

JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

Discussion sur la définition du mot « statistique »

Journal de la société statistique de Paris, tome 97 (1956), p. 253-258

http://www.numdam.org/item?id=JSFS_1956__97__253_0

© Société de statistique de Paris, 1956, 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/>

VI

DISCUSSION SUR LA DÉFINITION DU MOT " STATISTIQUE "

M. DUMAS expose l'essentiel de ce qui a fait l'objet de sa communication publiée dans le Journal de la Société d'avril-mai-juin 1955; on se rappelle qu'il avait alors distingué plusieurs acceptions du mot statistique, les premières étant relatives respectivement :

— à la *science* de méthodes en rapport avec un ensemble de valeurs numériques;

— à l'*art* d'appliquer les enseignements de cette science (de même que le calcul est l'art d'appliquer l'arithmétique);

— au tableau, numérique ou graphique, servant à consigner certains résultats.

Pour ce qui concerne ce que l'on désigne parfois par « statistique mathématique », une quatrième définition était mise en avant : *art d'utiliser le calcul des probabilités à la préparation des décisions*. Encore faut-il s'entendre sur ce qui est du domaine du calcul des probabilités.

M. Dumas voit dans le calcul des probabilités une science qui, quelles qu'aient été les circonstances dans lesquelles elle a pris naissance, est aujourd'hui purement axiomatique, remarque étant faite que, de ce point de vue, tout ce que l'on pourrait être tenté de rattacher au calcul des probabilités et qui ne découlerait pas des axiomes, devrait être délibérément écarté de cette science (il n'est

pas exclu d'ailleurs que cette opération conduise à rejeter certaines parties dans la statistique).

Cela étant, la statistique mathématique serait l'art d'utiliser la science axiomatique dite : calcul des probabilités, de même que, comme il est dit plus haut dans un autre cas, le calcul est l'art d'appliquer l'arithmétique. Est-il nécessaire d'ajouter cette précision : « à la préparation des décisions »? M. Dumas le pense, mais, reconnaissant que cela peut être mal compris, il serait tout prêt à accepter une expression indiquant qu'il s'agit d'une utilisation du calcul des probabilités à un cas concret. Cette utilisation se fait dans des conditions en tous points analogues à celles qui sont de pratique courante en ce qui concerne d'autres sciences axiomatiques : le géomètre rural, chargé d'évaluer la superficie d'un champ « rectangulaire », adopte comme modèle mathématique de ce champ un rectangle parfait et multiplie l'une par l'autre les dimensions des deux bases; mais s'il veut tenir compte de certaines dénivellations existantes, il changera de modèle et trouvera un autre résultat. De même, en accord avec la définition proposée, le statisticien est dans son rôle en adoptant un modèle probabiliste, correspondant au cas à traiter moyennant une hypothèse, et en déduisant de ce modèle, par simple application des enseignements du calcul des probabilités, la probabilité d'arrivée de tel ou tel événement; qu'il vienne à changer de modèle, donc d'hypothèse, et il trouvera une autre probabilité d'arrivée du même événement. S'il arrive que quelqu'un doive prendre une décision comprenant parmi ses éléments d'appréciation, la probabilité d'arrivée de cet événement, la décision ne sera peut-être pas la même suivant que la « valeur » de la probabilité aura été déduite de l'un ou de l'autre modèle.

La proposition de donner à « statistique mathématique » une définition du genre : art d'utiliser le calcul des probabilités... se heurte à une objection venant de ceux qui trouvent que le statisticien est très intimement mêlé à l'élaboration de la science que, d'après la définition, il n'a qu'à utiliser, dont il n'a qu'à appliquer les résultats acquis par d'autres. A cela, M. Dumas répond que dans sa conception ce travail d'élaboration est un travail de probabiliste et qu'aucun de ceux qui se nomment statisticiens ne pourrait se sentir diminué le jour où il se dirait qu'il partage son temps entre faire œuvre de statisticien et faire œuvre de probabiliste. Qu'à l'occasion d'un problème concret le statisticien voie l'intérêt de traiter à partir de bases axiomatiques un problème dont les données correspondent au modèle probabiliste qu'il a décidé d'adopter, voilà qui est fréquent; que ledit statisticien aborde lui-même ce problème, voilà également qui est fréquent, mais ce qui serait alors demandé à ce statisticien, ce serait de se dire qu'en faisant cela, il troque sa « casquette » de statisticien contre celle de probabiliste, quitte à reprendre la première l'instant d'après.

M. Dumas rappelle qu'un problème pratique se pose : donner une définition, formulée à la façon des définitions de bons dictionnaires, pour chacune des acceptions dans lesquelles le mot statistique peut actuellement, compte tenu des usages actuels, être considéré comme correctement employé. Ce problème est différent de celui qui consiste à délimiter un domaine qu'on devrait dorénavant reconnaître comme étant celui de la statistique. Dans la solution du problème posé, on doit se garder d'une part de passer sous silence un domaine

que les usages, regrettables ou non, admettent comme étant, lui aussi, du domaine de la statistique, et d'autre part de se contenter de mettre en évidence un lien logique existant entre les différents domaines correspondant respectivement au mot statistique.

M. PROT propose de distinguer pour le mot « statistique » trois définitions :

Statistique A : tout dénombrement d'êtres, d'objets, de faits ou de mesures constituant les espèces d'un genre.

Statistique B : technique de l'établissement des statistiques A et de la transformation de leur expression.

Statistique C : science résumant le langage, les lois et l'histoire des statistiques A et B.

Les définitions qui précèdent s'efforcent de satisfaire aux règles classiques de la définition :

- 1 — s'appliquer à tout le défini;
- 2 — ne s'appliquer qu'au seul défini;
- 3 — ne comporter que des mots du langage courant ou des mots spéciaux précédemment définis.

Si l'on veut bien accepter deux néologismes, la définition A pourrait, pour plus de précision, se diviser en deux parties :

A-1 : Statistiquer : dénombrer des êtres, des objets, des faits ou des mesures constituant les espèces d'un genre.

A-2 : Statistique : résultat d'un statistiquage.

M. GODIN. — Vu la difficulté de trouver une définition entièrement satisfaisante pour donner les différents sens du mot « statistique », je propose de garder le simple mot de « statistique » pour désigner les listes contenant, en plus ou moins grand nombre, les renseignements ou observations employés dans l'Administration, la Technique ou le Commerce, mais de réserver l'expression « statistique mathématique » pour désigner la science qui utilise les dits renseignements afin d'en déduire, par le calcul, des précisions, des prévisions et des lois.

M. LECLERC. — Il me semble que toutes les définitions qui viennent de nous être soumises présentent l'inconvénient de ne pas marquer assez nettement l'accent sur la notion de variation.

L'objet de nos études est de chiffrer les variations dans le temps ou dans l'espace de phénomènes donnés et c'est ce qui le distingue, entre autres, du dénombrement pur et simple.

M. DUMONTIER. — Je suis plus partisan de la définition de M. Prot que de celle de M. Dumas : j'estime que la première définition est assez exhaustive; cependant, pour être complète, il faudrait — il me semble — une cinquième colonne au tableau que nous présente M. Prot, et libellée : en vue de certains objectifs.

En effet, on rassemble des données en vue d'un objectif et c'est cet objectif qui permet la classification des données. L'exemple de l'indicateur des chemins de fer — qui n'est pas une statistique — éclaire mon propos : celui-ci comprend

— comme son nom l'indique — des renseignements rassemblés à titre indicatif. Au contraire, les mêmes renseignements qui eussent été classés dans un but différent (par exemple, pour totaliser le nombre de gares ou comparer certains kilométrages) auraient déjà constitué en eux-mêmes une opération statistique.

M. CHEVRY. — Je voudrais faire remarquer tout d'abord que les points de vue exposés tour à tour par MM. Dumas et Prot ne me paraissent pas devoir être opposés l'un à l'autre, puisque leurs intentions sont visiblement différentes.

M. Dumas, constatant que le substantif « statistique » est actuellement utilisé en lui donnant des sens divers, a voulu faire l'inventaire de ces acceptions, les classer en quatre ou cinq rubriques et montrer que la définition donnée par le dictionnaire est loin de les recouvrir toutes.

M. Prot au contraire a fait un effort de synthèse et a recherché une définition du mot statistique qui englobe ses divers sens.

Pour ma part, il me semble qu'il y a lieu surtout de faire la distinction entre :

— la Statistique, avec un S majuscule, qui est une science ou une technique et qu'il s'agit de définir,

— et la statistique, avec un s minuscule, ou mieux les statistiques, au pluriel, qui sont les résultats chiffrés de l'application de cette technique (1).

Pour la Statistique avec un grand S, il me paraît indispensable de rappeler l'excellente définition donnée par notre ancien président, M. Dugé de Bernonville dans son ouvrage « Initiation à l'analyse statistique » (page 9) et qui est la suivante :

« Ensemble des méthodes s'appliquant à l'étude numérique des collectivités, en entendant sous ce terme général, non point seulement des collectivités sociales, mais des groupes de faits nombreux de même nature, quels que soient, d'ailleurs, ces faits, pourvu qu'ils puissent être comptés et classés suivant certaines modalités distinctives. »

Cette définition est assez voisine de celle que propose M. Prot.

J. MARCZEWSKI. — Les définitions présentées par nos deux conférenciers se situent sur deux plans différents. Elles ne sont donc pas comparables. M. Dumas s'est attaché à cataloguer les définitions existantes. Celles-ci sont extrêmement nombreuses et se réfèrent à des contenus très divers. Il est évidemment intéressant d'en établir la liste. Mais je ne crois pas que ce soit le véritable but de notre réunion d'aujourd'hui. Le voudrions-nous, que le temps et la préparation nous manqueraient pour y parvenir. En effet, le terme « statistique » a déjà une longue histoire au cours de laquelle ses diverses acceptions ont considérablement évolué. C'est pourquoi il me semble préférable de concentrer notre attention sur celui des sens du terme qui désigne la méthode scientifique de recherche applicable à l'étude des phénomènes nombreux subissant une multitude d'influences qu'il est impossible de séparer expérimentalement.

(1) Notons en passant que l'édition 1956 du Nouveau Petit Larousse illustré donne du substantif « statistique » les définitions suivantes :

« Science qui a pour objet le groupement méthodique des faits qui se prêtent à une évaluation numérique (impôts, recrutement, condamnations, productions industrielles et agricoles, population, religion, etc...); tableau numérique d'un fait se prêtant à la statistique : statistique de la natalité ».

On retrouve bien là les deux acceptions fondamentales indiquées ci-dessus.

La définition proposée par M. Prot tend effectivement à répondre au but ainsi défini. Elle me semble très intéressante mais — à mon sens — elle n'est pas complète : elle n'insiste pas sur le caractère « nombreux » des phénomènes auxquels la statistique s'applique et qui constitue l'attribut essentiel de cette discipline.

Après avoir entendu l'intervention de M. Chevry, je pense que la définition de M. Dugé de Bernonville résume dans une formule remarquablement concise et élégante toutes les caractéristiques essentielles de la statistique. En particulier, le terme « collectivités » qu'elle contient exprime parfaitement le caractère nombreux des phénomènes étudiés par la statistique.

REMARQUES DE M. PENGLAOU. — N'ayant pas pu, à mon vif regret, assister à la séance de juin dernier, j'ai demandé à notre Secrétaire général de me communiquer les épreuves du compte rendu dont j'ai pris connaissance avec beaucoup d'intérêt.

Ayant maintes fois pris position sur la question au cours des précédentes années, et notamment par des communications à notre Société, je serai bref. Je n'indiquerai sommairement que ce qui ne me paraît pas être « statistique » et, ces éliminations faites, ce qu'il convient de retenir pour préciser l'acception du mot « statistique ».

a) *Ce qui doit être éliminé*

La première version de M. Prot : « tout dénombrement d'êtres, d'objets, de faits ou de mesures constituant les espèces d'un genre », doit être rigoureusement repoussée. Cette définition est loin d'être spécifique. En effet, l'admettre conduirait à estimer d'ordre statistique toute recherche scientifique ou technique comportant détection, sélection et comptage des données. De telles opérations, scientifiques ou techniques, qui ont pour but d'opérer un choix et une mesure d'unités disparates dans un réel donné sont communes à toutes les disciplines. Rechercher les éléments préalablement caractérisés comporterait-il un caractère statistique? Tous les savants, tous les techniciens seraient-ils des statisticiens qui s'ignorent? Opérer des sélections dans une masse de faits, de prime abord non différenciés, cela entraînerait-il l'application de procédés et de méthodes statistiques? Je n'y puis consentir, car ce serait ainsi conférer à n'importe quelle démarche un caractère abusif de statistique. Rien ne me paraît plus dangereux que ces généralisations outrancières, dont on a malheureusement trop d'exemples et qui faussent souvent les conclusions de la recherche.

Je sais bien que le sens commun attribue à tout tableau de chiffres assemblés et qualifiés l'appellation de « statistiques », l'exemple de l'indicateur de chemin de fer avancé pertinemment par M. Dumontier est caractéristique. Mais c'est un abus de langage contre lequel on ne saurait trop protester.

La seconde définition de M. Prot est non moins amphibologique. « Transformation de leur expression », cela doit vouloir signifier la quantification des données retenues par sélection ou dénombrement dans la définition A. L'enfant qui, à l'école, compte le nombre de bâtons inscrits sur le tableau noir serait-il, lui aussi, un statisticien précoce? Ce serait puéril de le penser!

Quant à la définition C, il me semble qu'elle ne correspond qu'à une vague conception de la méthode statistique.

Des propositions de M. Dumas, j'estime qu'il faut tout d'abord écarter la notion de science, quelque déplaisir que cela puisse constituer pour un statisticien. Sans doute le mot « art » convient-il mieux; je préfère pour ma part les mots « discipline », ou « méthode », ou encore « technique ».

J'attache une particulière importance à l'exposé de M. Marczewski, encore que je pense qu'elle est, elle, un peu trop exclusive. Écarter de la recherche statistique les phénomènes qu'il est possible de séparer expérimentalement — c'est ce qu'on peut déduire *a contrario* de sa déclaration — c'est ce qui me semble par trop restrictif. De nombreux travaux, statistiques de l'aveu de tous, portent sur des phénomènes qu'on a pu compter un à un : recensements démographiques, expériences de physique sur les unités matérielles, etc. Sans doute au laboratoire emploie-t-on des procédés statistiques pour déterminer le nombre de molécules contenues dans un récipient, alors que leur comptage est une impossibilité. Mais c'est un procédé qui ne doit pas faire exclure les méthodes de dénombrement un à un quand elles sont applicables. Au demeurant, ce qu'il faut retenir dans la déclaration de M. Marczewski, c'est avant tout le rôle qu'il dévolue à la notion de masse ou de collectivité. Nous affleurons ainsi à l'une des caractéristiques de la méthode statistique.

b) *Ce qu'il faut retenir*

Il reste que la statistique s'applique spécifiquement à établir des relations numériques entre des phénomènes considérés comme faisant partie d'une collectivité, c'est-à-dire en général nombreux; qu'elle est en somme une application de procédés mathématiques à des faits ou à des groupes de faits nombreux de même nature — ou, et j'ajouterai à l'excellente définition de M. Dugé de Bernonville, qui a été justement citée, à l'établissement de rapports, de corrélations entre ces faits ou groupes de faits. Les traités classiques montrent à l'évidence que c'est bien là le *quid proprium* de la méthode statistique et de la statistique tout court.

Ne recourons pas au *Nouveau Petit Larousse* ou à quelque Littré, même renouvelé. Le sens commun déforme en vulgarisant.

Je souhaite, en terminant, qu'un travail exhaustif soit entrepris, en attirant l'attention de notre Société sur l'opportunité de la publication d'un vocabulaire statistique (les vocabulaires parus au cours de ces dernières années ont eu un succès mérité et une audience étendue) qui établirait mieux que des dissertations les caractères spécifiques de la statistique.
