

JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

PONTUS E. FAHLBECK

La régularité dans les choses humaines ou les types statistiques et leurs variations

Journal de la société statistique de Paris, tome 41 (1900), p. 188-201

http://www.numdam.org/item?id=JSFS_1900__41__188_0

© Société de statistique de Paris, 1900, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

III.

LA RÉGULARITÉ DANS LES CHOSES HUMAINES OU LES TYPES STATISTIQUES ET LEURS VARIATIONS (1).

MESSIEURS,

Il est vraiment surprenant de constater combien de temps il a fallu à l'humanité pour arriver à cette idée qu'elle-même et tous ses actes, tous ses faits et gestes, pouvaient être soumis au nombre et à la mesure. Malgré sa connaissance profonde

(1) Cette communication se base essentiellement sur une étude publiée en 1897, par M. Fahlbeck, en suédois, sous le titre : *Den Statistiska typen eller regelmässigheten i de mänskliga foreteelserna.*

de l'homme, l'antiquité n'avait pas soupçonné l'existence d'un phénomène aussi remarquable, car les nombres mystiques de Pythagore ne correspondent à rien de semblable. Ce phénomène fut révélé au grand jour pour la première fois lorsque *Süssmilch*, guidé par les observations antérieures de *Graunt*, constata que les naissances, les mariages et les décès se représentent tous les ans avec une fréquence à peu près égale dans un seul et même groupe social important. Il compare avec orgueil sa découverte à celle de Christophe Colomb. Cependant plus de cinquante ans s'écoulèrent avant qu'on commençât à tirer parti des découvertes contenues dans l'ouvrage célèbre de *Süssmilch* sur *Die Göttliche Ordnung in den Veränderungen des menschlichen Geschlechts*. Ce furent d'abord les savants français, puis, surtout le Belge *Quételet*, qui étendit considérablement notre connaissance de la régularité des faits humains, en la constatant dans des domaines nouveaux, ceux de la statistique morale et de l'anthropométrie, et en cherchant à préciser sa nature et son essence.

Mais ces essais eurent des conséquences désastreuses pour les recherches qui suivirent, concernant les mêmes phénomènes. En effet, en identifiant la régularité dans les choses humaines à la régularité des lois naturelles, *Quételet* en vint à considérer la liberté d'action des individus simplement comme une « cause accidentelle » et « perturbatrice ». Il en résulta, à la fois parmi les initiés et parmi les non-initiés, une grande controverse sur la situation que créait au libre-arbitre la régularité nouvellement découverte, et cette dispute détourna les recherches du champ ouvert par *Süssmilch*. Au lieu de poursuivre leurs investigations, la grande masse des statisticiens qui se sont occupés de ces questions ont, avant tout, lutté *pour* ou *contre*, la plupart *contre*, les propositions de *Quételet*, oubliant ainsi pour une bonne part la question la plus importante pour le statisticien, à savoir le fait même de cette régularité découverte. Ajoutez à cela qu'à la même époque l'amélioration de la statistique administrative a accaparé presque complètement l'attention des hommes du métier. Elle avait besoin de méthodes plus parfaites, et d'un champ d'observations plus étendu ; par suite la statistique scientifique a dû rester en arrière.

Cependant il est temps que l'étude de la régularité elle-même dans les événements humains — ce qui est proprement l'objet de la statistique scientifique — soit reprise. Dans ces derniers temps, la méthode statistique et son application sont devenues l'objet de recherches approfondies d'une part, par des mathématiciens tels que feu FECHNER et PEARSON, et, d'autre part, par les biologistes. Sans doute, nous ne manquons pas absolument de tentatives faites dans ce domaine par les statisticiens eux-mêmes. Ainsi, les statisticiens doués pour les mathématiques, tels que *Knapp*, *Westergaard*, *Bortkewitz* et autres, ont appliqué le calcul des probabilités à l'explication des régularités statistiques, mais ils ne l'ont guère fait que comme « illustration ». Cependant, ce n'est pas tant du haut calcul mathématique que de l'examen inductif des régularités en question, dont nous avons besoin.

Les points que nous avons à examiner et à préciser, ce sont d'abord l'apparition et l'extension de la régularité statistique, puis le degré de sa *constance* ou de sa *variabilité*, ensuite *ses causes*, et enfin le rapport entre *la norme et les variations*, ainsi que les différentes espèces et les causes desdites variations. Il y a là tout un programme de recherches que l'étude de la régularité dans les faits humains doit se proposer. Le court exposé qui va suivre constitue simplement un certain nombre de

contributions destinées à l'éclaircissement de ces questions si différentes et dont chacune à elle seule embrasse un très vaste domaine.

Mais avant de m'y engager, je crois qu'il y a lieu de définir sommairement ce qu'est au fond la régularité statistique. Süssmilch la caractérisait comme « la répétition, dans un espace de temps déterminé, d'un nombre à peu près égal de cas dans un certain groupe d'individus » ; et après Süssmilch, d'autres auteurs ont donné des définitions analogues. Il est pourtant clair que ces définitions n'expriment que les caractères extérieurs ou l'apparition de la régularité, et non son essence intime. Car qu'est-ce qui, considéré à un point de vue général, se reproduit ? Une mesure, la mesure du fait en question. Ainsi donc, comme la statistique elle-même n'est que la science des mesures humaines en général, *la mesure égale et constante* est l'essence de la régularité statistique. Et si l'on fait un pas de plus dans l'analyse, on découvre que *cette mesure constante ou typique*, représentée d'une façon concrète par un chiffre moyen, une valeur moyenne, etc..., est la *mesure* du phénomène par opposition à celles des cas différents dudit phénomène, autrement dit est ce qu'on pourrait appeler une *idée générale numérique* ou quantitative, différant sur bien des points de celle de la logique ; mais, comme elle, *une loi* et *une norme* pour les individus et cas ou unités particuliers. Ces unités peuvent être d'ailleurs des individus, des groupes d'individus, des séries annuelles de naissances, de crimes et toute autre donnée que nous considérons comme un objet indépendant. La mesure égale et constante est pour ainsi dire leur *type primitif*, le type dont ils se rapprochent tous plus ou moins. C'est là que réside aussi, comme on le sait, la différence entre le nombre moyen commun purement abstrait, et la mesure-type. Le *nombre moyen abstrait* est simplement *une donnée numérique* à laquelle ne correspond pas un rapport déterminé entre des mesures particulières et le nombre moyen et qui par suite n'exprime aucune régularité ou norme. Au contraire, la *mesure-type* constitue avec les mesures particulières une courbe continue, comparable à celle que forment les observations faites sur une étoile ou bien les coups dans un jeu de hasard. Par rapport à la *mesure-type*, les mesures des cas particuliers sont des *variations* qui se groupent dans un ordre déterminé autour d'elle, comme autour de leur norme. Telle est donc la régularité statistique : *la mesure égale et constante* ou *le type*, c'est-à-dire *l'idée générale numérique au milieu de ses variétés*.

Si maintenant nous passons de cette détermination de la notion générale à une étude plus précise des « types », ainsi que j'appellerai par la suite la régularité statistique, on doit remarquer tout d'abord que nous n'y sommes pas encore arrivés à nous rendre clairement compte de leur *apparition*. On les a saisis là où ils se présentaient, pour ainsi dire par hasard, sur le chemin de l'observateur, mais où peut-on les trouver ? Dans tous les faits statistiques ou seulement dans quelques-uns, et en ce cas, dans lesquels ? Ce sont là des points sur lesquels aucune enquête générale n'a été encore entreprise. Pourtant, ce que nous ont appris les travaux de Süssmilch et de Quételet, ainsi que ceux de plusieurs de leurs successeurs, c'est que les types ne se présentent pas seulement dans la *démographie*, en ce qui concerne le mouvement de la population, etc., mais encore dans la *statistique morale* et dans l'*anthropométrie*.

Or, en dehors de ces branches de la statistique, la statistique *économique*, elle aussi, fournit beaucoup de cas de mesures plus ou moins constantes, par exemple en ce qui concerne le rapport du nombre des têtes de bétail au chiffre de la po-

pulation, la consommation par tête des denrées usuelles, la production industrielle, etc. De plus, il y a encore en dehors des faits qui ont été jusqu'ici l'objet de relevés statistiques, une foule d'autres faits qui présentent des mesures-types à un degré très caractéristique. Citons, par exemple, les *ouvrages humains* destinés à un but déterminé, depuis les outils et machines, les ustensiles de ménage et les maisons, les navires, etc., jusqu'aux établissements et institutions de toute espèce. Enfin il est certain que le développement intellectuel ainsi que le développement physique et les dispositions naturelles sont soumis à des mesures déterminées ; de même le mouvement d'échange entre les différentes classes sociales, ainsi que l'extinction des races, bien que nous sachions encore fort peu de choses sur toutes ces catégories de phénomènes.

D'un autre côté, il existe des faits statistiques pour lesquels il n'y a évidemment aucune mesure-type : de cette nature sont les faits concernant le territoire et le nombre de la population, la répartition de la population d'après les différentes religions ou d'après les nationalités, etc..., tous ces faits étant le résultat d'événements historiques qui se sont produits une fois, mais qui ne se reproduiront plus. De même il est clair que les phénomènes que l'on rencontre dans un mouvement rapide de progrès ou de décadence ne peuvent avoir une mesure-type qu'au sens modifié du mot, aussi longtemps que ce mouvement se poursuit. C'est le cas pour la plupart des phénomènes économiques et moraux de notre époque. Dans les époques plus anciennes comme au moyen âge, ces phénomènes présentaient sans nul doute une mesure fixe parfois pendant des siècles. Les mêmes faits peuvent montrer à une certaine période une grande régularité, et perdre cette régularité dans une autre période. Préciser ces observations, et d'une façon générale déterminer les cas où se présentent les types et ceux où ils ne se présentent pas, c'est là une des premières tâches que la science des types statistiques se trouve avoir à remplir.

A ces brèves indications sur l'*apparition* des types, nous devons en ajouter un autre sur leur *recherche*.

Un postulat qui revient souvent en statistique, c'est celui des « grands nombres », avec son corollaire de la « loi des grands nombres ». Ce postulat est invoqué précisément lorsqu'il est question de la régularité des faits statistiques, car on a considéré à peu près généralement que les grands nombres, d'autant meilleurs qu'ils sont plus grands, étaient nécessaires pour découvrir la régularité en question. Cependant, c'est là une erreur où le calcul des probabilités et la considération des jeux de hasard ont conduit les statisticiens. Pour ces calculs le postulat a raison ; mais il n'en est pas de même en matière de statistique et dans la recherche des types. Nous en avons une preuve historique dans ce fait que ni Süssmilch, ni Quételet n'avaient de grands nombres à leur disposition, ce qui ne les a pas empêchés de constater, avec beaucoup de précision, par exemple, la proportion des sexes dans les naissances. En outre, il est dans la nature des choses que les grands nombres en général ne sont pas favorables au dégagement des types statistiques. Ainsi, pour trouver la mesure de la mortalité, il ne serait pas à propos de prendre les décès non seulement pour un pays, comme cela se fait ordinairement, mais pour plusieurs pays, ou bien de les considérer pour une période de 100 ans, au lieu de se contenter d'un plus petit nombre d'années. En effet, ce que me donnerait cette masse de matériaux, ce ne serait pas le type de la mortalité dans sa réalité concrète en temps et lieu déterminés, mais une abstraction vaine, une expression arithmétique au lieu

d'un type. Par là aussi, je me priverais de toutes les espèces et sous-espèces particulières de mortalité qui peuvent se présenter, à savoir : mortalité masculine et mortalité féminine, mortalité aux différents âges, aux différentes époques de l'année, dans les divers métiers, etc..., variétés qui, seules, nous donnent la connaissance complète et réelle de la mortalité pour un temps et dans un pays déterminés. En effet, les types statistiques forment des espèces et des sous-espèces, comme les phénomènes eux-mêmes ; et ces subdivisions ne s'obtiennent qu'en limitant, comme il convient, la masse des matériaux statistiques. Ainsi donc, pour arriver au type, il ne faut pas de trop grands nombres, ni non plus, il faut le dire, de trop faibles, expression d'un trop petit nombre de variétés. L'importance « de la masse statistique » dont on a besoin pour en constater le type, varie de cas à cas ; elle ne peut être déterminée que par l'expérience.

Si nous passons maintenant à l'importante question de l'*invariabilité*, ou de la *variabilité* des types statistiques, en un mot de leur *constance*, nous trouvons devant nous un champ en grande partie inexploré. Il est tout naturel que les premiers chercheurs qui ont découvert la régularité des phénomènes humains, n'aient vu que leur immutabilité. Ils avaient aussi à leur disposition un si maigre bagage de matériaux que leur erreur sur ce point est fort excusable. Les chercheurs qui sont venus après eux, et en particulier tous ceux qui ont combattu la doctrine de Quételet sur le libre-arbitre, n'ont pas manqué de faire voir que spécialement les types fournis par lui dans le domaine des faits moraux étaient très sujets à changer. Cependant, il est clair, après ce que nous savons déjà de ces types, qu'ils présentent de grandes différences, et qu'il serait peut-être possible de les classer d'après leur plus ou moins grande constance. Au point de vue de la *constance*, nous placerions ainsi en première ligne certains phénomènes purement *physiologiques*, tels que le sexe des naissances, et la plupart des données anthropométriques, ces dernières étant toutefois limitées à chaque race ou à chaque nationalité particulière. Immédiatement après, nous pouvons placer, comme présentant le plus grand degré de constance, tout un ensemble de *créations proprement humaines*, destinées à un but déterminé : par exemple, l'importance des familles et des maisons, qui est encore un phénomène fort constant, mais qui l'a été surtout autrefois. La même observation s'applique à une foule d'objets matériels, sur lesquels nous manquons cependant de renseignements statistiques. Si nous pouvions, par exemple, donner les mesures des chaises et des tables, nous y trouverions sans doute une constance remarquable, presque aussi grande que dans les phénomènes naturels dont nous avons parlé. Il en est tout autrement d'un grand nombre d'autres types statistiques, tels que les *types économiques*, qui dépendent des besoins de l'homme (données quantitatives relatives au commerce, à l'industrie et à l'agriculture). Ces données sont, sans exception, très variables. Même remarque pour un grand nombre de types *démographiques* et pour la plupart de ceux qui relèvent de la *statistique morale*. Mais la variabilité que nous observons dans la mesure de tous ces phénomènes est permanente et prend volontiers pendant un temps assez long une seule et même direction, ce qui assurément ne l'empêche pas de faire volte-face pour prendre une direction opposée, mais alors encore pour une période assez longue.

Comme on peut le concevoir d'après cela, il règne la plus grande diversité en ce qui concerne la constance des régularités statistiques. Le diagramme I, que vous avez sous les yeux (voir diagramme, p. 193), vous en fournit une illustration, en

Suède.

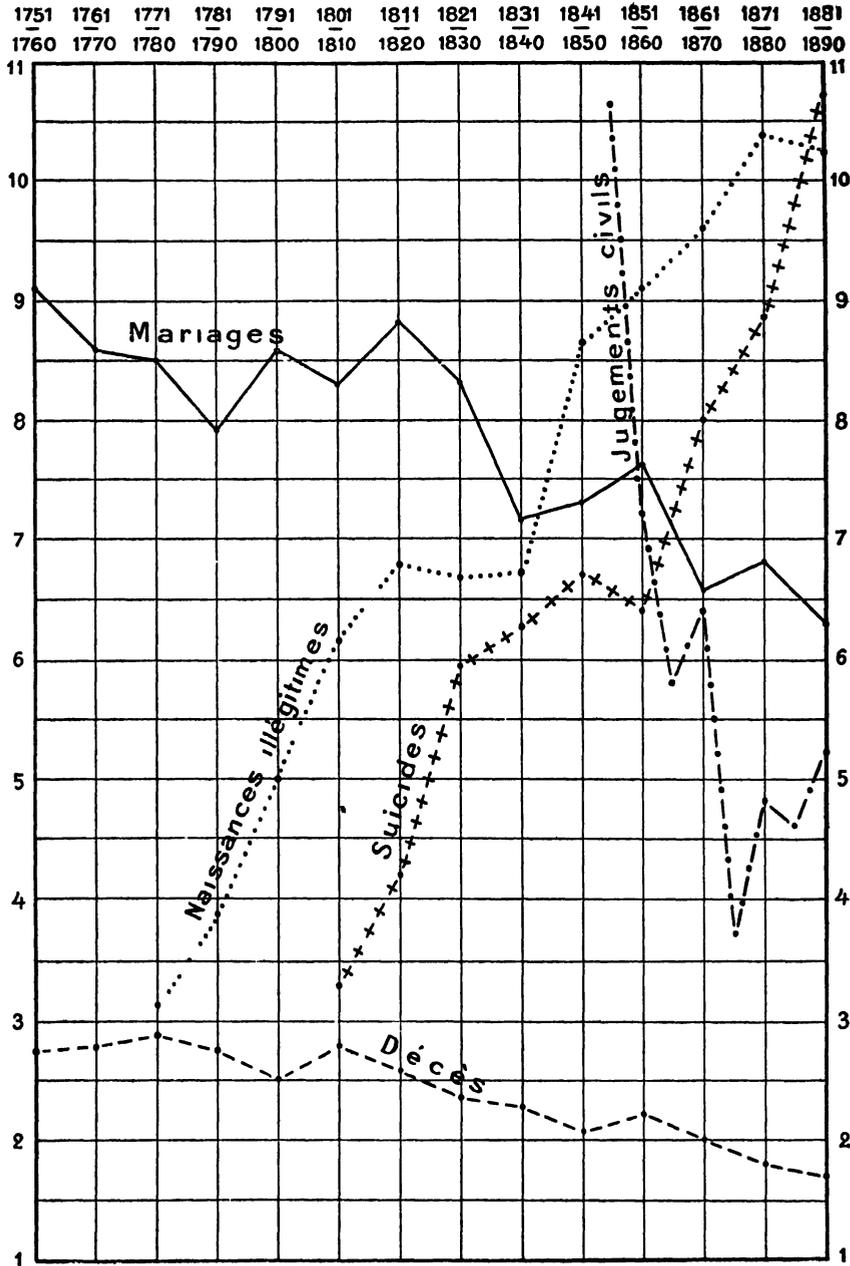


Diagramme I

- Mariages, pour 1 000 habitants sur la population moyenne calculée par moyennes décennales de 1751 à 1890
- - - - - Décès pour 100 habitants, sur la population moyenne calculée par moyennes décennales, de 1751 à 1890
- Naissances illégitimes pour 100 enfants nes vivants, en moyennes décennales, de 1871 à 1890
- · - · - · - Jugements civils de première instance, pour 1 000 habitants, sur la population moyenne calculée par moyennes quinquennales de 1851 à 1890
- + + + + + Suicides, pour 100 000 habitants, sur la population moyenne calculée par moyennes décennales de 1801 à 1890.

même temps que la confirmation de circonstances et de relations qui sont bien connues de vous. Cependant on se forme, en présence de ces schèmes, une toute autre idée de la régularité statistique que celle qu'on se faisait autrefois. Parler ici de lois et de mesures invariables est, pour une bonne part, une exagération.

Bien plus, une connaissance plus complète des données statistiques peut amener à se demander si l'on est bien en droit de parler de régularité, de mesure-type, dans une foule de cas, où sans doute le contingent annuel reste à peu près le même, mais présente cependant toujours certaines différences en plus ou en moins. Naturellement, on peut sur cette question différer d'appréciation; pourtant, il me semble qu'il ne faut pas exiger ici une constance trop rigoureuse. Toutes les choses humaines changent. Invariabilité et variabilité sont par suite des idées relatives. Même les mesures qui sont soumises sur certains points à des changements continus peuvent fort bien être considérées comme typiques, lorsque l'augmentation ou la diminution réitérée ne dépasse pas quelques pourcentèmes de la grandeur totale du fait considéré. On doit seulement faire une distinction entre les types *constants* et les types *variables*.

Ajoutons enfin une remarque au sujet d'une confusion que l'on fait assez souvent entre la *variabilité des types* et les *variations* des éléments individuels qui ont servi à établir ces types. Ce sont là deux choses bien différentes. La *variabilité* des types se présente fréquemment, mais *elle peut aussi cesser* pour un seul et même phénomène. La mortalité peut dans une certaine période croître ou diminuer, mais peut dans une autre période se maintenir constante. *Au contraire, les variétés existent toujours*, puisqu'elles sont les mesures correspondantes aux cas particuliers ou unités individuelles, lesquelles sont toujours dissemblables. Pour apprécier la régularité statistique, il est nécessaire de séparer les deux choses et c'est ce qu'on n'a pas toujours fait. Ainsi les variétés annuelles du *suicide* sont souvent très petites, beaucoup moindres que celles de la mortalité, et pourtant la variabilité du type du suicide est beaucoup plus considérable que celle de la mortalité, comme il ressort du diagramme I (voir le diagramme, p. 193). Aussi bien Quételet que Ad. Wagner, se sont principalement attachés au peu d'amplitude des variations dans le suicide, et beaucoup moins à la forte variabilité du type, et ont été amenés par là à exagérer la régularité du phénomène. Ajoutons comme remarque que la variation dans le suicide est loin d'être aussi faible que ces deux auteurs l'ont cru, et qu'elle est souvent très grande, comme le montre la statistique pour la Suède. Mais des erreurs de ce genre se sont souvent produites au cours des débats sur la régularité statistique.

Ce qui paraît plus difficile à expliquer que toute autre question relative à la régularité statistique, c'est *la question de ses causes*. Il n'en est pas moins vrai que, dans la statistique administrative et surtout dans la statistique scientifique, la recherche des causes est souvent le but de tous les efforts. Il arrive aussi très fréquemment que la statistique, tout simplement par un usage convenable des données dont elle dispose, peut révéler l'origine de tel ou tel changement dans la mesure des phénomènes, autrement dit, pénétrer jusqu'aux causes de certaines variations déterminées. Mais de là à découvrir les causes des régularités statistiques elles-mêmes, il y a une grande distance; et cette distance, personne ne l'a franchie. Pour expliquer ces causes, on a eu recours tantôt à la Providence ou à des forces plus ou moins mystérieuses, tantôt aux lieux communs. Ainsi on a allégué à cet égard la proposition que « des causes semblables produisent des effets semblables », ou,

avec Quételet, que « les effets sont proportionnels aux causes », ou enfin, avec un auteur moderne, Goldschmitt, qu'il serait beaucoup plus étonnant de ne pas trouver de régularité du tout qu'il ne l'est d'en trouver, toutes propositions qui sont évidentes par elles-mêmes, et dont la première peut rendre parfaitement compte de la cause générale de toute régularité, mais qui, précisément pour cela, n'expliquent absolument rien. Donner la raison des types statistiques, c'est montrer les causes particulières d'où ils proviennent dans chaque cas. Il est évident que c'est encore pour nous une chose impossible en ce qui concerne *tous* les types existants, mais c'est, dans *beaucoup de cas*, une chose assez facile, pourvu qu'on *spécialise* les régularités, qu'on cherche les causes de chaque groupe pris à part, et que de plus on n'aille pas naturellement demander à la statistique plus que ce qu'elle peut donner, c'est-à-dire *la cause la plus proche*. En effet, nous n'avons qu'à suivre les indications fournies par l'étude précédente sur la constance des types pour en trouver les causes saisissables par la statistique.

Ainsi il est clair qu'il faut chercher dans les *dispositions héréditaires* la cause de la constance observée dans le sexe des naissances, dans l'âge de la fécondité pour les femmes et pour les hommes, dans les mesures physiques de l'homme, et dans un grand nombre d'autres particularités physiques ou morales non encore soumises à la statistique. Chaque race ou chaque famille a sa mesure déterminée pour ce qui est de la taille, de la couleur des cheveux et des yeux, etc..., ainsi que d'autres groupes plus considérables encore, en ce qui concerne le sexe des naissances ; mais tout cela a pour fondement des particularités héréditaires de nature physiologique qui restent elles-mêmes constantes, et qui expliquent ainsi la raison d'être des types statistiques.

Il est évident que nous ne pénétrons pas ainsi jusqu'à la *cause dernière* de cette constance, mais c'est à la physiologie, et non à la statistique qu'il appartient de dire ce que c'est que l'hérédité, et il en est toujours de même des causes dernières de la régularité statistique. Il suffit à la statistique de connaître la *cause proche* et cette cause est dans le cas présent une disposition héréditaire. En vertu de cette origine, on peut appeler ces types des *types naturels*, en entendant par là qu'ils sont constants, comme la nature humaine elle-même. Car, en fait, ce sont là les seuls types invariables qui existent.

Une autre cause également évidente de la régularité statistique, c'est la *fin* ou l'*idée* qu'on trouve comme principe de toutes les œuvres humaines, physiques ou morales. Telle est la raison pour laquelle les données relatives à l'importance des familles, au train de maison, peuvent se ramener à des types, aussi bien que les fondations au sens idéal du mot (c'est-à-dire institutions), ou les fondations au sens matériel (par exemple, les usines) et de même encore la masse énorme des objets concrets : maisons, navires, machines à vapeur, etc... Tous ces types sont des *types de finalité*, tenant de la finalité qui les anime à la fois leur constance et aussi leur variabilité, surtout à notre époque, époque de crise si intense dans les désirs, dans la conception du but et si ardente dans la recherche des moyens propres à réaliser ces désirs ou à atteindre ce but.

Des types qui touchent de près à ces « types de finalité » sont ceux qui répondent à nos *besoins* : on pourrait les appeler *types de besoin*. Les mêmes œuvres, dont la grandeur est déterminée par leur but, reçoivent de nos besoins leurs nombres et leurs quantités. Toute la masse des productions du travail, telles que vêtements,

denrées alimentaires, matières premières, etc., présentent ainsi, en ce qui concerne les quantités produites, des mesures-types, lesquelles sont, il est vrai, fortement variables. La cause saisissable par la statistique de leur constance relative ainsi que de leur variabilité, ce sont les besoins qui leur ont donné naissance.

En outre des types que nous venons de signaler, avec leurs causes facilement observables, il en est d'autres auxquels nous ne pouvons, à l'heure actuelle, assigner de raison d'être, sauf qu'ils nous apparaissent comme des *résultantes* d'un grand nombre de forces agissant ensemble : dispositions physiques, finalité, besoin, en même temps que toutes sortes d'influences extérieures, physiques et sociales. Quelques-uns de ces types sont, parmi les plus attentivement observés, par exemple la nuptialité, la natalité, la mortalité, le suicide, le crime, etc.... Ici, il est impossible de trouver une cause déterminée à la régularité statistique. Elle semble être, comme nous l'avons dit, la résultante du jeu mobile de forces multiples. Tout ce qui paraît ressortir clairement de la relation de ces types avec les groupes délimités au point de vue social et politique (peuples, sociétés, communes, classes), c'est qu'ils dépendent, à un très haut degré, du milieu social. En attendant qu'on puisse mieux les expliquer, nous pouvons les ranger sous une dénomination commune, en les appelant *types-résultantes*.

Comme il ressort de ces indications sommaires, les causes de la régularité des faits de statistique sont multiples et très diverses. Il est assez remarquable que, comme cause primordiale, nous trouvons, parmi la masse innombrable des « types de finalité », le libre-arbitre humain, qui, croyait-on, avait été éliminé à tout jamais par la statistique. Ce n'est pas ici le lieu d'examiner comment cette liberté peut se combiner avec quelques-unes des régularités désignées en dernier lieu comme types-résultantes. Mais, si l'on réussit un jour à résoudre les résultantes en leurs éléments composants, on trouvera alors la réponse même à ce problème.

Après ces courtes remarques sur les types, il me reste à dire quelques mots de la *variation* des types et des *variétés* qui s'y rattachent. A l'opposé des types précédemment étudiés et qui sont les mesures générales des phénomènes, les variétés sont les mesures des cas particuliers ou unités individuelles. Pour nous, les cas particuliers sont toujours les éléments primaires en statistique comme en logique ; mais, dès qu'on est parvenu de ces unités à la mesure générale, de même qu'à l'idée générale, les choses changent de face. Les mesures particulières deviennent maintenant des variantes d'un type et se présentent vis-à-vis de ce type comme les individus vis-à-vis de l'espèce. C'est précisément en cela que consiste la régularité dans les faits en présence.

Dans le calcul des probabilités, on a donné le nom de *fautes* à ces mesures individuelles (observations particulières ou coups dans les jeux de hasard). C'est là une mauvaise dénomination, qui, sous l'influence de la théorie darwinienne, a été remplacée par celle que nous venons de donner : *variation* ou *variété*. C'est sur l'existence des variations que repose, comme on le sait, toute la théorie de la sélection naturelle. C'est pourquoi l'étude des variations est pour les biologistes un point capital, et cette étude a conduit dans ces derniers temps, comme nous l'avons dit plus haut, à la constitution d'une statistique mathématique des variations. L'étude

des *variations* n'est pas moins importante dans la statistique proprement dite, et elle est indissolublement liée à l'étude des *types*. Je dois cependant me borner à quelques remarques sur ce riche sujet.

Comme il est facile de le comprendre, ce sont les *relations des variétés au type et des variétés entre elles* qui sollicitent tout d'abord notre attention. On peut se représenter ces relations de trois façons différentes : ou bien la variation est égale à 0, ou bien elle est infiniment grande (absolument ou relativement), ou bien enfin elle est entre ces deux extrêmes. Dans le premier cas, il n'y a pas de variantes, dans le second, pas de type ; dans le troisième cas, il y a à la fois variantes et type. La plupart des phénomènes physiques appartient à la *première catégorie*. Le point de congélation de l'eau, la chute d'une pierre, présentent la même mesure dans tous les cas particuliers, tant qu'on reste dans les mêmes conditions extérieures. Au contraire, nous ne trouvons que des variantes et pas de types (*seconde catégorie*) dans une foule de phénomènes inorganiques ou organiques, à savoir dans ceux qui sont le résultat d'événements historiques qui se sont produits une fois et qui ne se reproduiront jamais plus : ainsi la formation de la croûte terrestre et son partage en mers et continents, vallées et montagnes ; de même encore l'extension et la répartition des États et des peuples, etc..., comme nous l'avons mentionné plus haut.

La *troisième catégorie* de grandeurs comprend celles qui possèdent à la fois type et variations. Ce sont les phénomènes physiologiques et biologiques, ainsi que la plupart des phénomènes sociaux. C'est de ces phénomènes que la statistique des variations dont nous avons parlé aussi bien que la science statistique proprement dite ont à s'occuper. Et le point principal est de préciser, aussi exactement que possible, la relation existant entre le type et les variations.

Je n'ai pas besoin, devant cette assemblée, de dire longuement en quoi consiste d'une façon générale cette relation qui est le groupement, autour d'une mesure moyenne, des mesures particulières en tant que variétés *supérieures* et variétés *inférieures*. Je ne vais pas non plus rendre compte des différentes manières dont est calculée et représentée ladite relation : arithmétiquement, par les chiffres moyens et les variations moyennes, par la valeur centrale et la valeur médiane, par les maxima et les minima, etc. ; géométriquement, par la courbe ordinaire de dispersion ou la courbe de valeur centrale de *M. Gallon*, ou bien par l'oscillation des variations autour d'une diagonale. Ce que je veux démontrer, c'est que le groupement des mesures particulières autour de la mesure moyenne a lieu dans des limites très différentes. Elles peuvent se presser étroitement contre cette dernière, ou bien se déployer et s'en éloigner considérablement, comme le représente schématiquement le diagramme II ci-dessous.

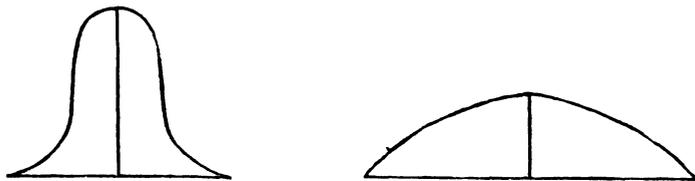


Diagramme II.

Dans le premier cas, les écarts sont faibles, dans le second, ils sont forts, ce qui revient à dire, en d'autres termes, que dans le premier cas la régularité est forte, et qu'elle est faible, dans le second. A cette observation se rattache une autre : c'est

que le nombre de ces unités et cas observés, autrement dit *la masse statistique*, d'où se dégagent les types, peut être importante ou faible, suivant que c'est le premier cas ou le second qui se présente. *Plus les écarts sont petits ou plus leur champ est restreint, moins la masse statistique a besoin d'être grande, et vice versa.* Quant à savoir quelles proportions elle doit avoir pour permettre la constatation du type, c'est là une question qu'il est impossible de trancher *à priori*, comme nous l'avons déjà dit plus haut.

Une autre observation à faire au sujet de la relation du type avec les variations, c'est que d'ordinaire, ces dernières ne se répartissent pas régulièrement autour du type. La courbe est presque toujours plus ou moins *asymétrique*. Et c'est principalement cette circonstance qui a conduit les mathématiciens dont nous avons parlé à donner à la théorie des jeux de hasard une extension générale. Dans la théorie des jeux, la courbe est, comme on le sait, symétrique, et cela d'autant plus que le nombre des coups est plus considérable. Mais dans la nature, une répartition aussi harmonique des mesures se produit très rarement, autant dire jamais. Il y règne une asymétrie plus ou moins grande. Et, si je ne me trompe, on devra toujours, en ce qui concerne toutes les mesures dépendant du développement et de la croissance (par exemple la taille humaine), arriver à cette constatation que le nombre des variétés inférieures est plus grand que celui des variétés supérieures. En effet, des circonstances multiples peuvent avoir pour résultat d'entraver ce développement et fort peu celui de l'encourager. Il n'est au pouvoir de personne d'ajouter un pouce à sa taille, mais des accidents de toute espèce peuvent la faire s'arrêter à un degré inférieur. Une illustration de ce fait nous est fournie par le tableau n° 1, qui représente la taille de 3 869 conscrits. Tandis que le nombre des variétés supérieures à la valeur médiane de 169-170 centimètres est de 1 513, celui des variétés inférieures est de 1 854. (Voir le tableau, p. 199.)

En outre, il est clair que tous les types *variables* doivent présenter une répartition asymétrique des variétés. C'est précisément parce que celles-ci se groupent de préférence d'un côté, comme variantes supérieures ou inférieures, que le type change dans le sens de ce mouvement. Nous en avons un exemple dans le diagramme III ci-dessous.

Fréquence des mariages, pour 1 000 habitants, sur la population annuelle de 1851 à 1895.

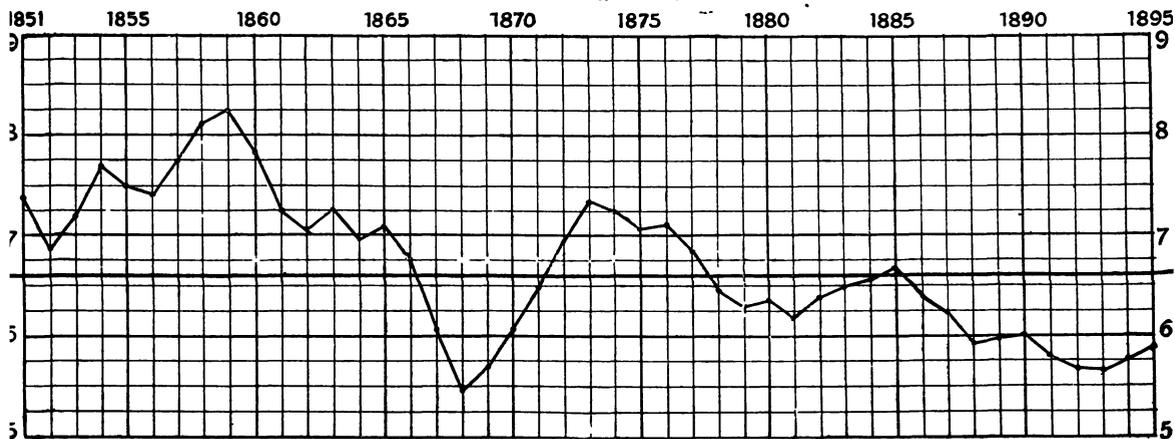


Diagramme III.

— Nuptialité en moyenne 1851-1895.

relatif à la répartition annuelle des mariages autour du type en diagonale.

TABLEAU I.
Taille des conscrits de Lund en 1890-1897.

Taille en centimètres (intervalle de 2 cm.).	Nombre d'individus.
188—187	5
186—185	9
184—183	26
182—181	67
180—179	122
178—177	173
176—175	278
174—173	402
172—171	431
170—169	502
168—167	454
166—165	446
165—163	345
162—161	218
160—159	168
158—157	114
156—155	51
154—153	24
152—151	12
150 et moins.	22

1513

1854



Hauteur moyenne : 169 cm.

S:e. 3869

Il y a beaucoup à dire et beaucoup d'expériences nouvelles à faire sur la relation du type aux variétés. Je dois cependant me borner aux courtes observations que je viens de faire, pour pouvoir ajouter quelques mots sur les *causes des variations*. Pour les biologistes, c'est là le nœud de la question, et c'est aussi pour la statistique proprement dite un point des plus importants. Partout où la statistique, elle aussi, recherche une relation causale, c'est la question de l'origine de telle ou telle variation qu'elle se pose. Comme on peut le comprendre, les causes des variations sont extraordinairement changeantes et multiples. Sur le nombre, il en est toutefois deux qui ont été déjà observées et qui méritent une mention spéciale. L'une se manifeste dans la marche *périodique* de la variation d'après les époques et se fonde par suite sur les variations de ces époques. Déjà *Villermé* avait observé cette cause et *Quetelet* en a démontré tout au long la grande importance pour la plupart des actions et circonstances humaines. Mais, outre ces variations périodiques, on a observé aussi de bonne heure que souvent à la suite d'une forte variation supérieure ou inférieure venait immédiatement une variation analogue en sens opposé. La statistique de tous les pays nous offre des cas de variations de ce genre, variations que j'appellerai *compensatrices*. Nous en avons deux exemples dans les diagrammes IV et V, l'un pour la Suède, l'autre pour la France. A une forte élévation succède une forte dépression dans la marche des variations. (Voir diagramme, p. 201.)

On soupçonna tout de suite qu'il y avait là une relation causale à découvrir, mais on n'y réussit pas. Comme on avait cru remarquer en même temps qu'après une guerre ayant amené une grande consommation d'hommes, il se produisait une augmentation dans le nombre des naissances masculines, on était conduit à voir là, comme en général dans les variations compensatrices, tantôt une disposition providentielle en vue de maintenir l'équilibre dans les choses humaines (Süssmilch), tantôt l'effet d'un mystérieux instinct de conservation des sociétés, d'une force analogue à celle par laquelle un organisme cherche à réparer les pertes subies dans une de ses parties (v. Ettingen). Que ces explications soient peu satisfaisantes, c'est ce que je n'ai pas besoin de dire, d'autant qu'elles ne conviennent guère au cas considéré, aucune variation inférieure dans le nombre des enfants mâles n'ayant précédé la prétendue variation supérieure. Ajoutons que le fait d'une augmentation dans le nombre des naissances masculines à la suite d'une guerre n'est rien moins que certain. On ne peut le constater ni pour la Suède ni pour la France. D'ailleurs, que les choses soient ou non ainsi, ce n'est pas ce qui nous occupe ici, et l'examen de cette question ne peut pas contribuer à l'explication des variations compensatrices. On ne peut pas non plus les expliquer « par une loi de compensation qui tend à rétablir l'équilibre lorsqu'il a été rompu par un accroissement ou une diminution subite ». En effet, ce n'est pas l'équilibre qui suit immédiatement un excédent survenant dans un sens, mais bien un nouvel excédent dans l'autre sens.

L'explication assez simple de ce phénomène semble être la suivante :

Tout excédent survenant dans la marche d'un phénomène, à un moment déterminé, consomme le nombre des cas possibles dans le moment suivant, et cela a pour conséquence la naissance d'une variation inférieure ou inversement. Par exemple, si pendant une année la mort a fait une moisson extraordinairement riche, un grand nombre d'existences faibles ont été éliminées et elles ne chargent plus le budget de la mortalité pour l'année suivante, et celui-ci se trouve extraordinairement réduit. Inversement, si la mort a été plus indulgente qu'à l'ordinaire pendant une période déterminée, la période suivante présentera nécessairement une mortalité plus forte, grâce à la masse conservée des vies faibles. Voilà le secret des variations compensatrices. Élévations et dépressions ont pour cause directe un phénomène contraire précédent, lequel a diminué ou augmenté le nombre des cas possibles.

Mais je dois m'arrêter ici, bien que le sujet soit aussi riche qu'intéressant et ouvre un large champ à de nouvelles expériences. A défaut d'autre résultat, c'est du moins ce qui ressort des indications rapides que j'ai pu donner sur la science des types statistiques. Elles montrent aussi que ces matières peuvent être traitées et exposées sans l'aide de hautes mathématiques, mais simplement avec les moyens ordinaires que la statistique a généralement à sa disposition.

En tous cas, je tiens à vous remercier, Messieurs, de l'occasion que vous m'avez gracieusement offerte d'exposer brièvement devant vous le résultat de mes recherches sur un sujet qui, selon moi, mérite de la part des statisticiens plus d'attention qu'ils ne lui en ont accordée jusqu'à ce jour.

Pontus E. FAHLBECK,
Professeur à l'Université de Lund (Suède).

Proportion des décès, pour 100 habitants, sur la population moyenne annuelle, de 1766 à 1780 (en Suède).

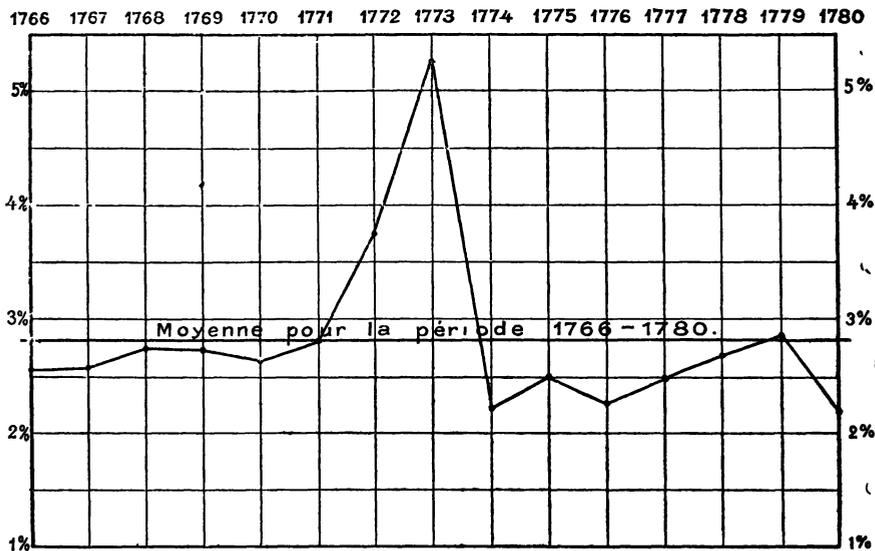


Diagramme IV.

Fréquence des mariages, pour 1 000 habitants, sur la population annuelle moyenne, de 1806 à 1820 (en France).

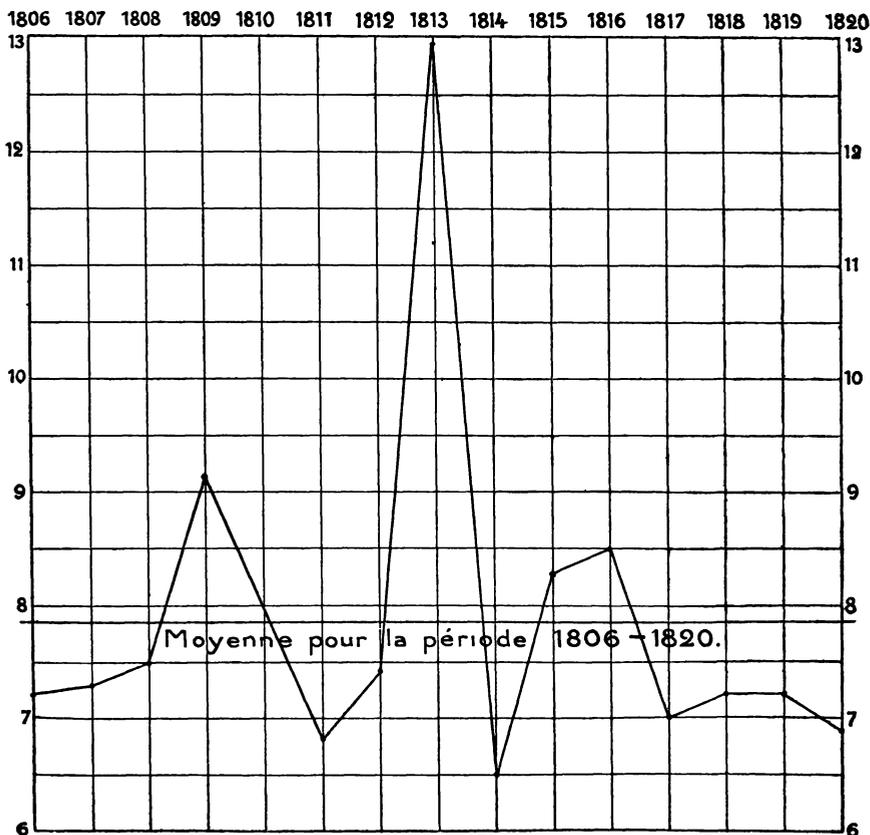


Diagramme V.