

Chapitre 1

Histoire de la représentativité statistique : quand le politique refait toujours surface

Emmanuel Didier

Introduction

La notion de représentativité est une émanation de la politique et des démocraties représentatives. Bernard Manin montre, par exemple, dans ses *Principes du gouvernement représentatif* (1995), qu'elle est apparue en même temps que la démocratie athénienne, et il ajoute qu'elle a toujours eu une multitude de sens, comme le pluriel de son titre l'indique. À la fin du XIX^e siècle, les statisticiens l'ont redéfinie et ont commencé à utiliser des expressions comme « statistiques représentatives » ou « échantillons représentatifs ». Ces expressions acquièrent leur caractère d'évidence durant la grande dépression des années 1930, d'abord aux États-Unis, puis dans le reste du monde.

Nous allons revenir sur le processus par lequel la représentativité a été définie. Nous verrons que les efforts effectués par les statisticiens et les probabilistes ont visé non pas à donner vie à une notion qui n'aurait pas existé, mais au contraire à restreindre le nombre de significations qu'elle recouvrait et à en épurer sa trop grande polysémie, en particulier politique. L'effort a moins consisté à ajouter un être dans l'arène statistique qu'à en soustraire tous ceux qui s'y invitaient indûment. Comme notre titre l'indique, il n'est pas certain que ces efforts n'aient jamais été couronnés de succès.

La genèse de cette évolution se situe aux États-Unis, connus pour être précurseurs en la matière. Nous progresserons en quatre temps, en démontrant comment, initialement, la représentativité possédait un très grand nombre de sens en politique comme en statistique, puis dans quelle mesure les efforts successifs pour unifier ces sens, pour rendre le mot monosémique ont, jusqu'ici, toujours échoué.

1. La représentation statistique

Dès le milieu du XIX^e siècle, c'est d'abord à la notion de *représentant* que les statisticiens ont eu recours. Parmi eux, les spécialistes de l'agriculture ont été les

premiers à avoir eu besoin d'informateurs locaux capables de quantifier la production agricole. Ces informateurs étaient « représentatifs », mais dans un sens complètement différent de celui que nous connaissons aujourd'hui, et directement influencé par la pratique démocratique. Analysons comment cette articulation entre politique et statistique a été initialement opérée.

À cette époque la question de l'asymétrie d'information entre les fermiers, producteurs de biens agricoles, et les acheteurs de ces biens, se posait déjà avec acuité. Les fermiers, que leur tâche quotidienne attachait à leur ferme et à leur entourage immédiat, avaient une connaissance des volumes produits globalement bien inférieure à celle des acheteurs, qui parcouraient de vastes étendues en allant de ferme en ferme. Le fermier connaissait les résultats locaux, alors que l'acheteur avait une vue d'ensemble sur toute une région et pouvait, en conséquence, spéculer sur les prix. Il a donc été décidé, pour pallier ce problème, de produire et de rendre publics les chiffres de la production agricole pour chaque région. Il suffisait au fermier de consulter ces données pour connaître les tendances du marché et vendre sa production au juste prix.

En 1862, ce sont les fermiers eux-mêmes, par le biais d'un journal, *The American Agriculturist*, qui décidèrent de « choisir dans chaque village un homme fiable pour la qualité de ses jugements et sa capacité générale à estimer avec une grande précision [la condition] des cultures les plus importantes ». Cet homme envoyait le résultat de ses propres enquêtes au journal, qui les publiait sous forme d'un tableau. Les fermiers, concernés par les résultats de l'enquête, choisissaient donc eux-mêmes celui qu'ils jugeaient digne de cette tâche. Une méthode largement inspirée par le scrutin démocratique.

En 1863, le ministère de l'Agriculture américain fut établi et reprit immédiatement cette tâche à sa charge. Il a alors peu à peu élargi l'ensemble des personnes susceptibles d'être représentantes, mais sans rupture nette avec le modèle de l'élection. Le ministère s'adressait à des correspondants volontaires, proches du monde agricole – soit parce qu'ils étaient eux-mêmes fermiers, soit parce qu'ils vivaient parmi eux – et capables, non seulement de mener des enquêtes sur la production de différentes cultures, mais aussi de synthétiser et de quantifier leurs résultats à partir de questionnaires préétablis. Parmi ces correspondants, on trouvait des notables locaux, des personnes désignées par les élus fédéraux et mêmes des élus. La notion de représentation politique était, elle aussi, susceptible d'être interprétée de mille façons, toutes convenables. L'objectif était de pouvoir obtenir un représentant par village, de façon que la somme de toutes les synthèses locales couvre la totalité du territoire, but qui fut largement atteint dès les années 1920 (Taylor et Taylor, 1952).

À la même époque, d'autres pays ont eu recours à des méthodes consistant à désigner des représentants statistiques répartis sur tout le territoire, en particulier, la Russie (Mespoulet, 2001) et la Prusse, où Max Weber lui-même mena des enquêtes sur l'agriculture (Pollak, 1986). Ainsi, avec la double contrainte de devoir récolter des informations à la fois sur un territoire très vaste et à un rythme assez rapide ne permettant pas de mettre en œuvre de véritables recensements, la statistique a permis de rendre possible le contrôle et le suivi de la croissance agricole en réutilisant un modèle de désignation des représentants, proche de la

démocratie ou inspiré par elle. On n'utilisait alors le terme de représentant, pas encore celui de représentativité.

2. La défense de la représentativité en statistique

La représentativité statistique apparaît comme telle à la toute fin du XIX^e siècle, en 1895, défendue par le directeur de la statistique nationale norvégienne, Anders Kiær, lors des sessions de l'Institut international de statistique (nous présentons ici les discussions reprises dans Kiær 1895, 1899, 1903, 1905, 1976). Kiær ne fait pas le lien avec les représentants décrits ci-dessus, mais cherche à démontrer la nécessité d'une troisième voie par rapport aux deux autres méthodes légitimées à l'époque : le recensement et les monographies. Comme l'explique Alain Desrosières (1993, chapitre 7), cette évolution prend forme avec l'apparition de l'État-providence, c'est-à-dire une nouvelle façon de penser la totalité de la Nation et son rapport à ses parties (l'Etat comme assureur des groupes sociaux), que les débats statistiques sur le lien entre l'échantillon et l'univers prennent corps et forme. Voyons comment Kiær introduit et justifie ce qu'il appelait « la méthode représentative ».

Kiær choisit l'Institut international de statistique, considéré alors comme l'une des instances les plus légitimes dans son domaine, pour présenter ses travaux. Fondé par Adolphe Quetelet en 1853, il regroupait les principaux statisticiens publics des pays occidentaux, au cours de sessions extraordinaires biennuelles, et constituait un lieu d'échanges et de mutualisation du savoir et des pratiques.

Entre 1895 et 1905, le directeur de la statistique norvégienne présenta plusieurs communications dans lesquelles il défendait sa « méthode représentative ». Il commença par présenter les deux méthodes canoniques, à savoir, le recensement et la monographie, qui donnaient (faussement à son avis) l'impression de répondre à l'ensemble des questions adressées aux statisticiens, puis il introduisit le coin de la représentativité dans cette alternative.

ou plutôt
"l'angle de la
représentati-
vité ?

Selon lui, le recensement avait comme avantage d'être exhaustif – c'est-à-dire qu'il renseignait sur la totalité d'un pays ou d'une région mais, comme il s'adressait à toute personne, de toute condition, les questions qu'il posait ne pouvaient être que brèves et simples. Pour Kiær, le recensement ne donnait donc à voir que le squelette d'une nation.

Il pouvait être complété par des monographies, dont le grand représentant était le Français Frédéric Le Play. Il s'agissait de choisir un type social – on connaît le travail de Le Play sur les différents types de famille et sa monographie de la famille Mélouga des Pyrénées – et de l'étudier longuement, sous le plus de rapports possibles, en utilisant prioritairement leur budget de famille. La monographie avait l'avantage de donner de la *chair* à l'objet étudié, mais ne permettait pas de le replacer dans la variété des types, chose rendue possible par le seul recensement.

Mais pour Kiær, cette métaphore du *squelette* et de la *chair* était trompeuse, de nombreuses questions ne pouvant être traitées ni par l'une ni par l'autre, ni

même en combinant les deux méthodes. Il réfléchissait en particulier à la création d'une caisse de secours et d'assurance destinée aux ouvriers de son pays. Pour ce faire, il avait besoin de données à la fois précises et complexes – en particulier sur les périodes de chômage, qui ne pouvaient être traitées que par de nombreuses questions échappant largement aux limites du recensement – et en même temps, il lui fallait savoir dans quels rapports proportionnels entraient les différentes professions du royaume.

La seule solution qu'il entrevoyait était la « méthode représentative » qu'il avait mise au point avec son directeur adjoint Jakob Mohn, oublié des manuels d'histoire, mais dont l'importance a été démontrée par notre collègue Einar Lie (2002). Cette méthode consistait à utiliser un questionnaire, technique relevant du recensement, mais de ne questionner qu'une petite partie de la population, ce qui laissait ainsi le temps matériel de traiter un questionnaire complexe, relevant de la monographie. Bien qu'il ait relevé l'utilité économique d'une telle méthode, moins coûteuse qu'un recensement, c'est plutôt qu'elle donnait accès à tout un ensemble de faits, échappant aux deux autres méthodes, qui était mis en avant dans son argumentation.

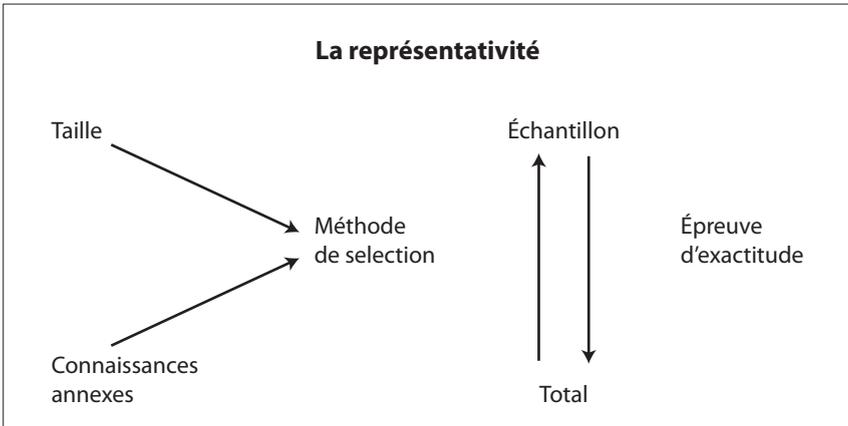
Cette méthode fut critiquée par Georg Von Mayr, directeur de l'institut de Prusse qui la trouvait inductive. Il ne considérait pas comme légitime de généraliser un argument à une totalité lorsque seule une de ses parties était observée. Von Mayr reconnaissait volontiers que cette méthode était utilisée en pratique, faisant probablement allusion aux enquêtes menées avec des représentants ou encore au fait qu'un recensement n'était jamais complètement exhaustif, mais il lui semblait peu concevable de laisser un organe – comme l'Institut international de statistique – institutionnaliser ce qui n'était, au fond, qu'un pis aller. Il lui semblait impossible de lui conférer un droit de « bourgeoisie » (pour reprendre l'excellente expression d'Edmund Wilhelm Milliet, représentant de la statistique suisse) comparable à celui des recensements et monographies.

Kiær lui accorde qu'il n'est pas légitime de généraliser de la partie au tout, mais il complète : sauf si cette partie est construite de telle façon qu'elle soit *représentative*. Lorsqu'on goûte une soupe, commente-t-il, on se fie à la petite cuillère que l'on ingère car on la pense *représentative* de toute la marmite. La question n'est donc pas de savoir si, oui ou non, l'Institut international de statistique doit autoriser les études représentatives, ce qui lui semble acquis, mais de déterminer les méthodes qui permettent de produire des échantillons *représentatifs*.

Or, pour Kiær, celles-ci sont infinies. Il en propose lui-même un certain nombre qu'il a mis en œuvre dans son pays, et qui reposent toutes sur des connaissances annexes portant, pour la plupart, sur des découpages géographiques du territoire. Il utilise plusieurs concepts qui lui permettent d'argumenter pour la représentativité. En particulier, il parle de bonne « couverture » du territoire en montrant une carte des localités sélectionnées. Il se réfère aussi à ce que nous appelons aujourd'hui une sélection systématique de l'échantillon (également dans un univers de lieux). Au cours de ces débats, Lucien March, alors directeur de la Statistique générale de la France, rappelle l'existence des travaux de Laplace qui avait utilisé une méthode baptisée la méthode du « coefficient multiplicateur », pour estimer la taille de la population du royaume et qui s'était servi du tirage

aléatoire pour estimer « l'erreur à craindre » dans un tel cas (Bru 1988). À vrai dire, March ne fait pas preuve d'une maîtrise très poussée des raisonnements laplaciens, mais cherche à rappeler que les Français ont été les premiers à l'utiliser. Ce qui, finalement, ressort de ces discussions au sein de l'Institut international de statistique est qu'il existe un nombre important de méthodes permettant de construire un échantillon représentatif, et qu'elles ont en commun la caractéristique d'utiliser des connaissances annexes.

Kiær ne se contente pas d'étudier les méthodes de sélection des échantillons pour établir leur représentativité. Il utilise ce que l'on peut appeler des épreuves d'exactitude, c'est-à-dire des comparaisons *ex post*, une fois l'échantillon observé, avec d'autres données comme celles du recensement. Par exemple, il construit un échantillon à partir d'un découpage géographique et compare ensuite sa structure à celle du recensement en termes d'activité professionnelle. Mise à part une catégorie qui lui semble mal représentée, il lui semble que la proximité entre les pourcentages marginaux permet d'affirmer que son échantillon est représentatif. Le schéma suivant synthétise la conception de la représentativité de Kiær.



La représentativité est visée par des méthodes de sélection qui permettent de passer d'une totalité à un échantillon (méthodes qui dépendent elles-mêmes en partie de la taille de la population et des connaissances annexes dont on dispose sur cette population). Dans le sens inverse, on peut constater ou vérifier la représentativité d'un échantillon en lui faisant subir une épreuve d'exactitude – une comparaison entre des résultats observés sur l'échantillon et d'autres statistiques.

Malheureusement, cette défense et illustration de la représentativité au niveau international se sont achevées pour Kiær par une défaite au niveau national. Les actuaires de son pays sont parvenus à lui ravir la responsabilité de la création des caisses d'assurances, pour lesquelles leurs techniques semblaient plus adaptées. Suite à cette défaite, Kiær disparaît des sessions de l'Institut international de statistique (Lie, 2002) qui cessa du même coup de traiter de cette question pendant un temps. **Si Kiær avait pu présenter son concept, les aléas de la vie donnèrent raison à Von Mayr et la représentativité disparut brutalement des discussions de la vénérable institution. Elle ne fut pas entièrement oubliée pour autant.**

il doit
manquer
un bout de
phrase : à
compléter

3. Tentative d'établissement d'une seule méthode représentative par les probabilistes

Alors qu'au début du xx^e siècle, l'Institut international de statistique se désintéressait de la question de la représentativité, des statisticiens différents des précédents, car davantage versés dans les mathématiques, s'en saisirent et approfondirent le concept pour, au milieu des années 1930 en proposer une définition unique et qu'ils espéraient stable. Deux statisticiens ont joué un rôle capital dans cette histoire : Arthur Bowley, puis Jerzy Neyman.

Arthur Bowley, de formation mathématicienne, enseignait à la London School of Economics depuis sa fondation en 1895. Ses recherches économiques portaient sur de nombreux thèmes et ceux qui nous intéressent ici traitaient de la pauvreté à Londres et au Royaume-Uni. Pour mener une étude qu'il publia en 1906, il recourut à une enquête partielle (il lui était impossible de *recenser* les pauvres). Or, en mathématicien qu'il était, il voulut recourir à l'erreur probable pour estimer la précision des mesures effectuées sur son échantillon (cette notion n'était pas révolutionnaire : Yule l'utilise par exemple dans son manuel publié en 1910). L'erreur probable étant, comme son nom l'indique, une notion probabiliste, Bowley dut introduire de l'aléa dans son enquête en tirant aléatoirement les individus de son échantillon dans des listes communales (que ce soit des listes d'adresses ou des listes d'usagers des services municipaux). Grâce au tirage au sort, Bowley pouvait estimer l'erreur commise par son échantillon. Il prolongea ses travaux avec sa collaboratrice Margaret Hogg avec laquelle il publia un livre (Bowley et Hogg, 1925) dont une partie entière, consacrée à l'échantillonnage, défendait très vigoureusement le tirage aléatoire.

Notons que Hogg partit ensuite aux États-Unis où elle joua un rôle capital, bien que totalement oublié du fait de sa mort précoce, dans le développement des sondages aléatoires (Didier, 2009, p. 263-272). Elle a été l'une des plus vaillantes défenseuses de la sélection aléatoire des échantillons (bien avant Neyman qui en retirera toute la gloire). L'histoire de la représentativité est jalonnée de personnages dont la notoriété fut gommée par de plus célèbres qu'eux.

Mais si Bowley parle de la précision de la mesure et, plus rarement, de la précision de la statistique, il n'évoque jamais la représentativité. La liaison entre ses instruments (de précision) et la notion défendue par Kiær eut lieu à l'endroit même où cette dernière avait été présentée au monde, à savoir, l'Institut international de statistique, mais des années plus tard, en 1925. En effet, après une éclipse de vingt ans, une session fut organisée à propos de la représentativité, sous l'impulsion de Bowley et de Jensen, directeur de la statistique danoise .

Kiær est mort en 1919, mais s'il avait pu y assister, il aurait été heureux de constater que la possibilité de construire des échantillons représentatifs n'était plus remise en cause, mais que la discussion portait entièrement sur le point de savoir comment les construire. La conclusion à laquelle les statisticiens parvinrent consista à légitimer non plus un grand nombre comme au tournant du siècle, mais seulement *deux* méthodes alternatives de construction d'échantillons aléatoires. La résolution de l'Institut est présentée *in extenso* dans l'encadré 1.

Cette résolution légitime définitivement l'usage de la représentativité en statistique et insiste sur le fait que deux méthodes principales doivent être approfondies. Rappelons-nous que Kïær pensait que les méthodes pouvaient être en nombre infini. En 1925, la notion de représentativité est donc précisée et clarifiée.

On ne s'étonne pas que Bowley, qui travaillait depuis 20 ans sur l'aléa, ait insisté pour que cette méthode soit retenue. On ne s'étonnera pas davantage que la méthode, alors appelée celle du « choix judicieux », le fût aussi. Il s'agissait, pour

Encadré 1. L'Institut international de statistiques

I. Attire l'attention sur les avantages très considérables qui peuvent être obtenus en appliquant la méthode représentative dans les conditions suivantes :

– Les résultats d'une investigation partielle ne peuvent être généralisés que si l'échantillon utilisé est par nature suffisamment représentatif de la totalité. Pour ceci, l'échantillon peut être sélectionné de différentes façons ; les deux principaux cas suivants peuvent cependant être distingués :

A) Sélection aléatoire : un certain nombre d'unités sont sélectionnées de façon à ce que l'exacte égalité de chance d'inclusion soit la règle dominante.

B) Sélection par choix judicieux : un certain nombre de groupes d'unités sont sélectionnés qui, ensemble, fournissent presque les mêmes caractéristiques que la totalité. Pour avoir connaissance de la précision des estimateurs, il est nécessaire que suffisamment de groupes soient inclus de façon à permettre la mesure des variations entre les caractères des groupes.

II. Recommande que les investigations soient menées de façon à permettre le traitement mathématique de la précision des résultats, et avec ces résultats une indication de la taille de l'erreur qu'ils sont susceptibles de faire commettre doit être indiquée.

III. Répète le souhait que dans les rapports donnant les résultats de toute investigation représentative, une description explicite et détaillée de la méthode adoptée pour sélectionner l'échantillon soit donnée*.

* "The International Institute of statistics I. Calls attention to the very considerable advantages which can be obtained by applying the Representative Method under the following conditions. The results of a partial investigation should only be generalized provided that the sample used is in its nature sufficiently representative of the totality. In such respects the sample may be selected in different ways; the following two main cases, however are to be distinguished: A- Random selection : a number of units are selected in such a way that exact equality of chance of inclusion is the dominant rule. B- Purposive selection: a number of groups of units which together yield nearly the same characteristics as the totality. In order to have any knowledge of the precision of the estimates, it is necessary that sufficient groups should be included to allow the variations between the characteristics of the groups to be measured. II. Recommends that the investigation should be so arranged as to allow of a mathematical treatment of the precision of the results, and that with these results should be given an indication of the extent of the error to which they are liable. III. Repeats the wish that in the reports of the results of every representative investigation an explicit account in detail of the method of selecting the sample adopted should be given." (Cité par Seng, *op. cit.*, p. 223.)

utiliser un anachronisme, de construire son échantillon d'une façon comparable à ce que nous appelons aujourd'hui les quotas, à la différence près que les variables géographiques (le lieu de résidence des enquêtés) jouaient un rôle fondamental par rapport aux autres variables ; l'échantillon était d'abord construit à partir de découpages spatiaux, puis à partir d'éventuelles autres variables. Or cette méthode avait été utilisée à de nombreuses reprises, et avec un succès particulièrement remarqué par les statisticiens italiens regroupés autour de Corrado Gini (Prévost, 2009). Par exemple, ils avaient dû faire de la place dans les archives des recensements pour pouvoir accueillir les questionnaires de la passation suivante, et avaient conservé un échantillon représentatif, au sens du choix judicieux, des précédents questionnaires.

Respectant les deux traditions, anglaise et continentale, l'Institut international de statistique avait encouragé deux méthodes différentes, sans trancher entre les deux. Ce n'est qu'une dizaine d'années plus tard, en 1934, que Jerzy Neyman, un Polonais alors émigré au Royaume-Uni travaillant dans le laboratoire des Pearson, écrivit un article dans le *Journal of the Royal Statistical Society* dans lequel il prétendait unifier toutes les méthodes représentatives. Cet article, qui rencontra un succès immense, est aujourd'hui considéré comme le fondement de la théorie probabiliste moderne de l'échantillonnage (Neyman, 1934).

Bien que le titre même de son article soit évocateur (« Sur les deux aspects différents de la méthode représentative: la méthode d'échantillonnage stratifié et la méthode par choix judicieux »), on oublie trop souvent que l'argument de Neyman porte sur la représentativité et sur le fait qu'elle ait, à son époque, deux sens. L'objectif de son auteur était de montrer que l'un des deux sens défini par l'Institut international de statistique – le choix judicieux – n'était qu'une version de l'autre, et une version non optimale. De sorte que la représentativité pouvait être définie, pouvait-il conclure, de façon univoque.

Pour parvenir à ce résultat, il mit au point la notion, aujourd'hui célébrissime, d'intervalle de confiance, qui n'est rien d'autre qu'un exemple – particulièrement réussi et formulé à l'aide de la théorie des probabilités – de ce que nous avons appelé ci-dessus une épreuve d'exactitude. Insistons sur le fait, assez contre intuitif aujourd'hui, que l'intervalle de confiance a été inventé dans le but de définir la représentativité, et non l'inverse ; il n'était alors qu'instrumental. L'intelligence du raisonnement de Neyman réside dans le fait qu'il est entièrement *déductif*. L'intervalle de confiance a pour vertu d'éviter au statisticien toute inférence de l'échantillon à la totalité, il permet de *déduire* des propositions sur la population globale à partir d'observations sur l'échantillon. Mais en quoi ces intervalles de confiance permettent-ils de parler de représentativité ?

Neyman explique qu'un échantillon est de qualité quand son intervalle de confiance est le plus étroit possible. Or, montre-t-il, pour cela, il ne faut pas le sélectionner de façon proportionnelle à la moyenne des strates, mais à la variance des strates – deuxième découverte importante. Autrement dit, si on a une population composée pour moitié de grands et pour moitié de petits individus, et si l'on sait qu'à une question tous les grands ont répondu exactement la même chose, tandis que les petits ont donné des réponses très variées, alors il ne faut pas faire un échantillon composé d'une moitié de grands et d'une moitié de pe-

tits, mais (à la limite) un seul grand (qui nous donnera l'information pour sa strate) et tout le reste de petits. Pour construire un bon échantillon, c'est moins la taille de la strate dans la population totale qui compte que les variations internes qui la composent.

Il ne reste plus alors qu'un pas à faire pour montrer que les échantillons par choix judicieux ne sont rien d'autre que des échantillonnages proportionnels à la moyenne de la strate, c'est-à-dire non optimaux, et qu'il faut donc les rejeter pour construire les échantillons de l'autre façon, c'est-à-dire aléatoire et proportionnels à la dispersion. En conclusion, Neyman se demande « s'il est possible de donner un sens précis aux mots "une méthode généralement représentative" ». Il répond clairement : « Je pense que oui⁽¹⁾. »

Si nous nous intéressons à un caractère collectif X d'une population π et utilisons des méthodes d'échantillonnage et d'estimation nous permettant d'attribuer à chaque échantillon Σ possible un intervalle de confiance $X1(\Sigma), X2(\Sigma)$ tel que la fréquence d'erreur dans l'affirmation que

$$X1(\Sigma) \leq X \leq X2(\Sigma)$$

ne dépasse pas la limite $1 - e$ déterminée à l'avance, quelles que soient les propriétés de la population, j'appelle cette méthode d'échantillonnage représentative. [...] Par conséquent, la méthode d'échantillonnage aléatoire [...] peut être dite représentative dans le sens du mot que j'utilise⁽²⁾.

Cette définition, qui associe tout simplement l'attribution d'un intervalle de confiance à la représentativité, ne bouleverse pas entièrement les résultats auxquels les prédécesseurs de Neyman étaient parvenus, puisque le schéma en aller-retour entre l'univers et l'échantillon, exposé dès les premiers débats, leur est commun. Pour Neyman aussi, la représentativité oscille entre une épreuve d'exactitude et une procédure de sélection. Cependant, l'ordre dans lequel se fait cet aller-retour est inversé. Le point principal de Neyman est que l'échantillon doit être construit de façon qu'il soit possible de lui associer un intervalle de confiance, c'est-à-dire qu'il puisse passer l'épreuve d'exactitude. Comme cette épreuve requiert de l'aléa, seuls sont représentatifs les échantillons aléatoires. La méthode de sélection est au service de l'épreuve qui détermine ses caractéristiques.

Ainsi, Neyman était alors convaincu d'avoir levé toutes les ambiguïtés de la notion et établi une définition unique de la représentativité, qui pouvait ainsi entrer de plein droit dans l'attirail des outils statistiques. Mais deux années plus tard, en 1936, l'histoire est venue compliquer ce beau mouvement de clarification.

(1) "The question whether it is possible to give any precise sense to the words 'a generally representative method'. I think it is" (p. 585).

(2) "If we are interested in a collective character X of a population π and use methods of sampling and of estimation, allowing us to ascribe to every possible sample Σ a confidence interval $X1(\Sigma), X2(\Sigma)$ such that the frequency of errors in the statements $X1(\Sigma) \leq X \leq X2(\Sigma)$ does not exceed the limit $1 - e$ prescribed in advance, whatever the unknown properties of the population, I should call the method of sampling representative. [...] Thus, the method of random [...] sampling may be called a representative method in the sense of the word I am using." (p. 585-6)

4. Les sondages d'opinion et les quotas

Pour suivre l'histoire de la représentativité, il convient de se replacer dans le contexte de la campagne présidentielle américaine de 1936, qui ne fut pas seulement l'occasion pour Roosevelt de présenter le bilan de son premier New Deal au peuple américain, mais aussi celle où, pour la première fois, sont apparus les sondages d'opinion, **apportant avec eux** une définition de la représentativité qui était très différente de celle de Neyman, et qui rencontra pourtant un important succès – au point qu'elle servit d'argument pour écarter les techniques statistiques utilisées jusqu'alors dans les prévisions électorales.

Aux États-Unis, depuis le début du siècle, les campagnes électorales étaient accompagnées de *straw polls*, des « votes de pailles », c'est-à-dire de grandes opérations au cours desquelles les journaux demandaient à leurs lecteurs de pré-jouer l'élection. Chaque lecteur était invité à renvoyer au journal un bulletin en faveur de son camp et le décompte permettait d'établir un rapport de force numérique entre les camps. C'était un simulacre d'élection, une « élection de paille ».

En 1936, trois instituts, ceux de George Gallup, Elmo Roper et Archibald Crossley remirent en cause cette méthode en faisant remarquer que la population des participants à ces « votes de paille » n'était en aucune façon *représentative* de la population des électeurs. Il leur semblait, au contraire, qu'elle avait toutes les chances de ne pas l'être. Ils présentèrent alors une méthode alternative appelée la « méthode des *quotas* » et qui permettrait, assuraient-ils, de construire une sorte de *miniature* de l'opinion américaine. Selon eux, la représentativité émanait de cette miniaturisation.

Pour comprendre comment réaliser une « Amérique en miniature », il faut insister sur le fait que la culture et les ressources intellectuelles de ces sondeurs étaient influencées par la science politique américaine. Ils se référaient en particulier aux travaux du philosophe James Bryce sur le rapport de l'opinion publique à la démocratie. Or l'opinion, selon eux, ne se modelait pas de n'importe quelle manière. Selon Gallup par exemple :

Le public américain est une mosaïque dont le patron complexe d'individus est divisé en une série de groupes sociaux et d'associations. Les individus ne vivent pas dans le vide. Des considérations de géographie, d'occupation, d'âge, de sexe, d'affiliation politique, de race, de religion et de culture sont les déterminants élémentaires de leurs expériences et de leurs opinions. Le jeu de telles forces divise le public en groupes et stimule des attitudes communes à l'intérieur des groupes eux-mêmes (traduction de l'auteur)⁽³⁾.

Ce qui comptait donc avant tout dans cette démarche de « miniaturisation » était de trouver des variables – évoquées dans ce paragraphe par la phrase sur les

(3) "The american public is a mosaic whose complex pattern of individuals is clustered together in a variety of social groups and associations. Individuals do not dwell in a vacuum. Considerations of geography, occupation, age, sex, political affiliation, race, religion and the cultural background are the basic determinants of their experiences and opinion. The play of such forces divides the public into groups and stimulates common attitudes within the groups themselves." (Gallup, 1940, p. 60.)

considérations – permettant de stratifier la population et de construire ainsi une matrice de l'Amérique « comme un tout » (une expression qui apparaît à cette époque). Les quotas permettent d'importer la notion statistique de strate vers une définition de l'opinion publique.

Dès lors, comment sélectionner les individus qui composeraient ces quotas ? Dans *The Pulse of Democracy*, Gallup commence par présenter le tirage aléatoire – qui lui a probablement été montré par Wilks (1951), alors professeur de statistiques à l'université de Princeton –, dont il reprend l'un des avantages sur lequel les acteurs ont insisté, de façon quasi invariable, pendant toute l'histoire que nous avons racontée : l'aléa permet de « calculer la taille de l'échantillon nécessaire pour assurer sa précision dans certaines limites⁽⁴⁾ ». L'avantage de l'aléatoire réside dans le fait de mesurer la précision des estimateurs. Gallup reprend des exemples alors bien connus : si, dans une urne qui contient 1 000 billes blanches et 1 000 billes noires, on tire 100 billes, on obtiendra des proportions de blanches et de noires qui seront proches de 50 % dans une limite que l'on peut calculer. Pour Gallup, l'un des intérêts du tirage aléatoire est qu'il permet de mesurer la précision des échantillons.

Mais cette défense n'est qu'apparente, car dans la pratique, les échantillons qu'il utilisait n'étaient pas tirés aléatoirement. Ce point nous a été confirmé au cours d'un entretien avec Frederick Mosteller en 1997. Au moment où nous l'avons rencontré, ce grand statisticien était professeur de statistiques émérite à l'université de Harvard⁽⁵⁾. Il a commencé à travailler pour l'institut Gallup à partir de 1938 comme consultant statistique et a continué durant toute la guerre. Il nous a expliqué que les premiers échantillons fournis par Gallup pour analyse n'étaient pas aléatoires, ce qui n'était pas un secret, mais que les traitements devaient se faire en les considérant comme tels. C'est-à-dire qu'il lui fallait faire l'hypothèse – qu'il savait fautive – que les échantillons étaient tirés aléatoirement et calculer des intervalles de confiance. Ainsi, avant la guerre, les membres de l'Institut Gallup et les universitaires supposaient que leurs échantillons avaient été tirés aléatoirement, sachant que cette hypothèse n'était pas vérifiée.

Le fait de ne pas procéder à des tirages aléatoires facilitait grandement l'entreprise car, comme nous l'avons montré par ailleurs, la construction d'une base de sondage exhaustive soulevait des difficultés considérables – surtout aux États-Unis, où l'état civil et les numéros de sécurité sociale étaient encore inexistantes (Didier, 2009). D'autre part, l'inconvénient de ne pas recourir à l'aléa était faible précisément pour la raison que la définition de la représentativité ne reposait en rien sur celle qu'avait proposée Neyman mais sur l'idée de *miniature*. Pour les sondeurs politiques qui raisonnent en termes d'opinion publique, l'échantillon se comporte électoralement comme la population, s'il en constitue une *miniature* au regard d'un certain nombre de variables décrivant les conditions de vie. La simplification de la procédure consistant à ne pas tirer aléatoirement l'échantillon, capitale pour les probabilistes, paraissait totalement secondaire pour les spécialistes de l'opinion publique.

(4) "To calculate the size of the sample needed to secure accuracy within certain limits" (Gallup, 1940, p. 58).

(5) Entretien du 11 mars 1997.

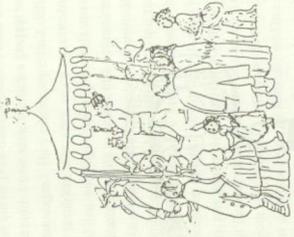
Or, comme l'ont montré Loïc Blondiaux (1998) et Sarah Igo (2007), il se trouve que, les spécialistes de l'opinion publique rencontrèrent un immense succès, notamment dans les universités, auprès de professeurs comme Hadley Cantril ou Samuel Stouffer, qui connaissaient leurs méthodes et les enseignaient. Ils conquièrent la presse qui publia leurs résultats et de nombreux hommes politiques commencèrent à recourir à leurs méthodes, dont Roosevelt lui-même qui eut, en la personne d'Emil Hurja, son « monsieur sondage ». Enfin, pour couronner le tout, ces sondeurs prédirent la réélection de Roosevelt en 1936, contre les « votes de paille », ce qui contribua énormément à leur notoriété. Les sondages se sont ainsi diffusés dans de très nombreux milieux et à un rythme soutenu. Après-guerre, forts de leur autorité au plan scientifique, mais aussi de leur apport au plan démocratique, ils furent relayés par la presse et même utilisés dans des entreprises et dans certaines administrations. Ces succès ont peu à peu montré à l'opinion (qui était non seulement un objet d'étude mais pouvait être juge) l'efficacité de la représentativité telle qu'ils la concevaient.

Notons que la réaction des probabilistes, comme celle de Frederick Stephan (1948), alors professeur de statistiques à l'Université de Chicago, fut extrêmement habile – bien qu'elle eût peu d'échos dans l'opinion publique. Pour eux, le succès des sondages d'opinion prouvait encore une fois que la technique des sondages était légitime et qu'ils pouvaient à bon droit remplacer recensements et monographies. Mais dans le même temps, ils présentaient comme archaïques les méthodes utilisées en sciences politiques puisqu'elles ne recourraient pas à l'aléatoire et se cantonnaient au traditionnel « choix judicieux ». Ces succès prouvaient donc la force supérieure de leurs propres méthodes aléatoires, utilisées par les gouvernants pour estimer le taux de chômage, le nombre de familles migrantes ou le nombre de personnes bénéficiant de l'aide publique. En effet, si les sondages d'opinion ont rendu célèbre la notion de représentativité statistique, ils ont néanmoins imposé une définition très différente de celle des probabilistes qui tentaient de trouver un consensus. Malgré les efforts de Neyman, il demeure bel et bien deux définitions de la représentativité.

Conclusion

L'histoire de la représentativité montre que cette notion semble toujours échapper à ceux qui cherchent à la consolider, à la rendre monosémique, à clarifier son sens. Soit les auteurs lui accordent plusieurs définitions soit, lorsque Neyman essaye de la définir de façon univoque, la société a imposé à nouveau une pluralité de significations – et *vice versa* pour les spécialistes de l'opinion publique qui ne purent davantage faire disparaître le sens probabiliste. La représentativité semble trop riche pour être contenue dans les limites d'un concept univoque. Le rapport de représentation semble être définitivement trop politique pour se voir cantonner à une définition unique.

Cette conclusion semble avoir été tirée par les statisticiens eux-mêmes. Certes, certains utilisent l'une des définitions qui en été établie en ignorant ou en faisant mine d'ignorer la polysémie du terme. On pense par exemple à cette antienne



Meaning 1. General, usually unjustified, acclaim for data: The emperor's new clothes.



Meaning 3. Miniature of the population: Model train set.



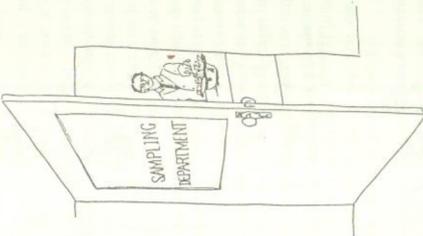
Meaning 2. Absence of selective forces: Justice balances the scales.



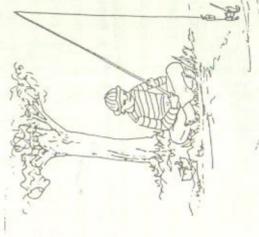
Meaning 4. Typical or ideal case: Superman and Supergirl and average man and average woman.



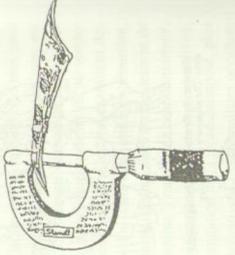
Meaning 6. Things are to be made precise: A hatching egg of unknown species.



Meaning 7. Some specific sampling method: The Sampling Department in action.



Meaning 9. Good enough for a particular purpose: Boy fishing with worm and safety pin.



Meaning 8. Permitting good estimation: Micrometer

Illustration extraite de l'article de Kruskal et Mosteller, "Representative sampling IV: the history of the concept in statistics, 1895-1939", International Statistical Review, 48(2), p. 170-171. Reproduction avec l'autorisation de la revue.

qui a cours lors des soirées électorales sur les sondages dits « représentatifs ». Mais la plupart la prennent en compte. Les statisticiens les plus versés dans les probabilités, ceux qui tendent le plus vers la constitution d'une science indiscutable, semblent ainsi préférer tout simplement l'ignorer. Par exemple, dans la *Théorie des sondages* de Christian Gouriéroux (1981), qui sert de support à la formation des administrateurs de l'Insee depuis les années 1980, cette notion n'apparaît tout simplement pas. Une autre position, par exemple celle de Kruskal et Mosteller, deux pionniers de l'aventure des sondages dont l'un deux conseilla Gallup en 1938, est de simplement lister ses différents sens. Ils concluent leur travail, constitué de quatre articles successifs consacrés à cette notion (Kruskal et Mosteller, 1976a, 1979b, 1979c, 1980), par une série de croquis (voir page précédente) illustrant les différents sens de la représentativité (1980, p. 170-171).

D'autres enfin, critiquent les sondages représentatifs, de façon parfois virulente. Bourdieu (1973) a par exemple pris appui sur le fait que les taux de non-réponse ne sont pas répartis également selon les catégories sociales – ce qui constitue un problème de représentativité – pour articuler une remise en cause radicale des sondages auto-qualifiés de « représentatifs ». Il arguait du fait que, non seulement, lorsqu'on calcule des pourcentages sur les seules opinions exprimées, on ne représente pas ceux dont l'opinion consiste à ne pas en avoir (qui refusent de répondre); mais en allant plus loin, si l'on pose les mêmes questions à tous les enquêtés, les sondages vont faire comme si les personnes interviewées comprenaient toutes le sens politique des questions, alors que, démontre Bourdieu, certaines catégories, souvent dominées, produisent leur réponse à partir de principes éthiques et non politiques. Il s'agit donc là d'une autre façon de dire que la diversité de la population n'est pas respectée ou que les sondages ne respectent pas une certaine définition de la représentativité.

Toutes ces positions semblent ignorer la spécificité mise en avant dans cet ouvrage : la polysémie de la représentativité permet de rendre les statistiques discutables (Desrosières, 2008, chap. 5). Elle permet de faire parler entre elles, non seulement plusieurs traditions statistiques, mais au-delà, de laisser une prise à la politique, à la philosophie, aux médias, bref, à la société entendue dans son sens le plus noble, d'entrer en interaction, que ce soit pour critiquer ou pour proposer des alternatives avec ceux qui se chargent de la quantifier. La représentativité n'est pas monosémique, et c'est peut-être grâce à cela qu'elle permet de tisser des liens avec d'autres mondes que celui des statisticiens.

Références bibliographiques

- BLONDIAUX L., 1998, *La fabrique de l'opinion*, Paris, Le Seuil.
- BOURDIEU P., 1973, « L'opinion publique n'existe pas », *Les temps modernes*, 318, p. 1292-1309.
- BOWLEY A.L., 1926, "Measurement of the precision attained in sampling", *Bulletin de l'Institut international de Statistiques*, 22, p. 6-62.
- BOWLEY A.L., BURNETT-HURST A.R., 1915, *Livelihood and Poverty*, London, G. Bell.
- BOWLEY A.L., HOGG M., 1925, *Has Poverty Diminished? A Sequel to Livelyhood and Poverty*, Londres, King and Son.
- BRU B., 1988, « Estimations laplaciennes », in Mairesse J. (dir.), *Estimations et sondages, cinq contributions à l'histoire de la statistique*, Paris, Economica, p. 7-46.
- CONVERSE J.M., 1987, *Survey Research in the United States. Roots and Emergence 1890-1960*, Berkeley, University of California Press.
- DIDIER E., 2009, *En quoi consiste l'Amérique ? Les statistiques, le New Deal et la démocratie*, Paris, La Découverte.
- DESROSIÈRES A., 1993, *La politique des grands nombres*, Paris, La Découverte.
- DESROSIÈRES A., 2008, *Pour une sociologie historique de la quantification*, Paris, Presses des Mines/Paristech.
- GALLUP G., RAE S., 1940, *The Pulse of Democracy*, New York, Simon Schuster.
- GALLUP G., 1948, *A Guide to Public Opinion Polls*, Princeton, Princeton University Press.
- GOURIEROUX C., 1981, *Théorie des sondages*, Paris, Economica.
- IGO S., 2007, *The Averaged American*, Cambridge, Harvard University Press.
- KIÆR A. N., 1895, "Observations et expériences concernant les dénombrements représentatifs", *Bulletin de l'institut international de statistique*, livre 2, pp. 176-183.
- KIÆR A. N., 1899, "Sur les méthodes représentatives ou typologiques appliquées à la statistique", *Bulletin de l'institut international de statistiques*, livre 1, p. 180-185.
- KIÆR A. N., 1903, "Sur les méthodes représentatives ou typologiques", *Bulletin de l'institut international de statistique*, livre 1, p. 66-78.
- KIÆR, A. N., 1905, "Discours sans titre sur la méthode représentative", *Bulletin de l'institut international de statistiques*, livre 1, p. 119-134.
- KIÆR A. N., 1976, *The representative method of statistical survey*, Oslo, Norwegian Central Bureau of Statistics.
- KRUSKAL W., MOSTELLER F., 1979, "Representative sampling I: non-scientific literature", *International Statistical Review*, 47(1), p. 13-24.

- KRUSKAL W., MOSTELLER F., 1979, "Representative sampling II: Scientific literature, excluding statistics", *International Statistical Review*, 47(2), p. 111-127.
- KRUSKAL W., MOSTELLER F., 1979, "Representative sampling III: the current statistical literature", *International Statistical Review*, 47(3), p. 245-265.
- KRUSKAL W., MOSTELLER F., 1980, "Representative sampling IV: the history of the concept in statistics, 1895-1939", *International Statistical Review*, 48(2), p. 169-195.
- LAPLACE P. S., 1785, « Discours du 30 novembre 1785 lu à l'Académie », in *Œuvres complètes de Laplace*, Paris, Gauthier-Villars.
- LIE E., 2002, "The rise and fall of sampling survey in Norway, 1875-1905", *Science in Context*, 15, p 385-409.
- MANIN B., 1995, *Principes du gouvernement représentatif*, Paris, Flammarion.
- MESPOULET M., 2001, *Statistique et révolution en Russie. Un compromis impossible (1880-1930)*, Rennes, Presses universitaires de Rennes.
- NEYMAN J., 1934, "On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection", *Journal of the Royal Statistical Society*, 97, p. 558-625.
- NEYMAN J., 1937, *Lectures and Conferences on Mathematical Statistics*, Washington DC, The Graduate School of the Department of Agriculture.
- POLLAK M., 1986, « Un texte dans son contexte. L'enquête de Max Weber sur les ouvriers agricoles », *Actes de la recherche en sciences sociales*, 65, p. 69-75.
- STEPHAN F.F., 1948, "History of the uses of modern sampling procedures", *Journal of the American Statistical Association*, 43, p. 12-39.
- TAYLOR H.C., TAYLOR A.D., 1952, *The Story of Agricultural Economics in the US, 1840-1932*, Ames, Iowa State College Press.
- WILKS S.S., 1951, *Elementary Statistical Analysis*, Princeton, Princeton University Press.