler *coefficients extérieurs de correction.*

Dans d'autre cas, ils sont établis sur les données mêmes de la recherche : nous les appellerons *coefficients intérieurs de correction.*

Par exemple, le coefficient de correction employé pour rendre comparable la taille de la femme à celle de l'homme, au lieu d'être établi pour les populations d'autres pays, pourrait être établi sur le groupe même d'hommes et de femmes dont on veut mesurer la ressemblance. Cela pourtant est licite seulement si ce groupe est assez nombreux, autrement on ne pourrait pas supposer que la taille moyenne des hommes du groupe égale celle des femmes, sans influencer la mesure même de la ressemblance.

Dans l'exemple que je viens de donner, la détermination du coefficient extérieur de correction est exceptionnellement simple. Généralement elle est plus complexe. Il s'agit en général d'établir, sur les données disponibles, un système d'équations dont la solution donne les valeurs des paramètres qui expriment d'un côté l'influence de la circonstance que l'on veut mesurer et de l'autre l'influence des autres circonstances qui sont considérées comme perturbatrices vis-à-vis de celle-là. On a recours généralement aux méthodes de la régression ou de la corrélation partielles. Par exemple, pour déterminer l'influence d'un certain fumage sur la récolte, on établit un système d'équations qui permet de mesurer l'influence du fumage à parité de lopins de terrain et de température, ces dernières étant considérées comme circonstances perturbatrices à caractère systématique. Pour déterminer l'influence sur la fécondité, de l'âge de la mère au moment du mariage on établit un système d'équations qui permet de mesurer l'influence de l'âge de la mère à parité de l'âge du père et de la durée du mariage.

La façon dont l'observation est organisée est naturellement importante pour pouvoir établir le système d'équations le plus approprié au but de la recherche. Le plan fixant la manière dont on doit recueillir les données dans les recherches statistiques (auquel on a toujours donné beaucoup de soins), et le plan des expériences (dont l'étude a pris un si grand développement ces derniers temps) donnent précisément les règles à suivre.

Les méthodes de régression ou de corrélation partielles, que l'on emploie généralement, sont astreintes pourtant à des hypothèses dont la portée pratique a été jusqu'à présent insuffisamment étudiée.

La première hypothèse concerne l'étendue des données sur lesquelles on établit le système d'équations : celle-ci doit être suffisante pour éliminer l'influence du hasard.

La deuxième admet que les influences des différentes circonstances jouissent de la propriété additive, ce qui n'est jamais vrai qu'approximativement et en général ne l'est même pas approximativement dans les recherches biologiques et sociales.

La troisième suppose que les régressions ou corrélations sont linéaires, et cela aussi ne peut être admis qu'approximativement, tout au plus.

Naturellement, la solution du système d'équations établi conduit toujours à des paramètres, mais dans quelle mesure les coefficients intérieurs, auxquels on parvient, représentent-ils les influences des facteurs à mesurer?

Pour résoudre la question, on peut comparer les coefficients intérieurs ainsi obtenus aux coefficients extérieurs déterminés sur la base de données qui ne font pas partie de la recherche.

S'est-on jamais donné cette peine? Je me la suis donnée pour déterminer les coefficients de correction qui expriment les besoins alimentaires de l'homme et de la femme aux divers âges par comparaison aux besoins de l'homme adulte. Les résultats concernant quatre séries relatives à trois pays (classes moyennes suédoises, ouvriers et employés berlinois, ouvriers lettons), n'ont pas été du tout encourageants.

Je ne veux pas les généraliser; il se peut que ce cas soit spécialement compliqué; mais il est certainement permis de conclure que l'application de la méthode de la régression et corrélation partielle couramment employée devrait bien être soumise à un contrôle rigoureux.

Les coefficients de correction dont nous avons parlé visent à éliminer des données recueillies, les perturbations systématiques. Que dire des perturbations accidentelles?

Les statistiques nous fournissent très souvent des données qui ne concernent pas la totalité des cas du phénomène ou qui sont affectées d'erreurs de dimension. Il s'agit en général de données recueillies dans des buts administratifs ou aussi, quelquefois, de données recueillies à l'occasion de travaux scientifiques, mais sans que l'on ait pris soin d'étudier préalablement la portée des erreurs accidentelles à prévoir et d'établir en conséquence le plan de collecte des données ou le plan de l'expérience. Rien n'empêche que ce qui n'a pas été fait préalablement soit fait plus tard. Naturellement les conditions ne sont pas les mêmes. D'un côté ce qui est fait est fait; il faut le prendre comme il est, même s'il n'est pas fait de la manière qui eût été la meilleure pour faciliter l'élimination des erreurs accidentelles. Dans ces conditions, même le calcul de l'erreur probable offre des difficultés qui dans une étude préalable ne se présentent pas : son calcul pourtant n'est pas toujours impossible. Lorsque les statistiques font connaître les valeurs individuelles des cas observés où celles-ci sont distribuées dans des classes très restreintes on peut étudier le groupement des données recueillies, ou de leurs classes, qui paraît le plus approprié pour éliminer les perturbations accidentelles, en négligeant les distinctions que l'on a raison de croire sans importance pour le but de la recherche.

Dans la plupart des cas pourtant la classification des valeurs n'est pas à ce point détaillée qu'elle suggère la constitution des groupements ultérieurs et on l'accepte telle qu'elle est. On prend alors la moyenne des valeurs dans chaque classe en admettant — généralement sans examen critique préalable — que les différences entre lesdites valeurs dépendent de circonstances perturbatrices à caractère accidentel et que partant les écarts de la moyenne correspondent à des erreurs, qui se compensent, affectant la valeur qui se serait vérifiée si lesdites circonstances perturbatrices n'étaient pas intervenues. C'est là l'application la plus courante et la plus simple de la méthode statistique.

*
* *

L'exposé qui précède sur la procédure à suivre pour épurer les observations des circonstances perturbatrices à caractère systématique et accidentel, montre assez clairement que les méthodes que l'on emploie sont toutes plus ou moins incertaines. D'où la nécessité d'avoir recours à tous les moyens possibles de contrôle et d'intégration.

5. **Contrôle des résultats.** — Parlons d'abord des contrôles. Dans les expériences il est difficile d'exclure qu'une circonstance perturbatrice se faufile et agisse en même temps que la circonstance dont on veut mesurer les effets, de façon que l'on attribue à celle-ci des effets qui en réalité sont produits totalement ou en partie par la circonstance perturbatrice inaperçue.

Pour éliminer cette éventualité, on peut utiliser des sujets ou objets de contrôle, c'est-à-dire des sujets ou objets qui sont soumis à toutes les circonstances des sujets ou objets sur lesquels porte l'expérience sauf la circonstance dont on veut mesurer les effets. Si ceux-ci ne se produisent pas dans les sujets ou objets de contrôle, l'expérience est doublement probante. On constate, en effet, non seulement que, à égalité de toutes autres circonstances, l'introduction de la nouvelle circonstance à éprouver est suivie des effets en question, mais aussi que ces effets ne peuvent pas être déterminés par une autre circonstance qui s'est introduite à notre insu, car les effets en question ne se réalisent pas en l'absence de la circonstance à éprouver.

Des contrôles sont souvent opportuns et parfois nécessaires pour ce qui concerne les conclusions obtenues par la méthode statistique. Il s'agit généralement d'applications du Calcul des Probabilités. Or il ne faut jamais oublier que le Calcul des Probabilités ne permet que des conclusions plus ou moins probables, jamais des conclusions certaines. Ajoutez que bien des fois ces applications impliquent des hypothèses qui peuvent ne pas correspondre à la réalité. Les contrôles en ce cas peuvent consister dans la répétition de l'observation sur des cas différents, ou bien dans le remaniement des données, en les groupant ou classifiant différemment, ou encore dans une augmentation du nombre des observations. Plus difficile est généralement l'élimination des hypothèses qu'impliquent les applications.

Si les phénomènes traités par la méthode statistique sont susceptibles d'être soumis à l'expérience, on ne doit jamais négliger de le faire.

* * *

6. **Intégration mutuelle des méthodes expérimentale et statistique.** — Avec cette dernière remarque nous sommes déjà entrés dans le domaine de l'intégration, intégration de la méthode statistique par la méthode expérimentale.

L'intégration peut être faite aussi dans la direction inverse, c'est-à-dire en intégrant les autres méthodes par la méthode statistique. En effet, l'élimination des circonstances perturbatrices faite soit préalablement dans l'expérience, soit successivement moyennant la classification des données observées n'exclut pas que dans les résultats se produisent des divergences qui, si l'élimination avait atteint pleinement son but, devraient dépendre exclusivement de perturbations accidentelles. C'est avec les méthodes de la statistique que l'on peut décider si cette présomption répond probablement à la vérité ou si, au contraire, les divergences révèlent des perturbations systématiques résiduelles.

Si, à la fin de l'expérience ou de l'observation, des perturbations systématiques sont prouvées ou seulement soupçonnées, et de même si elles sont prévues avant que l'expérience ou l'observation ne soit commencée, on peut bien tâcher d'en fixer des limites supérieures de façon à établir un intervalle de sécurité.

On peut avoir recours dans ce but à des méthodes qui sortent du cadre de l'expérience ainsi que de celui de l'observation massive propre de la statistique, bien que les statisticiens et les expérimentateurs s'en servent souvent. Tel est le cas pour le procédé des coefficients extérieurs de correction dont nous avons vu l'importance.

Mais, à mon avis, la discussion sur ce point de savoir si tel ou tel autre procédé entre ou n'entre pas dans les méthodes statistique et expérimentale, est tout à fait inutile.

En effet, il est essentiel de retenir que *la meilleure méthode, je voudrais même dire la seule bonne méthode — dans le domaine que nous venons d'explorer, comme dans tout autre domaine scientifique — est d'employer sans prévention toutes les méthodes possibles dans la mesure où elles sont profitables.*

CHAPITRE IV.

LES PROBLÈMES DE L'INVERSION STATISTIQUE.

I. — LES PROCÉDÉS CLASSIQUES.

Nous pouvons admettre, ainsi que nous l'avons montré précédemment (Chap. II : *De l'augmentation du nombre des observations et de ses effets sur les constantes statistiques*) que, si l'on augmente le nombre des observations, on peut pratiquement éliminer l'influence des erreurs accidentelles de fréquence sur la moyenne des grandeurs observées, et compenser, lorsque se réalisent certaines conditions, l'influence des erreurs accidentelles de dimension. Il y a lieu cependant de se demander : quels sont les rapports entre le nombre des observations et l'intensité de ces erreurs ? C'est là un autre problème fondamental de la méthodologie statistique que l'on doit résoudre quand on veut juger de la validité des données disponibles.

Or, lorsque d'un phénomène on connaît la probabilité qui correspond à sa fréquence dans la masse ou univers des observations, il est facile de tirer, du théorème de Bernoulli, la probabilité d'une erreur accidentelle de fréquence dans une ou dans plusieurs observations et, quand on connaît la valeur vraie et l'erreur quadratique moyenne des observations d'une grandeur absolue, il est facile de tirer, de la courbe des erreurs accidentelles, la probabilité d'une erreur de dimension dans une ou dans plusieurs observations.

Mais ordinairement on ne connaît ni la probabilité dans le premier cas, ni la valeur vraie et l'erreur quadratique moyenne de la grandeur absolue dans le second. Pour cette raison, on substitue habituellement à ces valeurs celles de la fréquence et, respectivement, de la valeur moyenne et de l'erreur quadratique moyenne tirées d'un grand nombre d'observations. Ces substitutions, et en particulier celle de l'erreur quadratique moyenne observée à l'erreur quadratique moyenne théorique, donnent lieu à quelques réserves, mais ce n'est pas de cela que dépend la difficulté la plus grande. Pour passer — même approximativement — de la grandeur observée à la valeur probable de la grandeur vraie, suffit-il d'avoir les mêmes connaissances qu'il faut posséder pour passer de la grandeur vraie à la valeur probable de la grandeur observée? Et, si elles ne sont pas suffisantes, quelles sont les autres connaissances nécessaires ? Voilà le point décisif de la question. Voilà le problème qui est à la base de toutes les discussions sur la probabilité inverse, sur les tests de significativité, sur les intervalles de confiance.

Les rapports de causalité et de probabilité ne sont pas intervertibles. — Pour résoudre cette question il faut partir de la constatation, indiscutable, de la non-invertibilité des rapports de causalité. Après la pluie les rues sont mouillées; mais le fait qu'une rue est mouillée ne prouve pas qu'il a plu : la rue peut avoir été arrosée, ou bien une inondation peut s'être produite, ou encore, l'eau peut s'être répandue par suite de l'éclatement d'une conduite. Bien plus : si nous n'avons quelque connaissance concernant les fréquences de la pluie, de l'arrosage des rues, des inondations et des éclatements de tuyaux, nous ne pourrons jamais dire avec quelle probabilité la circonstance que la rue est mouillée dépend plutôt de l'un que de l'autre des faits antécédents en question.

Les rapports de probabilité diffèrent des rapports de causalité en ceci qu'un même fait antécédent, ou un système de faits antécédents, ne produit pas toujours le même effet, mais produit une fois un effet, une autre fois un autre effet, toujours avec une certaine probabilité; mais, pour les uns comme pour les autres, il arrive qu'un même effet peut découler de divers faits antécédents ou systèmes de faits antécédents, et il est pareillement impossible de déterminer la probabilité qu'un effet soit la conséquence d'un certain fait antécédent plutôt que d'un autre, si l'on n'a pas (ou l'on ne suppose pas avoir) quelque connaissance de la fréquence avec laquelle ces divers faits antécédents se produisent. Cette connaissance n'est pas nécessaire, par contre, alors que, un fait antécédent étant connu, il s'agit de déterminer la probabilité de ses divers effets possibles. Par exemple, il est facile de calculer la probabilité pour que, dans un jeu de cartes ou l'autre, un joueur se trouve avoir quatre as; mais si vous voyez qu'un joueur a quatre as, vous ne pouvez calculer la probabilité que l'on joue à un jeu plutôt qu'à un autre, si vous ne savez quels sont les jeux que l'on fait habituellement et avec quelle fréquence.

Si l'on a égard à ces considérations évidentes, il est clair que l'entreprise d'invertir les théorèmes de la probabilité — et en particulier de construire des tests de significativité et de fixer des intervalles de confiance — sans avoir, ou sans supposer avoir, quelque connaissance de la fréquence des causes possibles de l'effet observé, est une entreprise vaine, et ce n'est que par erreur que quelqu'un peut croire y avoir réussi.

Voyons quelle peut être l'origine de cette erreur.

Probabilité pour qu'un résultat se produise dans un arrangement fortuit et probabilité pour qu'un résultat qui s'est produit dépende d'un arrangement fortuit. — Soit une famille avec dix filles. La probabilité qu'une telle combinaison des sexes ait lieu par hasard est bien petite : puisque les naissances de filles sont un peu moins fréquentes que celles de garçons, elle est inférieure à $\frac{1}{2^{10}} = \frac{1}{1014}$. Il y a des auteurs qui en déduisent qu'il y a une probabilité supérieure à $\frac{1013}{1014}$ pour que le couple procréateur ait tendance à mettre au monde des filles.

Dans une généalogie, on voit que l'aïeul, le père et quatre enfants sont morts d'une certaine maladie. Soit $\frac{1}{200}$ la probabilité qu'un décès se produise à cause de cette maladie ; admettant qu'il s'agisse d'une maladie non contagieuse, $\frac{1}{200^6}$ sera la probabilité pour que cette combinaison se produise par effet du hasard. Certains en concluent que la maladie en question, ou tout au moins la facilité à en être atteint, avait sûrement un caractère héréditaire, du moment qu'il n'y a qu'une possibilité sur 200^6 en faveur de l'hypothèse contraire. C'est là un raisonnement que l'on trouve fréquemment chez les généticiens pour démontrer l'importance de l'hérédité dans la détermination de certains caractères.

Une portée encore plus grande aurait le raisonnement suivant, qui se propose de démontrer l'impossibilité pratique pour que la matière organique soit dérivée de celle inorganique. La probabilité, dit-on, qu'une des plus simples molécules organiques, constituée d'à peine 2 000 atomes, se soit produite par un arrangement fortuit des atomes en question est inférieure à $\frac{1}{10^{600}}$. Il y a donc moins qu'une possibilité sur 10^{600} pour qu'elle ne soit pas la conséquence d'un dessein préétabli. C'est dans cette argumentation et dans d'autres analogues que, dans leur réaction à la conception matérialiste de l'origine de la vie, les spiritualistes cherchent un fondement objectif à leur croyance.

La validité de ces argumentations paraît incontestable à plusieurs ; mais nous apercevons clairement leur spéciosité si nous recourons à une argumentation analogue qui touche à nos intérêts, parce qu'alors notre sensibilité logique s'éveille et se fait plus aiguë.

En jouant à la loterie, vous constatez que vous avez gagné une quaterne (quatre des cinq numéros de votre billet étant sortis) et vous allez retirer votre lot. Vous trouvez au guichet un agent de police qui vous apostrophe ainsi : « La probabilité que la sortie de vos quatre numéros se produisit par hasard était de $\frac{1}{511\,038}$. Il y a donc 511 037 probabilités contre une pour que votre succès soit l'effet d'une supercherie. C'en est assez pour que je vous arrête ». Je suis bien sûr que ce raisonnement ne vous paraîtrait pas convaincant.

Pour se persuader du caractère fallacieux de ce raisonnement et des précédents, il suffit de considérer que des 1 014 arrangements des sexes possibles dans 10 naissances, des 200^6 arrangements des causes de mort qui peuvent se produire dans la suite de deux ascendants et de quatre frères, des 10^{600} et plus arrangements des 2 000 atomes, des 511 038 combinaisons de quatre des cinq numéros sortis, un arrangement ou combinaison devait toujours se produire et que celui qui s'est produit était tout aussi probable que chacun des autres. Le même raisonnement aurait pu être répété pour tout arrangement ou combinaison qui serait sorti, en concluant, avec autant de raison, qu'il ne pouvait être fortuit.

Où réside donc l'erreur ? L'erreur réside dans le fait de confondre la probabilité pour qu'un certain résultat se produise par hasard *dans* une série de combinaisons ou arrangements fortuits, avec la probabilité qu'il se produise *par* une combinaison ou arrangement fortuit, plutôt que par une combinaison ou arrangement dû à des causes systématiques. Cette erreur, qui paraît assez banale, a été et est encore commise par certains auteurs même des plus renommés. L'erreur est évidente, par exemple, dans un mémoire de Charles Pearson, auquel on peut faire remonter la construction moderne des tests de significativité due à l'École anglaise. Le mémoire est ainsi intitulé : *On the criterion that a given System of Deviations from the Probable in the Case of a correlated System of Variables is such that it can be reasonably supposed to have arisen from random Sampling* (*Philosophical Magazine*, 1900). « From random Sampling », c'est-à-dire « *par* un arrangement fortuit », dit le titre; mais voilà que dans le texte cette expression est remplacée par « on a random selection » « *dans* un arrangement fortuit », et c'est celle-ci qui correspond à la démonstration qui est effectivement donnée dans le mémoire. Pearson se proposait donc de fournir un critérium pour que l'on put raisonnablement supposer qu'un certain système d'écarts réellement observé avait été obtenu *par* un choix hypothétique fait au hasard, et il a, par contre, démontré un critèrium pour obtenir, *dans* un choix réellement fait au hasard, le système d'écarts supposé.

C'est là un saut logique qui, à travers les traités de Yule-Kendall et R. Fisher, a pris le droit de cité dans la littérature statistique, où il est accepté sous l'étiquette de « null hypothesis ». Le manuel de Miss F. N. David, qui se base invariablement sur la « null hypothesis » dans la vérification des hypothèses statistiques, date de 1953.

Or, entre la probabilité, que nous indiquerons par ${}_{a}P_{e}$, pour qu'un certain événement e se produise dans une combinaison fortuite et la probabilité que nous indiquerons par ${}_{e}\Pi_{a}$, pour que le même événement, résulte d'une combinaison fortuite, il y a un rapport extrêmement simple et bien connu ; ce rapport est

$$(1)\qquad {}_{e}\Pi_{a} = \frac{{}_{a}P_{e}}{{}_{a}P_{e} + \frac{1-p_a}{p_a}\,{}_{s}P_{e}},$$

où ${}_{a}P_{e}$ et ${}_{e}\Pi_{a}$ ont les significations ci-dessus précisées, p_a indique la probabilité d'intervention des causes accidentelles, $1-p_a$ la probabilité complémentaire pour qu'interviennent des causes non accidentelles et ${}_{s}P_{e}$ la probabilité pour que, les causes non accidentelles étant survenues, l'événement considéré se produise.

Appliquons maintenant cette formule au cas de la sortie de notre billet de loterie. En ce cas, ainsi que nous l'avons vu, ${}_{a}P_{e} = \frac{1}{511\,038}$; mais, étant admis que le jeu de la loterie a lieu en toute impartialité, il n'y a que l'intervention de causes accidentelles, c'est-à-dire que $p_a = 1$ et donc $(1-p_a) : p_a = 0$ et, par conséquent, ${}_{e}\Pi_{a} = 1$. Il y a donc certitude $\left(\text{et non pas une probabilité} = \frac{1}{511\,038}\right)$ pour que la sortie soit le résultat d'une combinaison fortuite. Naturellement, la conclusion serait différente si le jeu de la loterie n'avait pas lieu avec l'impartialité de rigueur.

Appliquons encore la formule à la formation de la molécule organique. Ici ${}_{a}P_{e} < \frac{1}{10^{600}}$; mais quelles sont les valeurs de p_a et de ${}_{s}P_{e}$? Le matérialiste se refuse à prendre en considération des causes autres que les causes naturelles et il nie la possibilité d'un dessein préétabli dans la nature : pour lui, $p_a = 1$ et partant ${}_{e}\Pi_{a} = 1$. Le croyant est convaincu du contraire : pour lui, $1-p_a = 1$ et aussi ${}_{s}P_{e}$ prend une valeur élevée, de sorte que ${}_{e}\Pi_{a}$ n'a qu'une valeur nulle.

Remarquons que, lorsque des facteurs non accidentels interviennent, de sorte que $p_a < 1$, $(1-p_a) : p_a > 0$, la valeur de ${}_{s}P_{e}$, et donc de ${}_{e}\Pi_{a}$, n'est pas égale pour tous les résultats possibles pour lesquels, par contre, la valeur de ${}_{a}P_{e}$ est égale. Pour celui qui admet que le processus de la nature est dirigé vers un but déterminé, il y a une forte probabilité ${}_{s}P_{e}$ (sinon la certitude même) que l'arrangement préétabli des atomes ait donné lieu à la molécule organique, mais il n'y a pas la même probabilité pour qu'elle ait donné lieu à un autre arrangement quelconque. Dans un pays hypothétique, où le jeu de la loterie ne se ferait pas honnêtement, il y aurait une probabilité plus ou moins grande pour qu'un lot important fût l'effet d'une supercherie, tandis qu'il n'y aurait pas de raison aussi forte pour supposer que l'effet de la supercherie soit la sortie de certains numéros dont personne ne pourrait profiter.

Il convient aussi de ne pas oublier, à ce propos, que l'utilité de la formule (1) n'est pas tant celle de permettre la détermination numérique de ${}_e\Pi_a$, qui en réalité ne peut s'opérer qu'exceptionnellement, que de donner le rapport entre ${}_aP_e$ et ${}_e\Pi_a$, ce qui est utile à un double point de vue : à un point de vue négatif, parce qu'il met en lumière combien il est arbitraire et combien il peut être dangereux de remonter de ${}_aP_e$ à ${}_e\Pi_a$ sans avoir d'autres connaissances ; à un point de vue positif, parce qu'il indique que dans des cas particuliers, où l'on a des connaissances qui concernent justement les valeurs de p_a (et par conséquent de $1-p_a$) et de ${}_sP_e$, on peut arriver à des conclusions bien fondées, même si elles ne peuvent pas toujours être exprimées numériquement, sur la valeur de ${}_e\Pi_a$.

La confusion entre problèmes directs et problèmes inverses de la probabilité ne s'est pas produite, cependant, pour la première fois dans l'École statistique anglaise moderne. En effet, elle remonte à une époque beaucoup plus reculée.

Jacques Bernoulli a démontré, comme on sait, que si l'on connaît la probabilité $\bar{v}$ d'un phénomène (supposée constante dans les observations successives), on arrive toujours, en augmentant le nombre des observations, à un point où il y a une probabilité, aussi voisine que l'on veuille de la certitude, que la fréquence observée f du phénomène s'écarte de $\bar{v}$ moins que d'une quantité donnée, aussi petite que l'on veut. C'est là le théorème, — un théorème de probabilité directe — connu comme le théorème de Bernoulli. Mais, en réalité, Bernoulli s'était proposé un problème de probabilité inverse, celui de remonter *a posteriori* de la fréquence observée $\bar{f}$ du phénomène à sa probabilité inconnue v avec une erreur inférieure à une intensité donnée, lorsque cette probabilité n'aurait pu être déterminée *a priori* au moyen du rapport des cas favorables aux cas possibles du phénomène.

Le contraste entre le théorème que Bernoulli voulait démontrer et le théorème qu'il a en réalité démontré est évident pour celui qui lit attentivement l'*Ars Conjectandi* et, s'il n'est pas généralement remarqué, cela vient — je pense — du fait que, si beaucoup de personnes parlent de Bernoulli, le nombre de celles qui l'ont lu est minime. Keynes, qui lui l'avait lu, a noté ce contraste et, pour l'expliquer, il pensait que, en réalité, Bernoulli s'était proposé, après avoir démontré le théorème direct, de démontrer le théorème inverse correspondant, mais qu'il en avait été empêché par la survenance de la mort et qu'il a laissé l'ouvrage inachevé. Or, que l'*Ars conjectandi* soit inachevé, ce n'est pas douteux, mais ce qui lui manque ce sont les applications aux matières civiles, morales et économiques, tandis que, en ce qui concerne la partie théorique, la lecture faite avec attention de l'ouvrage et des lettres échangées entre l'auteur et Leibniz ne peut laisser de doute quant au fait que Bernoulli croyait avoir complètement démontré le théorème inverse. Évidemment, il confondait un théorème de probabilité directe avec un problème de probabilité inverse. Bernoulli est considéré avec raison comme le fondateur du Calcul des Probabilités. C'est donc à l'origine même de ce Calcul que l'erreur remonte. On pourrait l'appeler le « péché originel du Calcul des Probabilités ». On comprend ainsi pourquoi, bien que ce ne soit pas la première fois qu'on l'ait mise en évidence, l'erreur soit si difficile à extirper. En fouillant dans les classiques, travail extrêmement utile qu'aujourd'hui on néglige à tort, j'ai en effet, trouvé que dès 1796 P. Prévost et S. A. Lhuilier la faisaient clairement ressortir. « Ainsi on conclue dans ce problème — remarquaient-ils, au sujet du théorème de Bernoulli — de la cause aux effets, et non des effets à la cause. » Mais leur voix se perdit, à ce qu'il paraît, dans le désert.

Du choix entre les hypothèses. — En faisant un pas en avant par rapport au procédé critiqué dans le paragraphe précédent, on a fait remarquer que la probabilité pour qu'un certain événement doive être attribué au hasard (ou, en général, à une certaine cause ou hypothèse) ne dépend évidemment pas seulement de la probabilité pour que, si cette hypothèse se réalise, l'événement s'ensuive, mais aussi de la probabilité pour que, une cause systématique (ou, en général, la cause ou l'hypothèse contraire) intervenant, l'événement se réalise. Alors, pour rendre maximum ${}_{e}\Pi_{a}$, il conviendra d'établir e de façon à rendre ${}_{a}P_{e}$ maximum et ${}_s\mathrm{P}_e$ minimum. C'est bien là la substance de la théorie de Neyman et E. S. Pearson sur le choix entre les hypothèses.

Or, cette condition est naturellement nécessaire, mais elle n'est pas suffisante. En effet, ${}_e\Pi_a$ ne dépend pas seulement de ${}_a\mathrm{P}_e$ et de ${}_s\mathrm{P}_e$, mais aussi de p_a et, tant que p_a n'est pas connu, on n'est pas à même de choisir en connaissance de cause entre les deux hypothèses contraires. Dans l'exemple du paragraphe précédent, relatif à la sortie des quatre derniers numéros à la loterie, ${}_a\mathrm{P}_e$ est certes très petit et ${}_s\mathrm{P}_e$ est certainement plus grand, et nous pouvons même supposer qu'il soit $= 1$; néanmoins, dans un jeu de loterie régulier, ${}_e\Pi_a = 1$ parce que $p_a = 1$.

D'autre part, lorsqu'on tient dûment compte de la valeur de p_a, le procédé de Neyman et E. S. Pearson conduit à la formule classique de la probabilité *a posteriori*.

Wald, qui dans son ouvrage *Sequential Analysis* (1947) prenait, comme point de départ de ses recherches, la théorie sur le choix entre les hypothèses telle qu'elle avait été formulée par Neyman et E. S. Pearson, dut s'apercevoir de l'insuffisance de ladite théorie, car dans l'ouvrage plus récent *Statistical Decisions Functions* (1950), il se base sur la formule de la probabilité *a posteriori* en vue de déterminer les fonctions de décisions qui rendent minimum le risque moyen.

Connaissances complémentaires et connaissances supplémentaires. — Nous avons fait allusion à la possibilité que les connaissances que l'on a sur les valeurs p_a et ${}_s\mathrm{P}_e$, en sus de celles de ${}_a\mathrm{P}_e$, permettent d'arriver à des conclusions fondées, mais non exprimables numériquement sur la valeur de ${}_e\Pi_a$.

D'autres fois on ne dispose pas des connaissances sur les valeurs de p_a, ${}_a\mathrm{P}_e$, ${}_s\mathrm{P}_e$ qui sont nécessaires pour déterminer la valeur de ${}_e\Pi_a$ par rapport précisément au phénomène qui nous intéresse directement, mais on a des connaissances relatives à un phénomène analogue, effectif ou hypothétique, et l'on sait qu'il diffère de celui qui intéresse par des caractéristiques connues, propres à influencer les valeurs de p_a, ${}_a\mathrm{P}_e$, ${}_s\mathrm{P}_e$, ou de quelqu'une d'entre elles, en un certain sens, mais avec une intensité qu'il n'est pas possible de préciser. En ce cas, il est possible d'exprimer numériquement une valeur ${}_e\Pi_a$ (celle pour le phénomène analogue), mais il n'est pas possible d'exprimer numériquement de combien elle diffère de la valeur de ${}_e\Pi_a$ précisément pour le phénomène qui intéresse. Et la valeur de ${}_e\Pi_a$, déterminée pour le phénomène analogue, peut servir comme valeur approchée de la valeur exacte de ${}_e\Pi_a$ pour le phénomène qui intéresse.

Les connaissances nécessaires pour déterminer la valeur approchée de ${}_e\Pi_a$ seront dites *complémentaires ;* les connaissances qui nous font juger que cette valeur s'écarte de la valeur exacte dans un sens ou dans l'autre seront dites *supplémentaires.*

Supposons que dans une île on ait introduit deux variétés A et B, de bétail également fécondes et résistantes et complètement interféscondes, dont A est totalement dominante sur B, et supposons que les représentants de la variété A étaient au nombre de 10 au moment de l'introduction et ceux de la variété B au nombre de 20, également répartis, pour les deux variétés, entre les deux sexes. Après un certain temps, pendant lequel le bétail s'est librement reproduit, le propriétaire débarque dans l'île et prend un produit x qui a les caractères de la variété A. Quelle est la probabilité que ce produit soit homozygote ? *A priori*, un produit peut être homozygote de la variété A, homozygote de la variété B ou hétérozygote. S'il y avait panmixie (c'est-à-dire si la tendance à l'accouplement entre individus ayant les mêmes caractères était la même qu'entre individus ayant des caractères divers), les probabilités p_a, p_b, p_{ab} des trois hypothèses seraient comme 1 : 4 : 4 ; d'autre part, la probabilité pour que, dans les trois hypothèses, on trouve un produit qui ait la caractéristique de celui choisi est respectivement

$$ {}_a\mathrm{P}_e = 1, \qquad {}_b\mathrm{P}_e = 0, \qquad {}_{ab}\mathrm{P}_e = 1_e, $$

de sorte que l'on aura ${}_e\Pi_a = 0,20$. Le propriétaire aura donc les connaissances complémentaires pour déterminer la valeur de ${}_e\Pi_a$ dans les conditions de panmixie. Mais, d'autre part, il sait (connaissance supplémentaire) qu'il y a, en général, chez toutes les espèces, une certaine homogamie, soit une tendance à l'accouplement entre les êtres semblables, de sorte qu'il est autorisé à croire que le nombre des individus homozygotes est en réalité supérieur à celui qui a été calculé dans l'hypothèse de panmixie, de sorte qu'il devra en conclure que ${}_e\Pi_a > 0,20$.

Permettez-moi d'appeler l'attention sur la circonstance que, pour tirer profit de ces connaissances supplémentaires, il ne suffit pas de savoir que le phénomène qui intéresse, (et pour lequel on voudrait déterminer la valeur de ${}_e\Pi_a$, sans toutefois pouvoir y arriver), diffère du phénomène analogue, pour lequel on est en mesure d'exprimer numériquement la dite valeur, il faut aussi connaître qu'il en diffère *dans un certain sens bien déterminé*.

Nous aurons l'occasion de revenir sur cette condition plus loin.

La formule générale de la probabilité des causes. — En indiquant par A, B, C, ..., Z les différentes causes, par p_a, p_b, p_c, ..., p_z les probabilités respectives *a priori*, et par ${}_a\mathrm{P}_e$, ${}_b\mathrm{P}_e$, ${}_c\mathrm{P}_e$, ..., ${}_z\mathrm{P}_e$ les probabilités que chacune d'elles donne lieu à l'événement e, la probabilité pour que, lorsque l'événement e s'est réalisé, il dépende de la cause x peut s'exprimer ainsi

$$(2) \qquad {}_e\Pi_x = \frac{p_x \, {}_x\mathrm{P}_e}{\Sigma p_i \, {}_i\mathrm{P}_e}.$$

Cette même formule reste valable si x peut résulter non pas d'une seule cause, a ou b ou c, mais d'un groupe de causes, par exemple, $a + c$, auquel cas on aura

$$p_x = p_a + p_c, \qquad {}_x\mathrm{P}_e = {}_a\mathrm{P}_e + {}_c\mathrm{P}_e.$$

Cette formule est connue sous le nom de formule de Bayes, mais en réalité Bayes avait considéré seulement l'hypothèse où les probabilités *a priori* étaient égales pour toutes les causes, tandis que c'est Laplace qui a considéré le cas de différentes probabilités *a priori*.

La formule (1) ne constitue qu'un cas particulier de la formule (2), dans lequel les causes sont groupées en deux catégories : causes accidentelles et causes systématiques.

L'inversion du théorème de Bernoulli opérée par Laplace dans l'hypothèse de l'équiprobabilité des causes. Proposition d'un schéma plus général et son application à la réception de parties de produits sur la base d'échantillons. — Lorsque les causes considérées sont, par contre, très nombreuses (théoriquement infiniment nombreuses) et toutes interviennent avec la même probabilité (suivant l'hypothèse de Bayes) et comportent des probabilités *a priori* de l'événement qui se distribuent uniformément entre les valeurs 0 et 1, Laplace a démontré que le théorème de Bernoulli peut être inversé, c'est-à-dire que la probabilité ${}_{;}\Pi_{+ef}$ pour que dans n cas se réalise une fréquence f, qui diffère de la valeur vraie correspondante (soit de la probabilité correspondante) é de plus de

$$e\sqrt{\frac{2\dot{v}(1-\dot{v})}{n}}$$

est égale, dans cette hypothèse, à la probabilité ${}_f\mathrm{II}_{+ev}$ que la probabilité inconnue v diffère de la fréquence $\dot{f}$ réalisée dans n cas, de plus de

$$e\sqrt{\frac{2\dot{f}(1-\dot{f})}{n}}.$$

Or l'hypothèse de l'équiprobabilité des causes, auxquelles correspondent différentes valeurs de 0 à 1 des probabilités *a priori* de l'événement, n'est pas généralement vérifiable, et souvent il est évident qu'elle ne correspond pas à la réalité ; de sorte que l'on comprend pourquoi plusieurs auteurs ont cherché à démontrer que cette hypothèse n'est pas nécessaire. Ils croient l'avoir démontré, pourvu que le nombre n des observations soit assez grand pour que puissent être négligés les termes de l'ordre de $\frac{1}{n}$; mais leurs démonstrations prêtent à des objections. Reprenant et développant un schéma proposé dès 1911, j'ai même démontré en collaboration avec le Dr G. Livada que, lorsque les causes correspondantes aux probabilités *a priori* de l'événement se serrent autour d'une valeur normale $\frac{k}{k+h}$ et présentent une valeur comprise entre x et $x+dx$ avec une probabilité

$$p_x = \frac{(k+h+1)!}{k!\,h!}x^k(1-x)^h\,dx, \tag{3}$$

les valeurs de ${}_f\mathrm{II}_{+ev}$ dépendent des valeurs k et h non moins que du nombre $\dot{f}n$ de fois où l'événement s'est produit et de celui $(1-\dot{f})n$ où il ne s'est pas produit. Et, comme il n'y a pas de restriction quant aux valeurs de k et de h, il paraît clair que, même lorsque le nombre n des observations est si grand que l'on peut négliger les termes de l'ordre $\frac{1}{n}$, les résultats dépendent essentiellement des valeurs de k et de h, c'est-à-dire de la distribution des causes.

Ce schéma est plus général que celui considéré par Laplace, auquel il se réduit dans le cas particulier $k=0$, $h=0$, et pratiquement il peut aussi s'appliquer, ainsi que je l'ai effectivement appliqué, au cas de réception de parties de produits effectués sur la base d'échantillons, lorsque les établissements intéressés à la réception ont précédemment présenté d'autres échantillons dont on peut tirer les valeurs de k et de h. J'ai montré que, en faisant différentes hypothèses sur les pourcentages d'éléments défectueux résultant d'échantillons présentés auparavant, la probabilité pour que dans la partie présentée le pourcentage d'éléments défectueux ne dépasse pas une certaine limite peut être très différente, suivant les pourcentages d'éléments défectueux résultant des échantillons précédents et, en tout cas, très différente de celle à laquelle on serait parvenu sur la base de l'hypothèse de Laplace concernant l'équiprobabilité des causes. Suivant cette hypothèse, on aurait dû prévoir, dans le cas de l'exemple que j'ai indiqué, une probabilité 0,1438 que le pourcentage d'éléments défectueux de la partie ne dépasserait pas 8 %, tandis que, suivant les différentes hypothèses considérées dans l'application de notre schéma, la probabilité serait 0,8789 ; 0,2968 ; 0,0004 ; 0,1607.

Il est à remarquer que l'hypothèse que toutes les probabilités *a priori* de l'événement, uniformément distribuées entre les valeurs 0 et 1, soient équiprobables correspond à l'hypothèse que toutes les *combinaisons* des facteurs favorables et contraires à l'événement soient équiprobables.

Or, c'est là une hypothèse tout à fait injustifiable, même en l'absence absolue de toute connaissance *a priori* sur la probabilité de l'événement,

Si, dans ce dernier cas, nous devions tout de même prendre une décision, la seule hypothèse compréhensible serait que soient équipossibles, voire équiprobables, non pas toutes les *combinaisons*, mais tous les *arrangements* des facteurs favorables ou contraires à l'événement, car tous ces arrangements sont également *imaginables*.

Or, cela aboutit, non pas à une équidistribution des combinaisons et, par conséquent, des probabilités *a priori* de l'événement, mais à une distribution binomiale qui ressemble beaucoup à la distribution β, donnée par la formule (3) [23].

L'introduction de ce schéma plus général, basé sur la distribution β des probabilités *a priori* et applicable à la réception des parties de produits ayant lieu sur la base d'échantillons, représente, si je ne me

---

([23]) *Voir*, à ce sujet, la note *Sulla rappresentabilità di una distribuzione binomiale mediante una distribuzione* β *e viceversa* présentée par le Docteur C. Benedetti au Séminaire de l'Institut de Statistique de l'Université de Rome (24 juin 1954), à paraître dans *Métron*, t. 18, nos 1-2, 1955.

trompe, un progrès réel pour l'étude des problèmes de la probabilité inverse, tant au point de vue théorique, qu'au point de vue pratique.

La détermination de la probabilité des événements futurs sur la base du schéma de la probabilité des causes. — De la formule (2) qui indique la probabilité des différentes causes pouvant avoir déterminé le fait e dans le passé, il est facile de passer à la probabilité pour que le fait y se produise à l'avenir.

La probabilité pour que, lorsque le fait e s'est produit dans le passé, le fait y se produise pour la cause x, sera ${}_e\Pi_x\,{}_xP_y$ et la probabilité qu'il se produise pour une cause quelconque ${}_e\Pi_y = \Sigma\,{}_e\Pi_i\,{}_iP_p$ ou, en substituant sur la base de la formule (2),

$$(4) \qquad {}_e\Pi_y = \frac{\Sigma p_i\,{}_iP_e\,{}_iP_y}{\Sigma p_i\,{}_iP_e}.$$

Si le fait e est constitué par la fréquence $\frac{m}{n}$ d'un certain événement dans n épreuves passées, et le fait y par sa fréquence $\frac{r}{s}$ dans s épreuves futures, et si l'on suppose que la probabilité de l'événement reste constante dans les épreuves passées et dans celles futures prises en considération, il est possible de calculer la valeur de ${}_e\Pi_y$ sur la base de l'hypothèse de l'équiprobabilité des causes à laquelle Laplace avait recours.

Pour le cas particulier de la fréquence de l'événement dans une épreuve future, on a démontré, ainsi qu'il est bien connu, que

$$(5) \qquad {}_{\frac{m}{n}}\Pi_{\frac{1}{1}} = \frac{m+1}{n+2}.$$

Une démonstration analogue peut être faite sur la base de notre hypothèse plus générale concernant la distribution des probabilités des causes : les formules auxquelles on arrive sont les mêmes, il faut seulement substituer $m+k$ à m et $n+k+h$ à n. Au lieu de la formule (5), on arrive, par exemple, à la formule suivante :

$$(6) \qquad {}_{\frac{m}{x}}\Pi_{\frac{1}{1}} = \frac{m+k+1}{n+k+h+4}.$$

Les formules (5) et (6) donnent, pour des hypothèses différentes, la *probabilité a posteriori de l'événement sur la base du schéma de la probabilité des causes.*

Puisque k et h peuvent atteindre des valeurs aussi élevées que l'on veut et indépendantes des valeurs m et n, la comparaison entre la formule (5) et la formule (6) nous dit encore une fois que, contrairement à l'illusion optimiste qui persiste chez bien des statisticiens et des spécialistes du Calcul des Probabilités, la distribution des causes n'est pas du tout indifférente même pour m et n très grands.

Schéma de la probabilité des résultats. — Si l'on s'en tient généralement à l'hypothèse de l'équiprobabilité des causes, c'est parce que manque pratiquement la connaissance de la probabilité des différentes causes. Aussi le schéma que nous avons proposé et appliqué présuppose-t-il, pour ses applications, des connaissances que nous ne pouvons avoir que dans des cas particuliers.

Pour cette raison, il est important non seulement théoriquement, mais aussi pratiquement, d'avoir un schéma différent de celui de la probabilité des causes, que nous avons proposé sous le nom de *schéma de la probabilité des résultats* et qui est susceptible d'être pratiquement appliqué dans de nombreux cas où la détermination de la probabilité des causes est impossible.

C'est ce que nous allons illustrer par un exemple.

Un tel a quatre filles : il voudrait avoir un garçon, mais il se doute que son ménage est un de ceux qui sont destinés à n'avoir que des filles. Quelle est la probabilité pour que se réalise l'événement qu'il souhaite ? S'il se basait sur la probabilité des causes, il ne pourrait jamais résoudre son problème, car la détermination des probabilités *a priori* pour qu'un ménage ait une probabilité plutôt qu'une autre d'avoir un garçon ou une fille dépasse les possibilités humaines. Mais n'y a-t-il pas un moyen plus pratique de poser la question ?

Les statistiques nous renseignent sur les fréquences des différentes combinaisons des sexes pour 5, 6, 7, ... enfants et, étant donné le nombre des observations sur lesquelles elles sont basées, elles peuvent être considérées comme des expressions approchées des probabilités respectives. Chacune de ces combinaisons constitue un *résultat* dont nous connaissons, bien qu'approximativement, la probabilité p_x. D'autre part, si un résultat x se produit, nous pouvons aisément calculer la probabilité ${}_x\mathrm{P}_e$ pour que se produise en même temps le fait observé, c'est-à-dire que les quatre premiers nés soient du sexe féminin. La formule valable, c'est encore la (2); seulement la signification des symboles change : p_x indique alors la probabilité du résultat x; ${}_x\mathrm{P}_e$ la probabilité pour que, le résultat x se réalisant, le fait e se réalise aussi. La formule se simplifie parce que seulement peu de résultats sont compatibles avec le fait, de sorte que ${}_x\mathrm{P}_e$ soit $\neq 0$; dans notre cas, par exemple, pour cinq enfants il suffit de considérer deux résultats : 5 f. et 4 f. 1 g.; pour six enfants, trois résultats : 6 f., 5 f. et 1 g., 4 f. et 2 g.

Voyons alors la réponse que les statistiques permettent de donner à notre père de famille.

Sur le fondement de relevés étendus effectués en Saxe pendant les dix années 1876-1885, nous pouvons dire que, parmi les familles qui ont eu une cinquième naissance, il y en avait 3 429 avec 5 filles et 16 851 avec 4 filles et 1 garçon. Si c'est le premier résultat qui se réalise, il y a certitude que les quatre premiers enfants seront du sexe féminin; si c'est le second résultat qui se réalise, il y a la même probabilité d'avoir le garçon à chacun des cinq naissances, et partant il y a la probabilité $\frac{1}{5}$ qu'il soit le dernier né, tandis que les quatre premiers accouchements ont donné des produits du sexe féminin. La probabilité pour que la cinquième naissance soit encore celle d'une fille sera donc :

$$ {}_e\Pi_5 = \frac{3\,429}{3\,429 + \frac{1}{5}\,16\,851} = \frac{3\,429}{3\,429 + 3\,370} = 0{,}50434. $$

Notre père de famille pourra donc compter sur une probabilité $= 0{,}49566$ d'avoir un garçon à la prochaine naissance.

Mais, si, à la place d'un garçon il aura une autre fille, sur quelle probabilité pourra-t-il compter d'avoir un garçon à la sixième naissance? Les mêmes statistiques nous informent que, des familles qui ont une sixième naissance, il y en a 1 408 avec 6 filles contre 8 012 avec 5 filles et 1 garçon. Un calcul analogue au précédent nous dit que la probabilité d'avoir un garçon à la sixième naissance, après qu'il y a eu 5 filles aux cinq premières naissances est 0,48676. Et, si la sixième naissance devait encore le désappointer, la probabilité d'avoir un garçon à la septième serait 0,47881. En somme, la probabilité d'avoir un garçon à la 5^{e}, 6^{e} ou 7^{e} naissance, en supposant que le père ne veuille pas insister, serait

$$\begin{aligned}&0,49566+(1-0,49566)\,0,48676\\&\quad+[1-0,49566-(1-0,49566)\,0,48676]\,0,47881\\&=0,49566+0,24548+0,12394=0,86508.\end{aligned}$$

Le père de famille, se fondant sur l'expérience, pourra donc compter sur une probabilité dépassant 86 % d'avoir le garçon désiré à une des trois naissances suivantes.

La réponse n'est qu'approchée, soit parce que l'on y arrive en substituant aux probabilités des différentes combinaisons leurs fréquences, qui en diffèrent par des erreurs accidentelles, soit parce que les fréquences ont été déterminées non pas précisément pour la population à laquelle appartient le père de famille, mais, ainsi que nous l'avons dit, pour la Saxe en 1876-1885. Mais, d'une part, on sait que les combinaisons des sexes dans les naissances ne varient que très peu d'une population à l'autre et d'une époque à une autre et, d'autre part, le nombre élevé des cas où les fréquences ont été établies permet de considérer satisfaisante leur approximation aux probabilités respectives.

Quoi qu'il en soit, une réponse approximative vaut toujours mieux que le manque de toute réponse.

Remarquons aussi que, dans le calcul, on suppose que la probabilité de la naissance d'un garçon reste constante dans les naissances successives. C'est une hypothèse implicite dans le théorème de Bernoulli et même, dans certaines limites, implicite dans toutes applications du Calcul des Probabilités.

Ce schéma peut être appliqué à plusieurs phénomènes : par exemple, à la probabilité d'une certaine composition par le sexe d'un accouchement multiple, lorsque le sexe du premier enfant est connu; à la probabilité que, dans un concours dont l'issue dépend d'un certain nombre d'épreuves, le vainqueur soit un des concurrents, probabilité déterminée après que l'on a appris le résultat d'une desdites épreuves; à la probabilité que la réunion d'un comité ne puisse avoir lieu lorsque quelques-uns des membres ont annoncé leur impossibilité d'être présents, et ainsi de suite.

La détermination de la probabilité des événements futurs sur la base du schéma de la probabilité des résultats. — Comme la formule (2) relative à la détermination des probabilités des causes peut être étendue à la détermination des probabilités des résultats, de même la formule (4) relative à la détermination des probabilités des événements futurs sur la base du schéma des probabilités des causes, peut s'étendre à la détermination de la probabilité sus-mentionnée sur la base du schéma des probabilités des résultats. En ce cas p_i indiquera la probabilité d'un résultat i, ${}_i\mathrm{P}_e$, la probabilité que ce résultat i ait donné lieu à l'événement passé e, et ${}_i\mathrm{P}_y$ la probabilité pour que ce résultat i donne lieu à l'événement futur y.

Dans le schéma des résultats, il y a cependant ceci de particulier que, quelquefois, les résultats peuvent être choisis de sorte que seulement l'un d'eux correspond à l'événement futur. En ce cas, pour une valeur de i, on a ${}_i\mathrm{P}_y = 1$, et pour les autres ${}_i\mathrm{P}_y = 0$, et la formule (4) se réduit à la formule (2). C'est là le cas sus-mentionné du père de famille qui, après avoir eu 4 filles, désire savoir la probabilité d'un garçon. Le résultat considéré (4 f. et 1 g.), dont nous avons déterminé la probabilité correspond précisément à l'événement futur (naissance d'un garçon) dont nous désirons déterminer la probabilité. En déterminant la proba, bilité du résultat, nous avons déterminé aussi la probabilité de l'événement futur. En ce cas, nous avons appelé *résultats directs* les résultats en question.

Mais il n'en est pas toujours ainsi : il pourrait se faire que nous ne connaissions pas les résultats directs, mais les résultats relatifs à un nombre plus grand de cas. Si, par exemple, nous ne connaissions pas la composition de familles avec cinq enfants, mais seulement celle des familles avec huit enfants, nous aurions cinq résultats pouvant donner lieu à 4 filles aux quatre premières naissances, et précisément : 4 f. 4 g., 5 f. 3 g., 6 f. 2 g., 7 f. 1 g., 8 f. Ayant trouvé par la formule (2) la probabilité des cinq résultats, nous pourrions calculer ensuite par la formule (4) les probabilités qui donnent lieu à un garçon à la cinquième naissance. En ce cas, où les résultats se rapportent à un nombre de cas plus grand que les résultats directs, nous avons parlé de *résultats indirects.*

Si nous indiquons par p_i la probabilité du $i^{\text{ième}}$ résultat indirect, par ${}_i\mathrm{P}_y$ la probabilité qu'il corresponde au $y^{\text{ième}}$ résultat direct et par ${}_y\mathrm{P}_e$ la probabilité pour que le $y^{\text{ième}}$ résultat direct corresponde à l'événement passé e, la probabilité pour que l'événement passé e corresponde au $y^{\text{ième}}$ résultat direct sera

$$(7) \qquad {}_e\Pi_y = \frac{\Sigma p_i\, {}_i\mathrm{P}_y\, {}_y\mathrm{P}_e}{\Sigma p_i\, {}_i\mathrm{P}_y}.$$

Dans le cas sus-mentionné, où les résultats directs ne sont pas connus, mais sont connus les résultats indirects relatifs à un nombre de cas plus grand (que nous nommerons dorénavant *résultats indirects supérieurs*), cette formule vient à coïncider avec la formule (4); elle présente cependant l'avantage de pouvoir être appliquée aussi à l'autre cas où l'on ne connaît ni les résultats directs ni les résultats indirects supérieurs, mais seulement les résultats indirects relatifs à un nombre de cas moindre que celui des résultas directs (que nous appellerons dorénavant *résultats indirects inférieurs*).

Par exemple, après qu'il a eu un garçon et une fille, le père de famille a intérêt à savoir la probabilité pour qu'il ait deux garçons aux deux naissances suivantes, mais il ne connaît la composition des sexes que pour les familles avec trois enfants. De cette composition on pourra alors calculer la probabilité de la composition des sexes dans les familles avec quatre enfants, en supposant que la probabilité d'avoir un garçon ou une fille soit, à la quatrième naissance, égale à ce qu'elle était pour les trois premières naissances, et cela quelle qu'ait été la combinaison qui s'était réalisée pour celles-ci. Connaissant la probabilité des quatre résultats directs : 3 g. 1 f., 2 g. 2 f., 1 g. 3 f., nous pourrons calculer la probabilité pour que chacun de ceux-ci donne lieu à deux garçons aux deux dernières naissances.

Remarquons à ce sujet que, pour passer de la probabilité des résultats indirects inférieurs à la probabilité des résultats directs, il faut une hypothèse de plus que pour passer de la probabilité des résultats indirects supérieurs à la probabilité des résultats directs, ou de la probabilité des résultats directs à la probabilité de l'événement considéré. Dans l'exemple de la composition des familles par le sexe, il est suffisant, dans ces deux derniers cas, d'admettre que la probabilité d'avoir un garçon ou une fille reste constante dans la suite des naissances qui donne lieu à une certaine combinaison des sexes, mais, dans le premier cas. il faut aussi admettre que la probabilité d'avoir un garçon ou une fille soit égale aux naissances futures, quelle qu'ait été la combinaison des sexes aux naissances précédentes.

L'inversion de la probabilité des erreurs dans le cas de grandeurs extensives. — Le théorème de Bernoulli, le schéma de la probabilité des causes, ainsi que celui de la probabilité des résultats et les problèmes de probabilité inverse considérés jusqu'ici, se rapportent tous à des grandeurs intensives et en particulier à des fréquences relatives.

On a moins étudié l'inversion de la probabilité des erreurs dans le cas de grandeurs extensives. Gauss avait bien opéré l'inversion de la probabilité de commettre une erreur d'observation, mais seulement par rapport a l'hypothèse que la distribution se conforme à la courbe qui a été nommée d'après lui. En outre, il subordonne l'inversion à l'hypothèse qu'*a priori* toutes les valeurs de la grandeur vraie soient également probables, hypothèse que l'on ne trouve pas précisée, par contre, chez autres auteurs (et, parmi eux, même Laplace) qui ont opéré cette inversion avant et après lui.

J'ai repris la question d'un point de vue plus général, c'est-à-dire en abandonnant la restriction de la forme gaussienne de la distribution et j'ai montré que la légitimité de l'inversion est subordonnée à deux hypothèses :

a. que la valeur vraie d'une grandeur θ puisse assumer *a priori* avec la même probabilité une quelconque des valeurs inférieures ou supérieures à $\dot{\theta}$ desquelles peut dériver la valeur observée $\dot{t}$;

b. que la courbe de distributions des valeurs observées t ne varie pas si la valeur de θ varie, au moins entre les limites dont peut dériver la valeur observée $\dot{t}$.

Les hypothèses susdites ne sont pas identiques à celles sur la base desquelles Laplace a démontré l'inversion du théorème de Bernoulli pour les fréquences relatives. Tandis que l'hypothèse *a* trouve en effet un pendant dans l'hypothèse analogue de Laplace sur l'équiprobabilité des causes qui correspondent aux probabilités *a priori* comprises entre o et 1, l'hypothèse *b* ne peut pas se réaliser dans le schéma de Laplace concernant des fréquences relatives, parce que la courbe binomiale suivant laquelle celles-ci se distribuent varie de forme, comme l'on sait, suivant que les termes du binome sont égaux ou différents et, en ce second cas, suivant qu'ils sont plus ou moins différents. L'hypothèse *b* ne peut se réaliser que pour des grandeurs absolues.

Les hypothèses en question sont d'accord, par contre, avec celles sur lesquelles se base la démonstration fournie par Gauss de l'inversion de la probabilité de commettre une erreur d'observation. Cette démonstration présuppose, en effet, explicitement l'hypothèse *a* et admet en outre que les valeurs observées se disposent toujours — quelle que soit la valeur vraie — suivant la courbe des erreurs accidentelles et avec la même variabilité, ce qui correspond à un cas particulier de l'hypothèse *b*. Partant, la démonstration de Gauss concerne le cas d'une courbe de distribution particulière; la mienne, le cas d'une courbe de distribution quelconque.

Au moyen d'un exemple approprié, j'ai aussi examiné séparément la portée que peuvent avoir sur les résultats la non-réalisation de l'hypothèse *a* et la non-réalisation de l'hypothèse *b*. Dans un cas et dans l'autre, la portée est notable.

Le système des hypothèses susdites, s'il est suffisant, n'est cependant pas nécessaire, en ce sens qu'il peut être remplacé par des systèmes d'hypothèses différents, qui sont cependant suffisants.

L'inversion dont il s'agit est, en vérité, également autorisée lorsque les deux hypothèses suivantes se réalisent :

α. que le logarithme de la valeur vraie de la grandeur θ puisse assumer *a priori* avec la même probabilité l'une quelconque des valeurs inférieures ou supérieures à $\dot{\theta}$ dont peut dériver la valeur observée $\dot{t}$.

β. que la courbe de distribution des logarithmes des valeurs observées ne varie pas alors que varie la grandeur θ, au moins dans les limites dont peut dériver la valeur observée $\dot{t}$.

Ces deux hypothèses peuvent se considérer réalisées (ce qui n'est pas le cas pour les deux précédentes *a* et *b*) au cas où la grandeur vraie θ est représentée par l'indice de variabilité d'un phénomène collectif et la grandeur observée $\dot{t}$, par l'indice de variabilité d'un de ses échantillons. Partant, le passage de l'indice de variabilité de l'échantillon à l'indice de variabilité du phénomène collectif est autorisé lorsque l'on envisage les deux hypothèses α et β,

* * *

Répondant à la question que nous nous étions posée en commençant, nous pourrons affirmer que le passage de la grandeur observée à la valeur probable de la grandeur vraie ou à la valeur probable d'une certaine erreur dans sa détermination, présuppose non seulement les connaissances nécessaires pour passer de la grandeur vraie à la valeur probable de la grandeur observée ou d'une certaine erreur de sa détermination, mais aussi d'autres connaissances (ou, à leur défaut, des hypothèses) concernant les différentes causes ou les différents résultats compatibles avec la grandeur observée.

La détermination de ces suppositions préalables est nécessaire pour que l'inférence inverse puisse avoir un caractère rigoureux; elle peut avoir ces deux buts :

A. Déterminer les suppositions préalables suffisantes pour que soit autorisée une certaine inférence inverse : par exemple, déterminer les hypothèses suffisantes pour que soit autorisée l'inversion du théorème de Bernoulli, ou pour que soit autorisée l'inversion du passage à l'écart quadratique moyen d'un phénomène collectif de l'écart quadratique moyen d'un de ses échantillons;

B. Certaines suppositions préalables étant données, déterminer quelle est l'inférence inverse autorisée : par exemple, déterminer la valeur probable de la probabilité p après que se soit réalisée dans n observations une fréquence $\frac{m}{n}$ dans l'hypothèse que, antérieurement aux n observations, toutes les valeurs de p étaient également probables (hypothèse de Laplace), ou qu'elles se distribuaient suivant la formule (3) (notre hypothèse). Comme nous l'avons dit, dans l'hypothèse de Laplace, la valeur probable de p est $\frac{m+1}{n+2}$, dans notre hypothèse elle est $\frac{m+k+1}{n+k+h+2}$.

Conclusion. — Dans ce chapitre nous avons traité le problème de l'inversion des relations statistiques en suivant les procédés classiques qui s'honorent des noms de Laplace et de Gauss. Nous avons montré, si nous ne nous faisons pas illusion, qu'ils peuvent être généralisés au point de vue théorique sur la base d'hypothèses plus amples que celles considérées par leurs auteurs et aptes à en autoriser des applications plus étendues. Nous avons, d'autre part, intégré les procédés basés sur la connaissanse des causes, avec les procédés basés sur la connaissance des résultats, procédés qui peuvent entrer dans les mêmes schémas mathématiques, mais permettent aussi des applications à de nouveaux domaines.

J'examinerai maintenant d'autres procédés qui ont pris pied ces derniers temps et je tirerai quelques conclusions générales sur l'inversion statistique [24].

II. — PROCÉDÉS NON CLASSIQUES.

Dans ce qui précède j'ai appelé *classiques* les procédés d'inversion statistique qui ont été introduits par Laplace et Gauss et qui ont été universellement acceptés, après. Ceux que j'ai introduits et appliqués moi-même, peuvent être considérés comme des généralisations, soit dans le sens qu'ils se basent sur des hypothèses plus générales que celles considérées par Laplace et Gauss, soit dans le sens qu'ils peuvent être exprimés par les mêmes formules, en attribuant pourtant aux symboles une signification différente qui les rend susceptibles de nouvelles applications.

Dans ce qui suit je traiterai le cas d'autres procédés que, pour employer une dénomination neutre, j'appellerai *non classiques*, mais que je serais tenté d'appeler *hétérodoxes*, parce que, à mon avis, les conclusions que l'on en déduit, ne sont pas acceptables, soit parce qu'elles sont erronées, soit parce que l'on prétend les appliquer au-delà du domaine de leur validité.

Parmi ces procédés, il y en a quelques-uns qui avaient été déjà proposés par des auteurs classiques, mais dont la proposition n'avait pas eu de suite. Seulement après avoir été nouvellement proposés par des

---

([24]) Plusieurs des sujets touchés dans ce chapitre IV.1 sont résumés dans l'article *Intorno alle basi logiche e alla portata gnoseologica del metodo statistico* (*Statistica*, 1945-1946, traduit en espagnol dans les *Actas del Instituto de Actuarios españoles*, 1946,). Une édition mise à jour et complétée a paru en langue portugaise sur la *Revista Brasileira de Estatistica*, n° 35, 1948. Dans cet article on peut trouver aussi les citations des travaux précédents dans lesquels les différents sujets ont été traités plus amplement. Qu'il suffise de rappeler ici que la généralisation du procédé d'inversion de Laplace, ainsi que le schéma des résultats, ont été introduit dès 1911, dans le mémoire *Considerazioni sulle probabilità a posteriori e applicazioni al rapporto dei sessi nelle nascite umane*, dans les *Studi economico-giuridici della R. Università di Cagliari* (reproduit dans *Metron*, t. 15, 1949) et que la généralisation du procédé d'inversion de Gauss a été introduite dans la communication *Sulla probabilità inversa nel caso di grandezze a distribuzione costante*, présentée à la VIII^e Réunion de la Société Italienne de Statistique (*Voir* Actes de ladite Réunion, Rome, 1945).

auteurs modernes, ils ont pris pied, de façon que l'on ne peut pas dire qu'ils descendent du filon classique.

* * *

Le premier des procédés à écarter, parce que ne représentant, en réalité, qu'une tautologie, est le suivant : p étant la probabilité d'un événement, soit Φ la probabilité pour que sa fréquence f dans n cas soit comprise entre les limites $p+\varepsilon$ et $p-\varepsilon$, c'est-à-dire que se produise l'inégalité

$$p-\varepsilon<f<p+\varepsilon. \tag{1}$$

L'inégalité peut être écrite aussi

$$f-\varepsilon<p<f+\varepsilon \tag{2}$$

laquelle, étant la même que la précédente sous une autre forme, a évidemment la même probabilité Φ.

Ce procédé ne représente pas du tout une inversion, comme quelque auteur l'a prétendu. C'est seulement un *passage formel* ou *tautologique*. La relation (2) comme la (1), nous enseigne à passer de la valeur connue de p à la valeur inconnue de f; pour avoir une inversion, il faudrait en déduire une relation qui nous enseigne à passer de la valeur connue de f à la valeur inconnue de p.

Si le malentendu a pu se produire, c'est à cause de la méthode usuelle d'indiquer de la même façon la variable indépendante et la fonction. Si l'on convenait — ainsi que je l'ai proposé [25] et que je le ferai dans la suite de ce chapitre — d'indiquer la variable indépendante par une lettre minuscule et la fonction par une lettre majuscule, le malentendu deviendrait impossible.

En écrivant la relation (1) de la façon suivante :

$$p-\varepsilon<\mathrm{F}<p+\varepsilon \tag{1 bis}$$

---

[25] *Voir* la communication, présentée à la XXVII^e Session de l'Institut International de Statistique (New-Delhi-Calcutta), *On some Symbols that may be usefully employed in Statistics* (*Bulletin of the Int. Statistical Institute*, t. 33, part II, 1951). Quelques erreurs typographiques qui n'avaient pas pu être corrigées dans le texte de cette communication, ont été corrigées dans les tirés à part et après dans l'édition italienne qui a paru dans *Metron*, t. 17, n^{os} 1-2, 1953.

et, par conséquent, la relation (2) de la façon suivante :

(2 *bis*) $$F - \varepsilon < p < F + \varepsilon,$$

il serait impossible de se méprendre sur la signification du passage. Pour avoir une inversion de la relation (1), il faudrait en déduire la relation (se produisant avec la même probabilité Φ)

(2 *ter*) $$f - \varepsilon < P < f + \varepsilon$$

ou l'autre équivalente

(1 *ter*) $$P - \varepsilon < f < P + \varepsilon.$$

Les relations (1 *ter*) et (2 *ter*) sont toutes les deux *inverses* par rapport aux relations (1 *bis*) et (2 *bis*) qui, toutes les deux, peuvent être dites *directes*.

*
* *

Retenons donc que l'*inversion* est une opération dans laquelle on passe de la relation entre une grandeur a considérée comme variable indépendante, et une autre grandeur B considérée comme fonction de a, à la relation entre b, considérée comme variable indépendante et A, considérée comme fonction de b.

Par conséquent, nous disons *inverses* deux *relations*, dont l'une est déduite de l'autre, entre deux grandeurs, si la grandeur, qui dans une des relations figure comme variable indépendante, figure dans l'autre comme fonction et vice versa.

Je crois qu'il est opportun de faire une distinction entre la notion que nous venons de définir, de relations inverses et celle de *relations réciproques*. Deux relations sont réciproques lorsqu'elles intéressent les mêmes grandeurs, mais la grandeur qui dans l'une des relations figure comme variable indépendante, figure dans l'autre comme fonction et viceversa. Or, il n'est pas nécessaire que, des deux relations réciproques, l'une soit déduite de l'autre; elles peuvent bien être établies indépendamment l'une de l'autre. La notion de relations réciproques est donc plus étendue que celle de relations inverses. Une relation inverse est une relation réciproque obtenue au moyen d'une inversion.

Par exemple, si, sur des données statistiques, j'établis la relation entre l'âge de l'épouse et l'âge moyen de l'époux, d'un côté, et la relation entre l'âge de l'épouse et l'âge moyen de l'époux, de l'autre, j'ai deux relations réciproques, non inverses.

Si, de la relation entre probabilité (variable indépendante) et fréquence (fonction), je déduis la relation entre fréquence (variable indépendante) et probabilité (fonction), la deuxième relation sera non seulement réciproque, mais aussi inverse de la première.

Même ainsi précisée, l'inversion comprend des opérations très différentes.

Il faut d'abord distinguer entre *inversion statistique* — qui nous intéresse spécialement — et *inversion analytique*. Et, pour bien en saisir la différence, il faut distinguer les *relations fonctionnelles* des *relations statistiques* et, parmi ces dernières, faire des subdivisions ultérieures.

Que toutes ces divisions et subdivisions et les définitions qu'elles comportent ne vous paraissent pas excessives ! Leibniz disait que la moitié des discussions seraient évitées si l'on définissait précisément les termes. Ce qui va suivre montrera — si je ne me trompe — comment des définitions précises et détaillées peuvent être utiles pour éliminer les divergences qui persistent dans la question de l'inversion des probabilités.

Dans une *relation fonctionnelle* à chaque valeur de la variable indépendante correspond une ou plusieurs valeurs individuelles de la fonction : une valeur si la fonction est biunivoque (par exemple, $A = kb$), plusieurs valeurs dans le cas contraire (par exemple, $A = k\sqrt{b}$).

Dans une *relation statistique*, au contraire, à chaque valeur de la variable indépendante correspond une collection de valeurs de la fonction et la relation s'établit entre la valeur ou plusieurs valeurs de la variable indépendante et une constante statistique que nous appellerons *valeur statistique*, concernant la collection des valeurs de la fonction (telle la moyenne ou le pourcentage des valeurs au-dessus ou au-dessous d'une certaine limite). Par exemple, aux époux d'un âge donné correspondent des âges très différents des épouses ; si par m on indique l'âge d'un groupe d'époux et avec $\bar{F}$ l'âge moyen de leurs épouses on trouve, dans certains pays, la relation statistique

$$\bar{F} = 10 + 0,5\, m.$$

Elle donne l'âge moyen de l'épouse en fonction de chaque âge de l'époux et peut, par conséquent, être appelée *relation statistique descriptive*.

Dans d'autres cas, on établit la relation entre le total de toutes les valeurs de la variable indépendante ou leur moyenne et une valeur statistique de la collection des valeurs de la fonction. Par exemple, si on désigne par M_m la moyenne des âges de tous les époux et par M_F la moyenne des âges correspondants de leurs épouses, l'expression

$$M_F = M_m - 4$$

représentera une *relation statistique globale* ou *moyenne*.

Il y a des cas intermédiaires dans lesquels la relation est établie entre un groupe de valeurs de la variable indépendante et une valeur statistique concernant la collection des valeurs correspondantes de la fonction. On peut l'appeler *relation statistique composite*. Si l'on indique par p la probabilité de tirer d'une urne une boule blanche, par f la fréquence des boules extraites dans n tirages, et par $d = f - p$ l'erreur de la fréquence par rapport à la probabilité, la probabilité $\overline{P}_{|d|}$ de commettre une erreur ayant une des valeurs symétriques $+d$ ou $-d$ est donnée par l'expression bien connue

$$\overline{P}_{|d|} = \frac{2}{\sqrt{2\pi np(1-p)}} e^{\frac{-d^2}{2np(1-p)}}$$

qui fournit un exemple de relation statistique composite.

Il est à remarquer que l'on peut aussi avoir des *relations fonctionnelles composites* et des *relations fonctionnelles globales* ou *moyennes*.

Si l'on a une courbe de fréquence symétrique et si l'on indique par $v_{+\varepsilon}$ une valeur qui présente l'écart $+\varepsilon$ de la moyenne arithmétique M et par $v_{-\varepsilon}$ la valeur symétrique qui présente l'écart $-\varepsilon$, on aura

$$M = \frac{1}{2}(v_{+\varepsilon} + v_{-\varepsilon})$$

qui fournit un exemple de *relation fonctionnelle composite*.

D'autre part, si l'on indique par v_i la valeur du terme $i^{\text{ième}}$ et par n le nombre des termes, l'expression

$$M = \frac{1}{n}\sum_{i=1}^{n} v_i$$

qui exprima que la moyenne arithmétique est égale à la somme des termes divisés par leur nombre, fournit l'exemple d'une *relation fonctionnelle globale ou moyenne*.

Par opposition aux relations fonctionnelles globales et composites, nous appellerons *relations fonctionnelles descriptives* celles qui donnent la valeur de la fonction pour chaque valeur de la variable indépendante.

En résumant, il y a lieu de faire la classification suivante :

A. *Relations fonctionnelles :*

a. descriptives;

b. composites;

c. globales.

B. *Relations statistiques :*

a. descriptives;

b. composites;

c. globales.

ou, si l'on préfère, la classification suivante :

A. *Relations descriptives :*

a. fonctionnelles;

b. statistiques.

B. *Relations composites :*

a. fonctionnelles;

b. statistiques.

C. *Relations globales :*

a. fonctionnelles;

b. statistiques.

* * *

Passons maintenant à l'application des divers procédés d'inversion auxdites relations.

L'inversion analytique peut être appliquée à toute relation descriptive (fonctionnelle ou statistique) qui soit monotone (c'est-à-dire qui ne présente pas des maxima ou des minima à l'intérieur de l'intervalle).

L'inversion analytique est le résultat d'un passage tautologique accompagné de la substitution de la variable indépendante à la fonction et vice versa. J'appellerai *rotation fonctionnelle* cette deuxième opération. Par exemple, l'inversion analytique de la relation fonctionnelle

$$\text{(1)} \qquad \mathrm{A} = \sqrt[3]{b}$$

porte à l'expression

$$\text{(2)} \qquad \mathrm{B} = a^3.$$

De la formule (1) on obtient en effet par un passage tautologique

$$\mathrm{A}^3 = b$$

dont, par une rotation fonctionnelle, on parvient à la formule (2).

De même, l'inversion analytique de la relation

$$\text{(3)} \qquad \overline{\mathrm{F}} = 10 + 0{,}5\, m$$

porte à l'expression

$$\text{(4)} \qquad \mathrm{M} = 2(\overline{f} - 10).$$

Il faut préciser la signification de la dernière expression. Elle signifie que, si l'on connaît l'âge moyen $\overline{f}$ d'un groupe d'épouses dont les époux ont tous le même âge, cet âge inconnu M des époux en question, est donné par la formule (4). La formule (4) ne serait pas applicable à un groupe d'épouses quelconque.

La formule (4) exprime une valeur individuelle M en fonction d'une moyenne de valeurs de la variable indépendante. Ce n'est plus une relation descriptive, mais une relation fonctionnelle composite.

Nous pouvons précisément dire d'une façon générale que l'inversion analytique appliquée à une relation fonctionnelle descriptive conduit à une autre relation fonctionnelle, tandis que, appliquée à une relation statistique descriptive, elle conduit à une relation fonctionnelle composite.

Si dans les formules (3) ou (4) on substitue à la valeur moyenne $\overline{\mathrm{F}} = \overline{f}$ de l'âge des épouses la valeur individuelle $\mathrm{F} = f$ et à la valeur individuelle $m = \mathrm{M}$ de l'âge de l'époux la valeur moyenne des âges des époux qui se sont mariés aux femmes ayant l'âge F, on obtient les formules suivantes :

$$\text{(5)} \qquad \mathrm{F} = 10 + 0{,}5\, \overline{m},$$

$$\text{(6)} \qquad \overline{\mathrm{M}} = 2(f - 10)$$

qui sont tout à fait différentes des formules (3) et (4) et ne sont pas du tout valables.

Nous appellerons *rotation statistique* l'opération avec laquelle on passe de la formule (3) à (5) ou de la (4) à (6). Elle consiste à substituer à la valeur statistique de la variable indépendante ou de la fonction, une valeur individuelle équivalente, et vice versa, à la valeur individuelle de la fonction ou respectivement de la variable indépendante, une valeur statistique équivalente.

Or, l'inversion statistique est le résultat d'une inversion analytique accompagnée d'une rotation statistique.

Il est indifférent d'appliquer d'abord l'inversion analytique et après la rotation statistique ou vice versa. Par exemple, pour faire l'inversion statistique de la formule (3) on peut passer par une rotation statistique de celle-ci à la formule (5) et après, par une inversion analytique, à la formule (6), ou bien on peut de la formule (3) obtenir par une inversion analytique la formule (4) et de celle-ci parvenir à la formule (6) par une rotation statistique. Pour que l'inversion statistique soit valable, c'est-à-dire pour que l'inversion statistique d'une relation statistique qui correspond à la réalité conduise à une autre relation statistique qui corresponde à la réalité, il faudrait :

1° que soit valable l'inversion analytique;

2° que soit valable la rotation statistique.

Or, la rotation statistique n'est pas valable en règle générale et par conséquent l'inversion statistique en général n'est pas du tout autorisée et conduit à des résultats erronés.

Vous vous en rendez compte tout de suite en considérant que la rotation statistique équivaut à la substitution de la ligne de régression de a sur b à la ligne de régression de b sur a. Elle ne serait autorisée que dans les cas où les deux lignes coïncideraient, c'est-à-dire lorsqu'il y aurait une seule ligne de régression. Or cela peut se produire seulement dans des cas exceptionnels et, en tout cas, lorsqu'il sagit de deux lignes finies, excepté pour leurs extrêmes ([26]).

([26]) *Voir*, pour des précisions à ce sujet, notre note *Sui limiti dell'invertibilità delle relazioni statistiche*, qui va paraître dans *Metron*, t. 18, n° 1-2, 1955.

On en conclut que, sauf dans ces cas exceptionnels et dans lesdites limites, l'inversion statistique d'une relation statistique descriptive n'est pas autorisée.

On entend par là qu'elle ne peut pas conduire à une relation statistique descriptive réciproque.

Dans le cas de la relation entre les âges des époux, la relation descriptive réciproque de celle exprimée par la formule (3) résulte à peu près la suivante :

$$\overline{M} = 15 + 0,5 f$$

qui est très différente de la (6) à laquelle on serait parvenu en appliquant à la (3) l'inversion statistique.

Tout cela — direz-vous — est bien élémentaire ! En effet, c'est très élémentaire; mais c'est précisément parce qu'on oublie les choses les plus élémentaires que se produisent les erreurs les plus répandues et les plus difficiles à extirper.

C'est parce qu'on ne comprend pas qu'une ligne de régression est une mesure unilatérale de la relation entre deux grandeurs que l'on continue à décider de l'aptitude au service militaire d'après le rapport du périmètre thoracique moyen à la taille individuelle.

J'ai en vain démontré, à la veille de la dernière guerre, que cela porte à classer d'une façon erronée environ 15 % des recrues ([27]). Tout le monde m'a donné raison, mais on a vite oublié ma démonstration et rien n'a été modifié — que je sache — quant aux critériums suivis dans le passé.

C'est encore pour ne pas se rappeler au juste moment de la diversité des deux lignes de régression que l'on persiste à faire, sur les relations statistiques descriptives, des inversions non autorisées.

Mais dirons-nous que l'inversion statistique des relations statistiques descriptives n'a aucune valeur ? Ce serait aller trop loin. En effet, lorsqu'une relation statistique descriptive est linéaire, elle est valable

---

([27]) *Voir* l'article *Principes à reviser. Les bases de la Biotypologie et les critériums du recrutement militaire*, paru dans la *Revue anthropologique*, juillet-septembre 1939. Cet article reproduit une conférence donnée à la Faculté de Médecine de Bruxelles. Un examen approfondi de la question avait été fait dans deux articles publiés sur la *Revue de l'Institut Int. de Statistique*, 1937, ayant pour titre : *Une question importante pour la science des constitutions et pour la médecine militaire : comment juger si les proportions d'un individu sont normales ?*

aussi comme relation statistique moyenne; mais une relation statistique moyenne est toujours invertible. Nous en pouvons conclure que l'inversion d'une relation descriptive linéaire conduise toujours à une relation qui, tout en n'étant pas valable pour les valeurs individuelles de la variable indépendante, est valable pour ses valeurs moyennes. Nous dirons que lesdites relations statistiques sont statistiquement *subinvertibles*.

Par exemple, nous avons vu que la relation

$$\overline{M} = 2(f - 10) \tag{6}$$

obtenue par inversion statistique de

$$\overline{F} = 10 + 0,5\,m \tag{3}$$

n'est pas valable comme relation descriptive. Elle est pourtant valable comme relation moyenne. En effet, en indiquant par A_m ou a_m les moyennes pondérées de toutes les valeurs de m et par A_f ou a_f les moyennes (pondérées également) de toutes les valeurs correspondantes de F, on obtient

$$A_f = 10 + 0,5\,a_m$$

qui est invertible et donne lieu à la relation

$$A_m = 2(a_f - 10).$$

Celle-ci n'est autre chose que la relation (6) appliquée aux moyennes. On pourrait aussi adopter le symbolisme suivant :

$$\overline{M} \lesseqgtr 2(f - 10)$$

qui signifie que la valeur $\overline{M}$ est quelquefois supérieure, d'autres fois inférieure à la valeur de $2(f - 10)$, mais les valeurs supérieures et inférieures se compensent de façon qu'en moyenne $\overline{M}$ est égal à $2(f - 10)$.

Il va sans dire qu'une relation moyenne nous dit beaucoup moins qu'une relation descriptive et, par conséquent, qu'une *subinversion statistique* est un procédé beaucoup moins efficace qu'une *inversion statistique;* mais en tout cas c'est plus que rien.

Nous reviendrons sur ce point.

*
* *

Vous espérez peut-être que je suis arrivé à la fin de toutes ces distinctions et sous-distinctions, mais il n'en est rien.

Si vous prenez le cas le plus célèbre d'inversion que l'on rencontre dans le Calcul des probabilités et dans la Statistique théorique — celui qui concerne le théorème de Bernoulli — vous trouvez qu'il ne s'agit pas d'une inversion analytique, ni d'une inversion statistique. ni d'une subinversion statistique, telles que nous les avons définies. C'est quelque chose de bien différent.

Le théorème inverse s'obtient en effet du théorème direct de Bernoulli en écrivant le symbole de la variable indépendante à la place du symbole de la fonction et le symbole de la fonction à la place de celui de la variable indépendante et en laissant inaltérée la relation telle qu'elle figurait dans la relation directe. Nous appellerons *symétrique* une telle *inversion.*

L'inversion symétrique peut être accompagnée ou non d'une rotation statistique. Si elle en est accompagnée, nous parlerons d'*inversion statistique symétrique.*

Dans le cas du théorème de Bernoulli, il s'agit d'une inversion statistique symétrique.

L'inversion symétrique peut s'appliquer à toute relation fût-elle fonctionnelle ou statistique, fût-elle descriptive ou moyenne ; l'inversion statistique symétrique, seulement à des relations statistiques : les deux ne conduisent qu'exceptionnellement à une relation valide.

Prenons la relation fonctionnelle

$$y = k(x - a); \tag{7}$$

l'inversion symétrique donne

$$x = k(y - a)$$

qui est conciliable avec (7) pour $k = -1$.

La relation

$$\overline{F} = 10 + 0,5\, m \tag{3}$$

a donné par inversion analytique, la relation

$$M = 2(\overline{f} - 10) \tag{4}$$

qui est valable comme relation fonctionnelle composite et, par inversion statistique

$$\overline{M} = 2(f - 10) \tag{6}$$

qui n'est pas valable comme relation statistique descriptive, mais seulement comme relation statistique moyenne.

Par inversion symétrique, on obtient au contraire

$$M = 10 + 0,5\,\overline{f} \tag{8}$$

et par inversion symétrique statistique

$$\overline{M} = 10 + 0,5\,f \tag{9}$$

qui ne sont pas valables dans le sens qu'elles ne correspondent pas à la réalité.

La relation descriptive réciproque de la formule (3), qui correspond à la réalité, est représentée, ainsi que nous l'avons dit, par l'expression

$$\overline{M} = 15 + 0,5\,f \tag{10}$$

très différente des expressions (4), (6), (8) et (9).

*
* *

Revenons à l'inversion du théorème de Bernoulli.

Le théorème de Bernoulli dit que, si un événement a une probabilité p — que l'on suppose constante — de se réaliser, il y a une probabilité Φ_t pour qu'en n observations l'événement se réalise avec une fréquence F comprise entre les limites suivantes :

$$p - t\sqrt{\frac{p(1-p)}{n}} < F < p + t\sqrt{\frac{p(1-p)}{n}}, \tag{11}$$

t étant lié à Φ_t par une relation fonctionnelle.

Par un passage tautologique, on tire de la relation (11)

$$F - t\sqrt{\frac{p(1-p)}{n}} < p < F + t\sqrt{\frac{p(1-p)}{n}} \tag{12}$$

qui n'est qu'une manière différente d'exprimer ladite relation (11).

L'inversion symétrique de la relation (11), obtenue en substituant à la variable indépendante p la fonction F et vice versa, donne

$$F - t\sqrt{\frac{F(1-F)}{n}} < p < F + t\sqrt{\frac{F(1-F)}{n}}. \tag{13}$$

L'inversion statistique symétrique est au contraire obtenue en substituant dans la relation (13), à la quantité individuelle connue p, la quantité statistique à atteindre P et à la quantité statistique à atteindre F la quantité individuelle connue f. Elle conduit, toujours avec la probabibité Φ_t, à l'inégalité

$$f - t\sqrt{\frac{f(1-f)}{n}} < P < f + t\sqrt{\frac{f(1-f)}{n}} \tag{14}$$

qui exprime précisément le théorème inverse de Bernoulli.

* * *

Il est bien connu que l'inversion statistique symétrique du théorème de Bernoulli n'est valable que dans certaines conditions. Laplace a démontré que cette inversion est valable dans le cas où toutes les probabilités de l'événement soient également probables *a priori*, c'est-à-dire avant qu'on n'effectue les n observations.

On parle, dans ce cas, d'inversion du théorème de Bernoulli sur la base du théorème de Bayes. Mais, dans de récents traités de Calcul des probabilités, une autre prétendue démonstration trouve place sous la dénomination d'inversion du théorème de Bernoulli sur la base du même théorème de Bernoulli.

Après avoir passé de la relation (11) à la (12), on remarque qu'en pratique la valeur de l'expression $\frac{p(1-p)}{n}$ ne peut pas différer sensiblement de celle de l'expression $\frac{f(1-f)}{n}$, où f est la fréquence observée de l'événement dans les n observations et n est un grand nombre. Par conséquent, en substituant dans la relation (12) la deuxième expression à la première, on obtient

$$F - t\sqrt{\frac{f(1-f)}{n}} < p < F + t\sqrt{\frac{f(1-f)}{n}}. \tag{15}$$

On prétend avoir démontré de la sorte le théorème inverse de Bernoulli. Il n'en est rien. La relation (15) ne doit pas être confondue avec (14). (15) ne fournit pas l'expression du théorème inverse de Bernoulli, mais rien de plus qu'une expression approchée de la relation (12), c'est-à-dire du théorème direct de Bernoulli. Il serait très commode d'inverser une démonstration en substituant une grandeur exacte par une grandeur approchée !

La confusion provient du fait que l'on n'a pas distingué par des notations appropriées une valeur lorsqu'elle joue le rôle de variable indépendante de la même valeur lorsqu'elle est une fonction. C'est un des cas dans lesquels le manque de symboles appropriés a conduit à des conclusions erronées [28].

Il faut dire que cette prétendue démonstration n'est pas du tout nouvelle ; on la trouve déjà chez Laplace, chez Poisson, chez Cournot; mais on l'avait, paraît-il, mise de côté. Elle a été remise en honneur par le traité de Bachelier (1912) et accueillie dans les dernières éditions de celui, qui est si justement renommé, de Czuber.

En Italie elle a été acceptée par plusieurs mathématiciens, même de haute valeur. Castelnuovo l'acceptait dans la substance, tout en qualifiant la fonction Φ_t du mot « attendibilità » (qu'en français, je crois, on pourrait rendre par « confiance ») plutôt que de celui de « probabilité »; mais l'expédient ne me paraît pas heureux, car, ou bien par le mot « attendibilità » on entend, selon sa signification courante, quelque chose d'équivalent à « probabilité », et la démonstration ne peut pas être admise, ou bien on lui donne un sens différent, et alors il faudrait le préciser pour savoir quel est le sens que l'on voudrait donner à l'énoncé du théorème inverse.

*
* *

Une autre forme d'inversion statistique — qui est à la mode — n'est pas neuve non plus et en général sa signification est mal interprétée. Dans un cas particulier, elle date tout au moins de Bessel et de Gauss.

Si v_c est la variance d'un collectif de m éléments, la relation

$$V_e = \frac{n-1}{n}\,\frac{m}{m-1}\,v_c \tag{16}$$

nous donne la valeur probable ou moyenne de la variance d'un échantillon de n éléments.

---

[28] *Voir* la communication citée à la note [25].

Si m est un grand nombre ou si l'échantillon est formé avec répétition, on peut écrire [29]

$$V_e = \frac{n-1}{n} v_c. \tag{16 bis}$$

C'est là une relation descriptive qui est valable pour les valeurs individuelles de v_c. L'inversion analytique de la relation (16 *bis*) qui conduit à

$$v_c = \frac{n}{n-1} \overline{V}_e \tag{17}$$

est parfaitement justifiée; c'est-à-dire que, si nous connaissons la variance de tous les échantillons possibles, nous pouvons tirer de leur moyenne la valeur exacte de la variance du collectif, et par conséquent, si nous connaissons les variances d'un grand nombre d'échantillons nous pouvons tirer de leur moyenne une valeur approchée de la variance du collectif.

Mais l'inversion statistique de la relation (16), qui est obtenue en appliquant à la relation (17) la rotation statistique et conduit à l'expression

$$\bar{V}_c = \frac{n}{n-1} v_e, \tag{18}$$

n'est pas autorisée, c'est-à-dire si nous connaissons la valeur de la variance d'un échantillon nous ne pouvons pas en tirer la variance probable du collectif.

La raison n'en est pas parce que nous ne connaissons que la valeur d'une seule variance, mais parce que le coefficient de correction $\frac{n}{n-1}$ s'applique

---

(29) Pour la démonstration des relations (16) et (16 *bis*), on peut voir notre mémoire *Variabilità e Mutabilità* dans les *Studi economico-giuridici publicati per cura della R. Università di Cagliari*, Anno III, Parte 2°, 1912; reproduite avec quelques additions dans *Memorie di Metodologia Statistica*, Milano, Giuffrè, 1935, p. 228-234. Des démonstrations de la relation (16 *bis*) sont aussi données dans le traité de J. BERTRAND, *Calcul des probabilités*, Paris, Gauthier-Villars, 1889, p. 202-204 et dans des traités modernes (par exemple, en G. CASTELNUOVO, *Calcolo delle probabilità*, vol. II, 2e éd., Bologna, Zanichelli, 1928, p. 23-25; en C. G. LAMBE, *Elements of Statistics*, Longmans, Green and Co., London, 1952, p. 71-72; J. F. KENNEY and E. S. KEEPING, *Mathematics of Statistics*, part two, 2e éd., Von Nostrand, New York, 1951, p. 160-168 et, avec une discussion plus approfondie, dans l'ouvrage de E. CZUBER : *Theorie der Beobachtungsfehler*, Leipzig, Teubner, 1891, p. 150-156).

à la moyenne des valeurs de v_e et non à une valeur individuelle. Par conséquent, si nous considérons un grand nombre, même un très grand nombre d'échantillons, ayant tous la même variance v_e ,nous ne pouvons pas dire qu'en appliquant à leur variance le coefficient de correction $\frac{n-1}{n}$ nous obtiendrons la valeur probable de la variance du collectif. En d'autres termes, la relation (18) n'est pas *l'inverse*, mais seulement la *subinverse* de la (16).

Bessel et Gauss ont recommandé la relation (18), c'est-à-dire la division de la somme des carrés de n écarts par $n-1$ plutôt que par n. Généralement on le fait; moi aussi je le fais — faute de mieux — et je crois bien de le faire, mais je me rends bien compte de ce que je fais. Dans l'échantillon que je considère ou même dans un grand nombre d'échantillons ayant la même variance, je sais que je m'expose non seulement à une erreur accidentelle du fait que je considère un seul échantillon ou un nombre limité — bien qu'il puisse être grand — d'échantillons, mais aussi à une erreur systématique. Pourtant, si je considère un grand nombre d'échantillons avec des variances différentes — échantillons tirés au hasard de tous les échantillons possibles — les erreurs systématiques dont sont affectées les diverses valeurs des variances tendent à se compenser. Je le fais faute de mieux, mais je me garde de croire que je fais une *inversion;* c'est beaucoup moins qu'une inversion, car dans l'opération que je fais — et que j'appelle *subinversion* — je renonce à utiliser la connaissance de la valeur de l'échantillon observé.

En 1930 R. A. Fisher a vanté d'une façon générale ce procédé, que j'appelle *subinversion*, pour passer de la valeur d'une constante quelconque dans l'échantillon à l'intervalle dans lequel la constante de la masse est contenue avec une certaine probabilité ([30]). Je n'ai pas bien compris si, soit Fisher, soit les autres auteurs (Neyman, E. S. Pearson, Clopper) qui s'en sont occupés à la suite ([31]) se rendent compte que le procédé est autorisé seulement pour l'ensemble des valeurs des constantes des échantillons. En tous cas, il paraît bien qu'ils ne s'aperçoivent pas qu'il y a là une cause d'infériorité de ce procédé vis-à-vis du procédé

([30]) *Voir* R. A. Fischer, *Inverse probability* dans *Proc. Cambridge Phil. Soc.*, vol. 26, Partie 4.

([31]) Ces auteurs et bien d'autres sont cités et leurs thèses discutées dans la communication : *I testi di significatività* citée dans la note ([30]).

de l'inversion. Fisher, au contraire, en fait même un mérite, car, dit-il, ce procédé est indépendant de la valeur de la constante observée dans l'échantillon et par conséquent plus général.

Ce serait comme si, en devant faire une prévision sur ce qui reste à vivre à une certaine personne, et ne connaissant pas la durée moyenne de la vie des personnes de son âge, on lui attribuait une durée de vie égale à la durée moyenne de la vie d'une personne d'un âge quelconque. Il est vrai que le procédé serait indépendant de l'âge de la personne et il est vrai aussi que, faute de mieux, on limiterait de la sorte l'erreur possible, mais évidemment la prévision serait foncièrement différente de celle basée sur la durée moyenne de vie des personnes ayant le même âge que la personne considérée.

Avant Fisher, en 1923, le procédé avait été adopté par Millot (32) et après, en 1933, par O. Anderson (33) dans l'inversion du théorème de Bernoulli, en procédant à l'inversion analytique accompagnée de la rotation statistique. La formule à laquelle on parvient avec l'inversion statistique qui en résulte est dans ce cas très différente de celle qui résulte de l'inversion statistique symétrique et que l'on désigne ordinairement sous le nom de théorème inverse de Bernoulli.

Au contraire, lorsqu'il s'agit de grandeurs extensives, les deux procédés de l'inversion statistique ordinaire et de l'inversion statistique symétrique, conduisent à la même formule.

L'essentiel qu'il importe de faire ressortir est que, soit la formule de Millot-Anderson dans le cas d'une fréquence relative, soit la formule obtenue par le procédé analogue lorsqu'il s'agit d'une grandeur extensive, ne sont pas applicables à des valeurs individuelles des grandeurs observées, mais seulement (dans l'hypothèse d'une relation monotone) à l'ensemble de ces valeurs.

Pour qu'elles soient applicables aux valeurs individuelles — pour qu'il s'agisse d'une véritable inversion et non d'une subinversion — d'autres hypothèses sont nécessaires que nous avons déjà mentionnées.

---

(32) Stanislas Millot, *Sur la probabilité à posteriori* dans *Comptes Rendus de l'Académie des Sciences*, Paris, t. 176, 1923.

(33) *Voir* l'article *Zum Problem der Wahrscheinlichkeit a posteriori in der Statistik* dans *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, t. 83, n° 6, 1947.

*
* *

La base des procédés classiques est la nécessité d'avoir ou de supposer d'avoir des connaissances *a priori* sur l'événement ou le caractère pour déterminer sa probabilité *a posteriori* après avoir observé sa fréquence ou telle autre valeur statistique dans un certain nombre de cas.

Tous les procédés que j'ai considérés dans ce chapitre visent à se passer de cette exigence, qui est assez incommode, car, dans la pratique, il n'est pas toujours possible d'avoir ces connaissances *a priori*.

C'est pour cela que je les ai appelés procédés non classiques.

Une autre façon de se soustraire à la tyrannie des connaissances *a priori* est de démontrer qu'elles n'ont pas d'importance.

Plusieurs auteurs modernes, et non des moindres (je citerai entre tous Edgeworth, Bowley et von Mises) ([34]), ont tenté cette évasion visant à établir que, lorsque le nombre des cas observés est très grand, l'inversion statistique basée sur l'hypothèse de l'égalité de toutes les probabilités *a priori* donne des résultats satisfaisants, même si cette hypothèse est loin d'être réalisée.

On adopte le procédé asymptotique, en démontrant que, si l'on augmente le nombre des observations, toutes autres grandeurs demeurant invariées, l'influence de la distribution *a priori* des probabilités diminue et tend à devenir nulle lorsque le nombre des observations tend à l'infini. Cela est incontestable.

On en conclut que, lorsque le nombre des observations est très grand, cette influence peut être négligée. Cela est plus contestable. En effet, il faudrait que, non seulement le nombre des observations fût très grand, mais aussi que, comparées à ce nombre, toutes les autres grandeurs qui figurent dans les formules fussent très petites. C'est là une condition dont généralement on ne se soucie pas, mais qui est nécessaire. Si l'on n'y fait pas attention, on risque d'arriver à des conclusions contradictoires.

En effet, il peut très bien arriver qu'en faisant tendre vers l'infini une grandeur α (toutes les autres β, γ, ..., ne variant pas), on arrive à une certaine conclusion. D'autre part, en faisant tendre vers l'infini une autre grandeur, par exemple β, toutes les autres — α comprise — res-

([34]) *Voir*, à ce sujet aussi, la communication *I testi di significatività*, la note ([50]).

tant invariables, on peut bien arriver à une conclusion tout à fait différente.

Si l'on applique l'une ou l'autre conclusion à des cas concrets, sans faire attention à la condition que toutes les autres grandeurs doivent être négligeables, on arrivera à des résultats contradictoires lorsque les deux grandeurs α et β présentent des valeurs élevées.

Du point de vue pratique, il y a lieu aussi de remarquer que, lorsque le nombre des observations est infiniment grand, la valeur que présentent les cas observés coïncide exactement avec sa valeur théorique et, lorsque le nombre des observations est très grand, on sait *a priori* que l'erreur est négligeable de façon que, dans ce cas, la mesure de sa grandeur n'a pas d'importance. L'important est de mesurer la probabilité des erreurs dans les échantillons qui comprennent toujours un nombre de cas plus ou moins limité.

Revenons à la formule

$$P = \frac{m+k+1}{n+k+h+2} \tag{19}$$

qui, ainsi que nous l'avons vu dans le chapitre précédent, exprime la probabilité *a posteriori* d'un événement qui a été observé m fois en n observations, si les probabilités *a priori* suivent la distribution β, dont les constantes k et h constituent les caractéristiques.

En maintenant invariables k, h et $\frac{m}{n}$ et en faisant augmenter indéfiniment n, la formule tend vers la valeur $\frac{m}{n}$; mais, en ne variant pas m, n et $\frac{k}{h}$ et en faisant augmenter indéfiniment k, la formule tend vers la valeur $\frac{k}{k+h}$. Il serait erroné de dire que, lorsque le nombre n des observations est très grand, on peut considérer que $P = \frac{m}{n}$, ainsi qu'il serait erroné de dire que, lorsque h est très grand, on peut considérer que $P = \frac{k}{k+h}$. En réalité, pour pouvoir admettre que $P = \frac{k}{k+h}$, il faut que, non seulement h, et par conséquent, k soient très grands, mais aussi que n, et par conséquent m, soient négligeables en comparaison. De même, pour pouvoir considérer que $P = \frac{m}{n}$, il ne suffit pas que n et par conséquent m, soient très grands. Il faut aussi que k, et par conséquent h, soient négligeables en comparaison. Si n et k sont tous les deux très grands, ni la première conclusion, ni l'autre ne peuvent être acceptées. Et, puisque les valeurs de k et h sont indépendantes des valeurs de m et n, la formule ci-dessus montre bien que la prétention, suivant laquelle l'influence de la distribution des probabilités *a priori* est négligeable lorsque le nombre n des observations est grand, est vaine.

*
* *

Lorsqu'on procède à l'inversion du théorème de Bernoulli en faisant l'hypothèse que les probabilités *a priori* suivent la distribution β, on parvient à une probabilité Φ_t pour que la probabilité P de l'événement soit comprise dans les limites ci-dessous

$$s - t\sqrt{\frac{s(1-s)}{n}} < P < s + t\sqrt{\frac{s(1-s)}{n}}, \tag{20}$$

où

$$s = \frac{m+k}{n+k+h}.$$

C'est là un résultat plus général que celui de la formule (14) qui intervient dans le théorème inverse de Bernoulli.

La formule (19) se réduit à la formule (14) dans le cas particulier $k = 1$, $h = 1$.

J'appelle *superinversion* ce procédé, par lequel on arrive à une relation inverse plus générale que la relation directe, comme j'ai appelé *subinversion* le procédé par lequel on arrive à une relation inverse moins générale que la relation directe.

*
* *

Avant de conclure, je désire examiner un dernier argument, que j'appellerai la dernière tranchée des équiprobabilistes.

Nous ne prétendons pas — un distingué collègue me l'a écrit une fois — que la conduite humaine soit dirigée exclusivement par nos formules et que l'on ne doive pas raisonnablement tenir compte aussi d'autres données de l'expérience.

Je veux bien me mettre moi-même à ce point de vue et le défendre dans la mesure du possible, en rappelant ce que j'ai dit sur les connaissances supplémentaires.

Nous avons parfois des connaissances — que j'ai appelées *connaissances supplémentaires*, qui ne peuvent pas être exprimées en chiffres et dont par conséquent on ne peut tenir compte dans les formules, mais dont on peut tenir compte dans une certaine mesure dans l'application de celles-ci. Une de ces connaissances peut concerner précisément les diverses probabilités des causes ou hypothèses *a priori*. Nous savons qu'en réalité elles sont différentes, mais nous ne savons pas en exprimer numériquement les différences; nous supposons alors qu'elles soient égales et nous déterminons dans cette hypothèse la probabilité de l'événement qui nous intéresse, sauf à tenir compte, après, de l'effet desdites différences. Nous pouvons dire, en d'autres mots, que nous déterminons de la sorte la probabilité d'un événement qui n'est pas exactement celui qui nous intéresse, mais qui lui est analogue, sauf à examiner, après, les variations que nos conclusions doivent subir par l'effet de ce manque d'identité.

D'un point de vue pratique, il y a lieu pourtant de remarquer que cet examen *après formule* est fait bien rarement.

D'un point de vue théorique, il y a lieu d'ajouter que, même lorsque cet examen a été fait, il ne peut pas toujours autoriser une correction — fût-elle seulement qualitative — de nos conclusions.

Pour permettre de tirer profit des connaissances supplémentaires, il est nécessaire, non seulement de connaître le fait que le phénomène qui intéresse et que l'on voudrait mais que l'on ne peut pas exprimer numériquement, diffère du phénomène analogue qu'on est en condition d'exprimer par des nombres, mais aussi qu'il en diffère *dans un certain sens déterminé*.

Or, dans certains cas, il est possible de préciser le sens dans lequel l'hypothèse de l'équiprobabilité des causes s'écarte de la réalité et alors il est possible de corriger la conclusion à laquelle cette hypothèse nous amènerait.

J'ai eu l'occasion d'adopter ce procédé lorsque — il y a presque 50 ans — je me suis occupé des prétendues exceptions à la régularité de l'excédent des garçons dans les naissances humaines ([35]).

([35]) Voir *Il sesso dal punto di vista statistico*, Sandron, Palermo, 1908 (en vente à la Faculté de Statistique de l'Université de Rome), p. 71-72.

Dans un traité de démographie et de politique de la population, A. von Fircks remarquait que, dans les pays les plus septentrionaux, les naissances semblent se partager dans des proportions presque égales entre les deux sexes, et que parfois même ce sont les filles qui présentent un excédent. Pour le démontrer il rappelait qu'en 1895 dans le Groenland, il y avait eu 222 nouveau-nés de sexe masculin et 227 de sexe féminin.

Or, je faisais ressortir le peu de vraisemblance de la conclusion qu'il voulait en tirer. En admettant qu'*a priori* tous les rapports des sexes dans les naissances du Groenland avaient été également possibles, le Calcul des probabilités permet de calculer, sur la base des données relatées par von Fircks, la probabilité pour qu'au Groenland il y ait plus de chances d'enfanter un garçon qu'une fille. Elle est de 37 %.

« En réalité, pourtant — remarquais-je — on ne peut pas admettre qu'*a priori* tous les rapports des sexes dans les naissances soient pour le Groenland également possibles : le fait qu'en tous les autres pays on ait toujours observé plus de naissances de garçons que de filles doit faire considérer *a priori* que très vraisemblablement au Groenland il arrive aussi la même chose. Par conséquent, la probabilité pour qu'au Groenland les chances de la naissance d'un garçon l'emportent sur celles de la naissance d'une fille aurait dû être évaluée à bien plus que 37 % par celui qui, comme von Fircks, connaissait seulement qu'en 1895 il y avait eu 222 naissances de garçons et 227 de filles. En effet, les naissances au Groenland dans les 20 années 1884-1903, montrèrent la proportion, tout à fait normale, de 106 garçons pour 100 filles ».

Il est à remarquer à ce sujet que le théorème inverse de Bernoulli donne toujours une valeur supérieure pour la probabilité que la différence entre deux rapports de fréquence se réalise dans le sens dans lequel elle a été observée par rapport a la probabilité pour qu'elle se réalise dans le sens contraire.

Cela revient à dire qu'il donne toujours une probabilité plus élevée, pour que la différence soit significative plutôt qu'elle ne le soit pas. Cela montre bien la faiblesse du procédé basé sur l'équiprobabilité des causes et la nécessité de tenir compte, au contraire, des différences de leurs probabilités.

Il est à peine nécessaire de faire remarquer que la conclusion qu'on est autorisé de tirer sur la base de certaines observations peut être tout à fait différente selon que les probabilités *a priori* se condensent sur des valeurs plus élevées ou au contraire moins élevées que la fréquence observée. Si nos connaissances supplémentaires nous avaient dit que les rapports des sexes dans les naissances donnaient en général un excédent de filles, les déductions à tirer des données de von Fircks auraient été évidemment bien différentes.

Lorsque nos connaissances supplémentaires nous informent donc que les probabilités *a priori* ne sont pas équiprobables sans nous préciser dans quel sens elles s'écartent de l'équiprobabilité, nous ne sommes pas en mesure de corriger les conclusions tirées du théorème inverse de Bernoulli.

Quatre établissements ont présenté des parties de produits dont on a extrait des échantillons de 100 éléments. Dans chaque échantillon, 10 % des éléments se sont avérés défectueux. On demande les probabilités pour que le pourcentage d'éléments défectueux dans les quatre parties entières dépasse la limite tolérée de 8 %. Sur la base du théorème inverse de Bernoulli, la probabilité serait, pour les quatre parties, 0,144. C'est l'exemple que nous avons mentionné au chapitre précédent. En admettant, au contraire, que les probabilités *a priori* suivent la distribution β, la prévision à faire pourrait être tout à fait différente dans les quatre parties selon les échantillons — dont notre formule permet de tenir compte — présentés par les établissements en question dans des occasions précédentes. Au lieu de 0,144 la probabilité monterait pour un établissement à 0,878 — c'est-à-dire six fois autant — et descendrait pour une autre à 0,0004, c'est-à-dire 360 fois moins ([36]).

Cela montre bien à quelles erreurs peut donner lieu, dans la pratique, l'application du théorème inverse de Bernoulli sur la base de l'hypothèse de l'équiprobabilité des causes et l'utilité par conséquent d'avoir recours — toutes les fois que cela est possible — au procédé de la superinversion sur la base de la distribution β des probabilités *a priori*.

Conclusions. — Nous pouvons maintenant récapituler les lignes essentielles de ce que nous avons dit sur l'inversion statistique.

Il y a une distinction fondamentale à faire entre méthodes classiques et méthodes non classiques ou modernes.

([36]) *Voir* la communication *Sulla probabilità inversa nel caso di grandezze intensive e in particolare sulle sue applicazioni a collaudi per masse a mezzo di campioni* (en collaboration avec G. Livada), présentée à la VII^e Réunion scientifique (27-30 juin 1943) de la Société italienne de Statistique et publiée dans les *Actes* de la dite Société, Rome, 1945.

La caractéristique des méthodes classiques est de tenir compte des connaissances *a priori*, c'est-à-dire antérieures aux observations que l'on a faites en dernier lieu sur l'événement considéré.

Les connaissances *a priori* représentent le fruit des observations antérieures et de nos réflexions. On ne saurait nier leur importance. Souvent elles sont même plus importantes que les observations actuelles. C'est pourquoi les procédés classiques représentent la grande route à suivre dans l'inversion statistique. Même, lorsque leurs formules ne peuvent être appliquées, faute de renseignements nécessaires, elles montrent le chemin à suivre et dans certains cas elles peuvent indiquer dans quelle direction on s'écarte de la vérité en ne les suivant pas.

Dans la première partie de ce chapitre, nous avons d'abord examiné les bases philosophiques des procédés d'inversion statistique et nous avons ensuite exposé les méthodes introduites par Laplace et par Gauss, en proposant des généralisations théoriques qui permettent des applications plus étendues ou nouvelles, soit sur la base d'hypothèses plus générales et plus proches de la réalité, soit en faisant entrer dans les schémas traditionnels d'autres données de l'observation.

Malgré ces généralisations, il faut reconnaître que l'on ne dispose pas toujours de renseignements *a priori* suffisants pour l'application des procédés classiques, ce qui justifie les efforts, qui caractérisent les recherches modernes dans le domaine de l'inversion statistique, visant à se libérer de la nécessité desdites connaissances *a priori*.

Il y a lieu aussi d'ajouter que parfois, surtout dans les sciences expérimentales, des problèmes nouveaux se présentent, pour la résolution desquels on ne peut pas profiter des connaissances *a priori*.

Les procédés non classiques, que nous avons examinés dans la deuxième partie de ce chapitre, se proposent précisément de parvenir à l'inversion statistique, en se passant des connaissances *a priori* nécessaires à l'application des procédés classiques. Certains d'entre eux ont leurs racines dans les classiques, mais ce n'est que récemment qu'ils ont pris pied et ont été développés.

Nous avons analysé le mécanisme logique des divers procédés en question, en mettant en évidence les méprises, facilitées parfois par des notations inadéquates, sur lesquelles certains d'entre eux se basent, ainsi que les méprises sur lesquelles se basent les applications de quelque autre procédé.

J'ai en particulier distingué les procédés de l'inversion statistique proprement dite, du procédé que j'ai appelé de *subinversion*, lequel conduit à des relations inverses qui ne sont pas valables pour les valeurs individuelles, mais seulement pour l'ensemble des valeurs observables, ainsi que du procédé, que j'ai appelé de *superinversion*, qui conduit à des relations inverses ayant une valeur plus générale que les relations directes d'où elles sont déduites.

Le premier est un procédé qui n'est pas tout à fait dépourvu d'utilité, mais qui a un rendement bien moindre que les procédés classiques d'inversion que, par conséquent, il ne peut pas remplacer.

Le deuxième est un procédé qui a une utilité pratique remarquable, car il permet d'exprimer numériquement des connaissances sur la distribution de probabilités *a priori* qui peuvent être décisives pour les conclusions à tirer des observations que l'on vient de faire.

L'ignorance des procédés classiques de la part de bien des jeunes statisticiens et la persistance avec laquelle ils s'attachent à des procédés modernes imparfaits ou moins efficaces, sans prêter l'oreille aux critiques qui leur ont été adressées de divers côtés, a probablement une base psychologique dans la tendance de la jeunesse actuelle à agir à tout prix et en se passant de l'expérience antérieure, comme si le monde venait de commencer [37].

J'ai été frappé par la réponse caractéristique qu'un statisticien américain assez jeune — les Américains d'ailleurs sont toujours jeunes — et certainement distingué, faisait en 1947, au Congrès de Washington à la critique que je venais d'adresser aux méthodes modernes d'inversion statistique. — Nous sommes des disciples de William James — disait-il — nous sommes des pragmatistes. Notre programme est l'action. Les méthodes statistiques doivent être dirigées vers l'action. C'est là leur avenir, c'est là leur progrès.

Or, il n'y a pas de doute qu'en face des mêmes faits le comportement de personnes différemment douées d'initiative, des jeunes et des vieux, est différent. Mais cela dépend des divers jugements de valeurs que les uns et les autres font des faits en question; cela ne dépend pas de la circonstance que les procédés logiques ou les formules du Calcul des

---

([37]) *Voir*, à ce sujet, le discours d'ouverture de la XIV^e Réunion scientifique de la Société italienne de Statistique (1954) ayant pour titre *Cause e significato della recente evoluzione della metodologia statistica*, et publié dans les Actes de ladite Réunion (1955).

probabilités seraient différents pour ceux-ci et pour ceux-là. Il n'y a pas deux logiques ou deux mathématiques : une pour les pragmatistes et l'autre pour les non-pragmatistes.

Il faut reconnaître de même l'impossibilité de prétendre que l'homme agit toujours suivant les préceptes de la science. Au contraire il peut bien être nécessaire de procéder par tentatives et, d'autre part, il est évident que la disposition à faire face au risque que ces tentatives comportent, dépend essentiellement du tempérament personnel. Mais cela ne peut pas changer les préceptes de la science, rendre rigoureux des procédés qui ne le sont pas ou changer les conclusions que des procédés rigoureux autorisent. La mission de l'homme d'études est non seulement de proposer des procédés bien fondés, mais aussi de mettre en garde contre les procédés qui ne le sont pas, en traçant de la sorte les limites de la science : il ne peut pas prétendre qu'au-delà de ces limites la vie cesse; mais il peut prétendre que l'on ne vend pas sous l'enseigne de la science des produits qui ne sont pas les siens ([38]).

CHAPITRE V.

GÉNÉRALISATION DE LA THÉORIE DE LA DISPERSION.

1. La théorie de la dispersion fournit les critériums pour déterminer si les différences que les intensités d'un phénomène donné présentent dans divers espaces de temps, ou dans diverses circonscriptions territoriales, ou dans des catégories qui se distinguent par les modalités d'autres phénomènes, ont caractère accidentel ou systématique. Dans ce but on détermine ordinairement un indice de variabilité des intensités qui se présentent dans les diverses catégories et on le compare à sa valeur théorique telle qu'elle serait par le seul effet du hasard.

([38]) Pour les sujets traités dans ce chapitre IV.2, on peut consulter, outre les travaux cités dans les notes, la communication *I testi di significatività*, présentée en juin 1943 à la VIIe Réunion scientifique de la Société italienne de Statistique et publiée dans les *Actes* de ladite Réunion (1945); l'autre communication *On Statistical Relations and their Inversions*, présentée à la 25e Session de l'Institut International de Statistique (Washington, 1947) et publiée en anglais dans la *Revue de l'Institut International de Statistique*, 1947, et en allemand dans la *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, Bd 83, 1947, H. 6; *voir* aussi la note dans la même Revue, Bd 84, 1948, H. 4, et l'article *Passaggio dall'indice di variabilità di un campione all'indice di variabilità della massa*, dans *Metron*, 1951, nos 3-4.

Les premières applications de cette théorie — faites par Dormoy (1874), qui employait comme indice de variabilité l'écart simple moyen, et par Lexis (1876), qui employait l'écart quadratique moyen — se rapportaient à des fréquences relatives (par exemple : coefficients de mortalité, rapports des naissances de garçons au total des naissances etc.).

Von Bortkiewicz étendit ces applications aux moyennes de grandeurs mesurables, telles que la stature, le revenu, etc. On peut examiner, par exemple, si c'est accidentellement ou systématiquement que la stature moyenne des conscrits ou le revenu moyen des contribuables varie d'une année à l'autre ou d'une circonscription à une autre.

2. Le schéma adopté pour déterminer la valeur théorique de l'indice de variabilité choisi était le même dans les deux cas.

En indiquant par $1, 2, 3, \ldots, s$ les sections (espaces de temps, circonscriptions territoriales, catégories qualitatives) dans lesquelles on divise le champ d'observation; par $n_1, n_2, \ldots, n_s$ et en général par n_k (où $k = 1, 2, \ldots, s$) le nombre des observations faites dans chacune d'elles, par $N = \sum_{k=1}^{s} n_k$ leur total; par $1, 2, \ldots, t$ les modalités du caractère (sexe, survivance, stature, etc.); par $M_1, M_2, \ldots, M_t$ le nombre des fois où elles se présentent et par $N = M_1 + M_2 + \ldots + M_t$ leur total correspondant au nombre total des observations, on suppose que dans une urne aient été mises N boules de t couleurs, correspondant en nombre aux t modalités possibles et l'on procède ensuite à l'extraction au hasard et sans répétition, d'abord de n_1 boules que l'on attribue à la section 1, puis de n_2 boules que l'on attribue à la section 2 et ainsi de suite, jusqu'à n_s boules que l'on attribue à la section s.

Ce *schéma* peut être appelé des *tirages au sort des modalités appartenant à chaque section* ou, plus brièvement, *schéma des tirages au sort*.

3. On peut avoir recours aussi à un autre schéma. Supposons que dans une urne on ait encore mis N boules, mais de s couleurs différentes correspondant aux sections, et en nombre égal à celui des observations faites dans les sections respectives, et que l'on tire de l'urne, au sort et sans répétition, d'abord M_1 boules que l'on attribue à la

modalité 1, puis M_2 boules que l'on attribue à la modalité 2, et ainsi de suite jusqu'à M_t boules que l'on attribue à la modalité t.

Ce schéma peut être appelé *schéma des tirages au sort des sections entre lesquelles doivent être réparties au hasard les modalités* ou, plus brièvement, *schéma des répartitions au hasard*.

Les résultats que l'on obtient par les deux schémas sont identiques.

4. Si l'on indique par a_{1k}, a_{2k}, ..., a_{nk} et en général par a_{ik} (où $i = 1, 2, \ldots, n$) les valeurs observées dans les observations relatives à la $k^{\text{ième}}$ section, par

$$A_k = \frac{1}{n_k}\sum_{i=1}^{n_k} = a_{ik}$$

les moyennes des n_k valeurs observées dans la $k^{\text{ième}}$ section, et donc par

$$A = \frac{1}{N}\sum_{k=1}^{s} n_k A_k = \frac{\sum_{k=1}^{s}\sum_{i=1}^{n_k} a_{ik}}{N}.$$

la moyenne des N valeurs observées dans toutes les k sections, l'écart quadratique moyen effectif des s valeurs de A_k sera

$$\sqrt{\frac{1}{N}\sum_{k=1}^{s} n_k (A_k - A)^2}$$

et l'écart quadratique moyen qu'il y aurait lieu de prévoir par l'effet du hasard serait

$$\sqrt{\frac{s-1}{N-1}\,\frac{\sum_{k=1}^{s}\sum_{i=1}^{t}(a_{ik} - A)^2}{N}},$$

dont on tire l'indice de dispersion

$$(1) \qquad Q_1 = \sqrt{\frac{(N-1)\sum_{k=1}^{s} n_k (A_k - A)^2}{(s-1)\sum_{k=1}^{s}\sum_{i=1}^{t}(a_{ik} - A)^2}}.$$

On dit que Q_1 indique une *dispersion subnormale*, *normale*, *supernormale* selon qu'on a

$$Q_1 \lesseqgtr 1.$$

Dans le cas où il s'agit de grandeurs intensives, soit de caractères énumérables pour lesquels il n'y a que seulement deux intensités 1 ou 0, les moyennes A_k et A sont constituées par des rapports de fréquence, que nous désignerons par $\frac{m_k}{n_k}$, $\frac{M}{N}$, et la formule (1) peut être transformée ainsi

$$(2) \qquad Q_2 = \sqrt{\frac{(N-1)\sum_{k=1}^{s} n_k\left(\frac{m_k}{n_k} - \frac{M}{N}\right)^2}{(s-1)M\left(1-\frac{M}{N}\right)}}.$$

Il y a une différence importante entre les formules (1) et (2) au point de vue de leur applicabilité. Pour appliquer la formule (2), il suffit, en effet, de connaître (outre les valeurs de m_k et de N) la fréquence moyenne $\frac{M}{N}$; pour appliquer la formule (1), il faut, par contre, connaître (outre les valeurs de A_k et de n_k) non seulement l'intensité moyenne A, mais aussi les écarts $a_{ik} - A$ que présentent les intensités singulières à partir de cette moyenne. Cela dépend du fait que, pour les fréquences relatives, une fois que l'on connaît l'intensité moyenne $\frac{M}{N}$, on en déduit immédiatement l'écart quadratique moyen, tandis que cela n'est pas vrai pour les autres moyennes.

5. Si les deux schémas considérés ci-dessus des tirages au sort et des répartitions au hasard conduisent au même résultat, le deuxième offre cependant un intérêt particulier, parce que, avec de légers ajustements, il peut être étendu au cas des phénomènes susceptibles de se répéter, auxquels le premier n'est, par contre, pas applicable.

Les deux schémas peuvent être appliqués, par exemple, pour juger si les décès se distribuent au hasard, dans un espace de temps donné, entre les habitants des diverses circonscriptions territoriales ou, dans une circonscription territoriale donnée, entre les espaces de temps successifs, mais ils ne peuvent, par contre, être utilisés pour déterminer si, entre, ces habitants les rhumes se distribuent au hasard. Cela parce que, dans le premier cas, la somme des cas présentant les diverses modalités du phénomène est égale au total des individus, comme il est supposé dans les deux schémas ci-dessus considérés (par exemple, la somme des morts et des survivants est égale au nombre des habitants exposés à mourir, étant donné que tout individu ou meurt ou survit et naturellement ne peut mourir qu'une seule fois); tandis que, dans le second cas, un individu peut avoir été ou non enrhumé dans un espace de temps, mais peut aussi avoir été plusieurs fois enrhumé dans cet espace de temps, de sorte que les rhumes plus les individus qui n'ont jamais été enrhumés ne sont pas équivalents au nombre des habitants exposés aux rhumes, et par conséquent les formules ci-dessus ne sont pas applicables.

Le schéma des répartitions au hasard devient cependant applicable si, au lieu de considérer des tirages au sort sans répétition, nous considérons des tirages au sort avec répétition.

Supposons pour cela une urne contenant des boules en nombre égal à celui des N habitants du pays considéré, et de couleurs différentes selon les circonscriptions territoriales en lesquelles nous supposons que la population soit divisée. Nous nous demandons si les M rhumes qui se sont manifestés dans une année sont répartis accidentellement entre les habitants des circonscriptions. S'il s'agissait de décès au lieu de rhumes, nous pourrions, sur la base du schéma des répartitions au hasard ci-dessus considéré, tirer au sort de l'urne sans répétition M boules qui seraient de couleurs diverses correspondant aux circonscriptions où les décès auraient lieu par l'effet du hasard. Mais nous ne pouvons pas en faire autant pour les M rhumes qui pourront être, et seront vraisemblablement, plus nombreux que les habitants et donc que les boules contenues dans l'urne. Du moment que le rhume peut se reproduire sur un même individu, nous devons considérer le cas où la boule qui a déjà été extraite, et correspond au rhume d'un individu donné, peut être extraite de nouveau en correspondance à un rhume ultérieur de cet individu.

Si nous indiquons par $m_1, m_2, \ldots, m_s$ les rhumes des circonscriptions ayant respectivement $n_1, n_2, \ldots, n_s$ habitants, la distribution des rhumes sera accidentelle lorsque les rapports $\frac{m_1}{M}, \frac{m_2}{M}, \cdots, \frac{m_s}{M}$ des rhumes dans les diverses circonscriptions aux rhumes de tout le pays considéré s'écarteront par des différences accidentelles des rapports $\frac{n_1}{N}$, $\frac{n_2}{N}, \ldots, \frac{n_s}{N}$ des habitants des diverses circonscriptions aux habitants de tout le pays.

Au lieu de comparer, comme dans les schémas précédents, des rapports théoriques à des rapports effectifs de dérivation, nous sommes amenés ainsi à comparer des rapports théoriques à des rapports effectifs de composition.

Nous appellerons ce schéma : *schéma de répartition au hasard sur la base des rapports de composition* ou simplement *schéma des rapports de composition* par opposition aux deux précédents, que nous appellerons *schémas des tirages au sort* et, respectivement, des *répartitions au hasard sur la base des rapports de dérivation* ou simplement *schémas des rapports de dérivation.*

6. Selon le schéma des rapports de composition, le carré de l'écart effectif de la première circonscription sera $\left(\frac{m_1}{M} - \frac{n_1}{N}\right)^2$ et sa valeur théorique $\left(\frac{n_1}{N}\right)\left(1 - \frac{n_1}{N}\right)\left(\frac{1}{M}\right)$. Pour toutes les circonscriptions, l'indice de dispersion sera donc

$$(3) \qquad Q_3 = \sqrt{\frac{M\sum_{i=k}^{s}\left(\frac{m_k}{M} - \frac{n_k}{N}\right)^2}{\sum_{k=1}^{s}\frac{n_k}{N}\left(1 - \frac{n_k}{N}\right)}} = \sqrt{\frac{M\sum_{k=1}^{s}\left(\frac{m_k}{M} - \frac{n_k}{N}\right)^2}{1 - \frac{1}{N^2}\sum_{k=1}^{s} n_k^2}}.$$

Dans le cas où n_k est const. $= n = \frac{N}{s}$, en faisant $\frac{M}{s} = m$, l'expression (3) devient

$$(3.1) \qquad Q'_3 = \sqrt{\frac{\sum_{k=1}^{s}(m_k - m)^2}{(s-1)m}}$$

tandis que la formule (2) devient

$$(2.1) \qquad Q'_2 = \sqrt{\frac{N-1}{N}\,\frac{\sum_{k=1}^{s}(m_k - m)^2}{(s-1)m\left(1 - \frac{m}{n}\right)}}.$$

En faisant abstraction du facteur $\frac{N-1}{N}$, que l'on peut ordinairement négliger, il y a entre Q'_3 et Q'_2 le rapport très simple

$$Q'_3 = Q'_2 \sqrt{1 - \frac{m}{n}},$$

où n indique la limite maximum que m peut atteindre dans le cas de phénomènes reproductibles. Dans le cas où le phénomène est infiniment reproductible, $n = \infty$ et donc Q'_3 devient égal à Q'_2.

La formule (3.1) est également valable lorsqu'il s'agit de phénomènes très rares, de sorte que m peut être considéré négligeable vis-à-vis de n, quelle que soit la grandeur de celui-ci.

La diversité entre la formule (3.1) et la formule (2.1) est due au fait que, dans le schéma des rapports de composition, on admet que tous les M cas observés peuvent être attribués à un seul terme et, puisque le nombre M dépend du nombre des termes et rien n'empêche que l'on considère un nombre de termes de quelque grandeur que l'on veuille, cela revient à dire que l'on peut attribuer à un seul terme un nombre illimité de cas du phénomène.

Par contre, dans le schéma des rapports de dérivation on admet qu'il y a, pour chaque terme, un nombre limité de cas réalisés, dont le maximum est indiqué par le nombre n des cas possibles, d'entre lesquels on suppose tirer au sort les cas réalisés.

Or, il faut remarquer que, dans la théorie de la variabilité et de la concentration, l'hypothèse a déjà été prise en considération qu'il y ait une limite supérieure à la valeur des termes singuliers de la série, et il a été trouvé qu'il faut alors introduire dans les indices relatifs un coefficient de correction qui est précisément celui qu'il faut introduire dans la valeur de Q'_3 pour obtenir la valeur de Q'_2 ([39]).

7. Le schéma des rapports de composition étend notablement le champ d'application de la théorie de la dispersion. Sur la base des formules (1) et (2), cette théorie pouvait s'appliquer aux fréquences

---

([39]) *Voir* l'article : *Sul massimo degli indici di variabilità assoluta e sulle sue applicazioni agli indici di variabilità relativa ed al rapporto di concentrazione* dans *Metron*, t. 8, 1930, n° 3; reproduit dans *Memorie di Metodologia statistica*, Giuffrè, Milan, 1939. Une nouvelle édition de cet ouvrage, mise à jour par MM. E. Pizzetti et T. Salvemini est sous presse (Veschi, Rome, Viale dell' Università, 7).

relatives (rapports de natalité, de nuptialité, de mortalité, de masculinité etc.) et aux autres grandeurs relatives ayant au dénominateur des quantités énumérables, telles que les moyennes arithmétiques (par exemple, stature moyenne, poids moyen, revenu moyen, etc.) : autrement dit, aux grandeurs relatives énumérables ou mesurables rapportées à des grandeurs énumérables.

Le schéma des rapports de composition permet d'étendre les applications à toutes les grandeurs énumérables absolues, telles que maladies, incendies, tremblements de terre ou secousses de tremblement de terre, éruptions volcaniques, pluies, chutes de neige, averses de grêle, nombre de poissons pêchés, des têtes de gibier tués, de diamants découverts et ainsi de suite ([10]). Des applications que nous avons faites restent exclues les grandeurs mesurables absolues (telles que : distances parcourues, espaces de temps passés, quantité de pluie tombée, montant des profits réalisés, valeur des diamants découverts, etc.) ou rapportées à d'autres grandeurs mesurables (telles que : dommages causés par les incendies rapportés à la valeur des bâtiments exposés, dommages causés par la grêle rapportés à la valeur des récoltes pendantes, etc.).

L'impossibilité de ces applications dérive du fait que les formules ci-dessus amèneraient à des résultats différents selon l'unité de mesure adoptée.

Or, la théorie de la dispersion peut ausssi être étendue comme nous le verrons, à des grandeurs de ce type, mais pour cela elle a besoin d'une généralisation ultérieure, à laquelle nous allons procéder en introduisant le concept de *dispersion d'ordre supérieur*.

8. Dans les schémas de la dispersion dont il a été question jusqu'ici, une urne seulement était considérée : la dispersion ainsi mesurée peut être appelée *dispersion de premier ordre*. Par opposition, nous

---

([10]) Pour le schéma des rapports de composition et pour les applications qu'il rend possibles, *cf.* l'article : ***Estensione della teoria della dispersione e della connessione a serie di grandezze assolute*** (***Giornale dell' Istituto Italiano degli Attuari***, t. 15, 1952, n^{os} 1-2, § 1-4. Rome). Des applications étendues des schémas des rapports de composition et de dérivation sont données dans l'article *Généralisations et applications de la théorie de la dispersion* qui va paraître dans *Metron*, t. 18, n° 1-2, 1955. Elles concernent les oiseaux tués, les thons pêchés, les tremblements de terre, les cas d'incendie, les jours de pluie, de neige, de grêle et en général de précipitations atmosphériques, les cas de maladie.

appellerons *dispersion d'ordre supérieur* celle qui est mesurée sur la base de schémas impliquant le recours à plusieurs urnes.

Pour illustrer ce concept, nous prenons en considération les grandeurs mesurables dont nous avons donné des exemples ci-dessus : distances parcourues, espaces de temps passés, quantité d'eau tombée, valeur des diamants découverts, montant des profits réalisés, etc.

Retenons ainsi que, pour les fins de la statistique qui s'occupe de phénomènes collectifs, ce qui intéresse ce sont des collections ou masses de ces grandeurs, par exemple : distances parcourues non pas par un seul train, mais par tous les trains qui ont fonctionné sur une ou plusieurs lignes durant un certain espace de temps; durée non pas d'une pluie seulement, mais de toutes les pluies qui se sont produites au cours d'une année; d'une façon analogue, quantité de l'eau tombée au cours non pas d'une seule, mais de toutes les précipitations atmosphériques de l'année; valeur non pas d'un diamant seulement, mais de tous les diamants trouvés dans un certain territoire pendant un espace de temps donné; montant des profits réalisés non pas dans une seule entreprise, mais dans toute une masse d'entreprises. Nous nommerons en général *montant extensif* le total des grandeurs mesurables ou extensives comprises dans le phénomène collectif. Les montants extensifs que l'on trouve dans la statistique peuvent donc toujours être figurés comme les produits de deux grandeurs, variables l'une et l'autre : nombre des grandeurs qui composent le phénomène collectif et moyenne arithmétique des intensités de ces grandeurs. Le montant extensif peut être défini comme une *grandeur mesurable à composantes énumérables*.

9. Nous dirons qu'une certaine série de montants (par exemple : la série des montants des lots de la loterie sortis dans chaque année d'un siècle) varie accidentellement, c'est-à-dire qu'elle présente une *dispersion normale du deuxième ordre* lorsqu'il n'est pas possible d'apercevoir aucun facteur systématique dans les variations des grandeurs énumérables et mesurables dont résulte ce montant.

La dispersion sera dite du deuxième ordre parce que deux urnes interviennent dans le schéma des probabilités qui lui correspond.

Dans le cas des lots de la loterie, la première urne — l'urne des grandeurs énumérables — contiendra des boules de deux couleurs différentes, disons blanc et noir, les boules blanches correspondant aux billets gagnants, les noires, aux billets perdants et en nombre proportionnel à celui des billets gagnants et perdants dans les tirages effectués pendant la période considérée. La deuxième urne — l'urne des grandeurs mesurables — contiendra autant de boules qu'il y a de lots, chaque boule portant la mention du montant correspondant au lot respectif.

Si $n_1, n_2, n_3, \ldots, n_{100}$ sont les nombres des fois que l'on a joué pendant la 1^re^, la 2^e^, la 3^e^, ..., la 100^e^ année du siècle, nous extrairons avec répétition de la première urne successivement, $n_1, n_2, n_3, \ldots, n_{100}$ boules, en attribuant les boules blanches extraites (au nombre de $m_1, m_2, m_3, \ldots, m_{100}$) aux années respectives. Elles représentent le nombre des lots qui seraient sortis par l'effet du hasard dans lesdites années.

A chacun de ces lots nous associerons chaque fois une boule prise au hasard sans répétition dans la deuxième urne. Nous obtiendrons ainsi $m_1 + m_2 + \ldots + m_{100} = \mathrm{M}$ grandeurs qui donnent lieu aux 100 montants $\mathrm{T}_1 + \mathrm{T}_2 + \ldots + \mathrm{T}_{100} = \mathrm{T}$. La dispersion de la série de ces 100 montants constituera la dispersion théorique des montants des lots sortis dans les 100 années du siècle, laquelle, comparée à la dispersion effective, nous dira si celle-ci est normale, ou bien supérieure ou inférieure à la normale.

Dans ce cas, le nombre des lots annuels possibles est limité au nombre des billets joués. Le schéma peut donc être dit de la *dispersion du deuxième ordre sur la base des rapports de dérivation.*

Si, par contre, le nombre des cas réalisés dans chaque section n'avait pas une limite supérieure définie, il nous faudrait recourir au schéma de *la dispersion du deuxième ordre sur la base des rapports de composition.*

Ce serait le cas pour la quantité d'eau tombée au cours des précipitations atmosphériques des différentes années d'un siècle.

Il n'y a pas, en effet, de limite définie aux précipitations atmosphériques d'une année.

L'urne des grandeurs énumérables devrait, dans ce cas, contenir M boules (soit un nombre égal à celui des précipitations atmosphériques ayant eu lieu dans le siècle) marquées des numéros de 1 à 100 (soit le nombre des années du siècle) avec une fréquence proportionnelle à l'espace de temps correspondant à chaque année, c'est-à-dire au nombre des jours qu'elle contient (365 pour les années normales, 366 pour les années bissextiles). De cette urne on extraira, l'une après l'autre avec répétition, M boules, en les attribuant à l'année marquée sur chacune d'elles. A chaque boule ainsi extraite de l'urne des grandeurs énumérables sera associée une boule extraite sans répétition de l'urne des grandeurs mesurables qui contiendra M boules correspondant aux précipitations atmosphériques et portant la mention de la quantité d'eau tombée au cours de chacune d'elles. On obtiendra ainsi une série composée de 100 montants, dont la dispersion pourra être comparée à la dispersion effective des quantités de l'eau tombée au cours de chaque précipitation atmosphérique dans chaque année du siècle.

10. Considérons d'abord le cas plus simple où les s sections soient équivalentes, dans le sens qu'il y a la même probabilité pour qu'une grandeur entre dans l'une ou l'autre d'entre elles (comme il adviendrait dans l'exemple ci-dessus, si toutes les années avaient le même nombre de jours). La somme des carrés des écarts des montants des s sections considérées $T_1, T_2, \ldots, T_s$ (en général T_k) de leur moyenne $\frac{T}{s}$ où $T = \sum_{k=1}^{s} T_k$ sera

$$\sum_{k=1}^{s} T_k^2 - \frac{T^2}{s}.$$

La valeur théorique de cette expression, sur la base du schéma des rapports de composition, est [41]

$$\sum_{k=1}^{s} \sum_{i=1}^{m_k} a_{ik}^2 - \frac{T^2}{Ms},$$

où a_{ik} indique l'intensité du phénomène dans le $i^{\text{ième}}$ cas ($i = 1, 2, \ldots, m_k$) de la section $k^{\text{ième}}$.

On en tire l'indice de dispersion

$$(4) \qquad Q_1 = \sqrt{\frac{\sum_{k=1}^{s} T_k^2 - \frac{T^2}{s}}{\sum_{k=1}^{s} \sum_{s=1}^{m_k} a_{ik}^2 - \frac{T^2}{Ms}}}.$$

[41] *Voir* l'article cité : *Estensione della teoria della dispersione* ..., § 8.

On démontre que la dispersion qui en résulte est normale lorsque sont normales, soit la dispersion du nombre des grandeurs, soit celle de leurs intensités moyennes, et qu'il y a, d'autre part, corrélation nulle entre ce nombre et ces intensités moyennes ([42]).

Pour $a_{ik} = \text{const.}$, la formule (4) se réduit à la formule (3.1).

Sur la base du schéma des rapports de dérivation, la valeur de l'expression

$$\sum_{k=1}^{s} T_k^2 - \frac{T^2}{s}$$

est au lieu

$$\frac{n(s-1)}{ns-1}\left(\sum_{k=1}^{s}\sum_{i=1}^{m_k} a_{ik}^2 - \frac{T^2}{ns}\right).$$

De là on déduit l'indice de dispersion

$$Q_5 = \sqrt{\frac{\displaystyle\sum_{k=1}^{s} T_k^2 - \frac{T^2}{s}}{\displaystyle\frac{n(s-1)}{ns-1}\left(\sum_{k=1}^{s}\sum_{i=1}^{m_k} a_{ik}^2 - \frac{T^2}{ns}\right)}} \tag{5}$$

qui, pour $a_{ik} = \text{const.}$, se réduit à la formule (2.1).

Il est évident que les valeurs des indices de dispersion qu'on tire des formules (4) et (5) sont indépendantes de l'unité de mesure.

11. Passons au cas, plus général, dans lequel les s sections ne sont pas équivalentes, mais présentent une probabilité différente pour qu'une grandeur rentre dans l'une ou dans l'autre d'entre elles.

C'est, par exemple, le cas lorsque nous voulons examiner la dispersion des montants des dommages causés par les tremblements de terre dans diverses circonscriptions territoriales ayant une étendue différente, ou bien la dispersion des montants des indemnités à payer aux assurés qui ont été malades dans les diverses circonscriptions administratives qui comprennent des nombres différents d'assurés.

Dans ce deuxième exemple nous devons appliquer le schéma des

([42]) *Ibid.* § 9.

rapports de dérivation, car le nombre des assurés qui ont été malades présente une limite définie par le nombre des assurés; dans le premier exemple, nous devons au contraire appliquer le schéma des rapports de composition car on ne peut pas fixer une limite au nombre des tremblements de terre dans une circonscription territoriale.

L'indice de dispersion du deuxième ordre sur la base des rapports de composition est donné par la formule

$$(6) \qquad Q_6 = \sqrt{\frac{\sum\limits_{k=1}^{s}\left(\frac{T_k}{T} - \frac{n_k}{N}\right)^2}{\frac{\sum\limits_{k=1}^{s}\sum\limits_{i=1}^{m_k} a_{ik}^2}{T^2} - \frac{\sum\limits_{k=1}^{s} n_k^2}{MN^2}}}.$$

et l'indice de dispersion du deuxième ordre sur la base des rapports de dérivation par la formule

$$(7) \qquad Q_7 = \sqrt{\frac{\sum\limits_{k=1}^{s}\left(\frac{T_k}{T} - \frac{n_k}{N}\right)^2}{\frac{N(s-1)}{(N-1)s}\left(\frac{\sum\limits_{k=1}^{s}\sum\limits_{s=1} a_{ik}^2}{T^2} - \frac{1}{N}\right)}}.$$

12. Il importe d'observer qu'une même série de montants peut être scindée de plusieurs manières en ses composantes énumérables et mesurables. Par exemple, la quantité de l'eau tombée au cours des précipitations atmosphériques de l'année peut être considérée comme le produit du nombre des précipitations par la quantité moyenne de l'eau tombée au cours de chaque précipitation, ou bien comme le nombre des jours où se sont produites des précipitations atmosphériques multiplié par la quantité moyenne de l'eau tombée dans chacun d'eux.

Dans le premier cas, on applique le schéma des rapports de composition, puisqu'il n'y a pas de limite définie au nombre des précipitations qui peuvent avoir lieu dans une année; dans le second cas, on applique le schéma des rapports de dérivation, puisque le nombre de jours où des précipitations se produisent a une limite définie par le nombre des jours de l'année.

Or, il n'est pas dit que l'application des deux schémas conduise au même résultat. Effectivement, le concept de dispersion normale dans un cas est différent du concept de dispersion normale dans l'autre. Dans le premier cas, il faut que les nombres de précipitations se distribuent accidentellement au cours de l'année; dans le second, ce sont les nombres des jours avec précipitation qui doivent être ainsi distribués; ce n'est pas la même chose, car il peut arriver que le nombre des jours où des précipitations se produisent varie accidentellement, mais non pas celui des précipitations, du moment que, dans un même jour, des précipitations peuvent se produire en grand ou en petit nombre et il peut arriver qu'une seule précipitation se produise.

13. Dans quelques cas, la même série de grandeurs énumérables peut être traitée sur la base du schéma de la dispersion du premier ordre ou sur la base de celui de la dispersion d'ordre supérieur.

Prenons l'exemple du nombre des secousses de tremblement de terre ayant eu lieu dans diverses années. Elles constituent des grandeurs énumérables dont il est possible d'étudier la dispersion du premier ordre sur la base des rapports de composition. Mais nous pouvons aussi scinder le nombre de ces secousses en deux composantes : le nombre des tremblements de terre et la moyenne des secousses pour chaque tremblement. Nous pourrons alors considérer ces grandeurs comme des *grandeurs énumérables à composantes énumérables*.

Nous donnerons à ces grandeurs la dénomination de *montants intensifs*.

Nous pourrons étudier leur dispersion du deuxième ordre sur la base du schéma des rapports de composition et appliquer la formule (6).

Évidemment, les résultats auxquels on arrive sur la base de ces deux schémas sont différentes. Nous avons déjà mis en lumière la relation qui existe entre les indices de dispersion que l'on obtient au moyen des deux schémas ([43]).

([43]) *Voir* l'article déjà cité : *Estensione della teoria della dispersione* ..., § 10, et l'autre article également cité, *Généralisations et applications de la théorie de la dispersion.* Dans ce deuxième article, on peut trouver aussi plusieurs applications des schémas de dispersion du second ordre aux montants intensifs et extensifs et précisément aux secousses des tremblements de terre, aux nombres des bâtiments endommagés par les incendies et aux montants des dommages, aux nombres des Communes et aux superficies endommagées par la grêle, à la durée des maladies et aux remboursements liquidés.

14. De ces considérations et de ce que nous avons dit au paragraphe 11, il apparaît clairement que les conclusions sur la dispersion normale, supernormale et sous-normale d'une série n'acquièrent un sens précis que lorsqu'est précisé le schéma sur la base duquel on juge la dispersion.

15. Dans les paragraphes précédents nous avons parlé de dispersion d'ordre supérieur, mais, en réalité nous n'avons fourni que des exemples de dispersion du deuxième ordre. Il est aisé de donner des exemples de grandeurs énumérables ou mesurables dont on peut étudier la dispersion du troisième, quatrième, . . . , ordre.

La série des nombres des tremblements de terre peut être traitée suivant le schéma de la dispersion du premier ordre, et la série des nombres des secousses du tremblement, ainsi que nous l'avons dit, suivant le schéma de la dispersion du deuxième ordre, puisque la seconde urne contient des boules avec la mention du nombre moyen des secousses pour chaque tremblement. La série des nombres des bâtiments endommagés par les tremblements de terre peut être traitée suivant le schéma de la dispersion du troisième ordre, puisque la troisième urne contient les mentions relatives au nombre moyen des bâtiments endommagés par chaque secousse. La série des montants des dommages causés par les tremblements de terre, enfin, peut être traitée suivant le schéma de la dispersion du quatrième ordre, la quatrième urne fournissant les indications relatives au dommage moyen causé à chacun des bâtiments endommagés.

De façon analogue, la série du nombre des assurés qui sont tombés malades pendant une année peut être traitée, aux fins de la dispersion, par le schéma de la dispersion du premier ordre; la série de leurs maladies, par le schéma de la dispersion du deuxième ordre, la deuxième urne fournissant les indications relatives au nombre moyen des maladies par malade; la série des journées d'absence pour cause de maladie, par le schéma de la dispersion du troisième ordre, la troisième urne fournissant les indications relatives au nombre moyen des jours d'absence pour chaque maladie; la série des dommages qui en sont dérivés pour l'administration, par le schéma de la dispersion du quatrième ordre, la quatrième urne fournissant les indications relatives au dommage moyen que l'administration a subi pour chaque jour d'absence.

16. Il est évident qu'une même série peut être traitée par des schémas d'ordres divers, puisque la grandeur résultante peut être scindée en un nombre divers de composantes. Nous avons vu que le montant des dommages pour cause de maladie peut être scindé en quatre composantes, mais il pourrait être scindé aussi en trois composantes, auxquelles, aux fins de la mesure de la dispersion théorique, correspondraient trois urnes : nombre des maladies, nombre moyen des maladies par malade, moyenne des dommages que l'administration subit pour chaque malade, ou bien aussi : nombre des malades, nombre moyen des absences par malade; moyenne des dommages pour chaque jour d'absence; ou encore : nombre des maladies, nombre moyen des jours d'absence pour malade, moyenne des dommages pour chaque jour de maladie. Et il pourrait aussi être scindé en deux composantes, auxquelles correspondraient deux urnes qui fourniraient les données concernant le nombre des malades et la moyenne des dommages pour chaque malade, ou bien concernant le nombre des absences et la moyenne des dommages pour chaque absence, ou encore concernant le nombre des maladies et la moyenne des dommages pour chaque maladie.

17. Remarquons que les différentes composantes qui interviennent dans l'étude de la dispersion d'ordre supérieur peuvent être toutes limitées ou bien toutes illimitées ou, enfin, en partie illimitées et en partie limitées. Par exemple, est illimité le nombre des précipitations atmosphériques ainsi que le montant de l'eau tombée; est limité, par contre, le nombre des jours où des précipitations se produisent, mais est illimitée la quantité d'eau tombée par jour. Sont limités, au contraire, le nombre des jours de grêle ainsi que les dommages causés par elle aux récoltes. Illimité sont, soit le nombre des tremblements de terre soit le nombre des secousses pour chaque tremblement; sont limités, par contre, le nombre des bâtiments endommagés et le montant des dommages qui leur ont été causés. Est limité le nombre des malades, qui ne peut dépasser le nombre des individus exposés à tomber malades, et est illimité le nombre des maladies, mais limité le nombre des jours d'absence ainsi que le montant des dommages incombant à l'administration par suite de ces absences.

18. Les considérations développées dans ce chapitre montrent que

le concept de dispersion peut être convenablement étendu de façon à devenir applicable à des séries de grandeurs auxquelles on n'avait pas pensé par le passé; elles montrent pareillement que la dispersion d'une série de grandeurs énumérables et mesurables peut souvent être mesurée de diverses façons, de sorte que les concepts de dispersion normale, ou supernormale, ou sous-normale n'acquièrent de signification qu'en relation avec le schéma sur la base duquel la dispersion est mesurée.