

JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

JEAN STOETZEL

La statistique et l'étude des opinions

Journal de la société statistique de Paris, tome 87 (1946), p. 52-60

[<http://www.numdam.org/item?id=JSFS_1946__87__52_0>](http://www.numdam.org/item?id=JSFS_1946__87__52_0)

© Société de statistique de Paris, 1946, 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 STATISTIQUE ET L'ÉTUDE DES OPINIONS

I

Au mois de juillet 1824, une petite revue américaine, le *Harrisburg Pennsylvania*, avait l'idée de lancer la première enquête d'opinion, et concluait au succès d'Andrew Jackson dans l'élection présidentielle dont John Quincy Adams devait sortir vainqueur. Cent onze ans plus tard, George Gallup et Elmo Roper prédisaient correctement, contre l'avis de fort bons juges de la situation politique, la première réélection du président Roosevelt. Or, non seulement Gallup et Roper osaient parler avec assurance, mais encore ils avaient travaillé indépendamment. On était en présence d'une technique objective. La base scientifique de cette technique, il faut la chercher d'abord dans la théorie de l'échantillonnage et les résultats satisfaisants obtenus dans la « méthode représentative » qui en dérive. Adolphe Jensen a présenté, en 1926, à la XXI^e session de l'Institut International de Statistique, un rapport remarquable sur les utilisations de cette méthode. Jensen fait surtout état d'enquêtes plus ou moins officielles, et de nature plus ou moins démographique ou sociale, car la méthode représentative n'est guère sortie de ce domaine jusqu'à l'époque où il écrit. Depuis ce moment, toutefois, de nombreuses études de marché, œuvres d'organisations purement privées, devaient à leur tour s'inspirer du principe de l'échantillonnage, et c'est plutôt, semble-t-il, à cette dernière source de connaissances et d'expériences qu'ont puisé, à partir de 1935, les organismes d'étude de l'opinion.

Il me semble pourtant que ce ne soit pas la seule. Partie du journalisme avec la revue de Harrisburg, revenue au journalisme dans les colonnes de la revue *Fortune* où se publient les enquêtes de Roper, et dans celles des quotidiens abonnés aux services de Gallup, l'étude de l'opinion a fait dans l'intervalle un crochet par les Universités. Entre 1925 et 1930, les premiers travaux pour l'évaluation des attitudes et des opinions, dus à Allport, à Katz, à May, à Rice, à Symonds, à Thurstone, par exemple, ont préparé la voie à d'innombrables recherches qui ont fait des opinions, à partir de 1930, le domaine peut-être le plus exploré de toute la psychologie sociale. La seule présence du nom de Thurstone dans cette liste — pour ne point parler de Rice — nous garantit suffisamment l'importance du rôle que la statistique, encore que sous une forme très différente, ne pouvait manquer de jouer dans ces études universitaires. Nous avons fait leur place à ces travaux dans des publications antérieures.

Que les instituts d'opinion aient profité de ces recherches, cela n'est pas douteux : indi-

rectement, par l'intérêt qu'elles avaient éveillé pour ce genre de problèmes dans le public cultivé, dont les sentences, en l'occurrence, peuvent avoir tant de poids; directement, par les leçons sorties du laboratoire, qui allaient pouvoir s'appliquer dans les vastes expériences de plein air. Pourtant, jusqu'aux premières années de cette décennie, et malgré les attaches académiques de certains des observateurs de l'opinion publique, on peut dire que les laboratoires psycho sociaux et les instituts d'opinion ont en grande partie vécu dans une sorte d'isolement réciproque. Il en résultait, dans les cas les plus défavorables, qu'une puissante technique était provisoirement perdue pour la science, et qu'inversement cette technique renonçait temporairement à la coopération des spécialistes scientifiques qualifiés et aux importants progrès qu'elle pourrait lui valoir.

Cette situation n'est, toutefois, déjà plus celle des États Unis. Il s'est d'abord trouvé qu'en 1940, un « Bureau de recherches sur l'opinion publique », grâce à une fondation et à des subsides universitaires, s'est créé à l'Université de Princeton. Cette organisation, dirigée par Hadley Cantril, un des meilleurs psycho sociologues américains de la jeune génération, se consacre à des recherches sur la méthode des sondages et sur les lois des opinions. L'an dernier, Cantril a pu publier, sous le titre de « La mesure de l'opinion publique » (*Gauging public opinion*), le résultat des premiers travaux de ses collaborateurs pendant quatre ans. En 1941 s'est constitué d'une manière analogue, au sein de l'Université de Denver, un « Centre national de recherches sur l'opinion publique », dont les préoccupations sont surtout sociologiques. D'un autre côté, le Gouvernement fédéral a fait appel, pour leur confier la direction de recherches par sondages dans différents domaines, à des savants hautement qualifiés, Samuel Stouffer, par exemple, à l'État Major général de l'armée, Rensis Likert au Département d'agriculture, Edwards Deming au Bureau du président. Ils ont eu la préoccupation d'attaquer les problèmes qui leur étaient posés, en les prenant avec les implications les plus larges. Ils ont réussi, non seulement à s'acquitter des charges précises qui leur étaient confiées, mais encore à faire progresser la théorie et la technique des études d'opinions. Enfin, le 22 décembre dernier, à New York, il nous a été donné d'assister à la première session d'une commission réunissant les théoriciens et les praticiens des études d'opinions et qui aura pour rôle de poser en termes communs les problèmes rencontrés par tous, d'en répartir l'étude et les charges, et finalement d'en diffuser les solutions.

La seule création de cette commission montre assez que les difficultés présentées par l'étude des opinions sont loin d'être toutes surmontées et beaucoup d'entre elles sont de nature statistique. C'en est assez dire pour qu'on s'attende ici à un exposé incomplet. Il reste que d'appréciables progrès ont été accomplis, depuis l'époque du premier sondage entrepris parmi ses lecteurs par le *Harrisburg Pennsylvania*.

II

Parmi les cinq étapes que l'on peut abstraitement distinguer dans l'exécution d'un sondage d'opinions, la première est évidemment l'échantillonnage lui-même. Un sondage peut être défini comme l'opération qui consiste à porter un jugement sur les caractéristiques d'un ensemble (« univers » ou « population ») après examen d'une partie seulement de cet ensemble : il est clair que toutes les autres précautions que l'on aura pu prendre seront vaines, si le sondage n'est pas d'abord basé sur un échantillonnage correct. Il ne s'agit pas seulement du nombre des sujets examinés à titre d'échantillons, mais surtout du mode de sélection des échantillons. En pratique, les instituts d'opinion examinent habituellement de deux à trois mille sujets, choisis en proportions convenables dans certaines catégories, au point de vue de quatre ou cinq critères, tels que la région géographique, l'habitat, le sexe, l'âge et la condition économique sociale. La réalisation de cet échantillonnage constitue le « plan d'enquête ».

La deuxième étape du sondage consiste dans l'élaboration du questionnaire. Une étude préliminaire approfondie du champ du problème d'opinion à étudier s'impose évidemment si l'on veut assurer la pertinence du questionnaire. A ces exigences de fond, s'ajoutent des conditions de forme auxquelles le questionnaire ne doit pas moins satisfaire : universelle intelligibilité dans le domaine sociologique mis en observation, univocité des questions, objectivité du libellé. Presque toujours un essai préalable des questions projetées permet de dépister, par des moyens très simples, qu'il n'est en général pas besoin de raffiner statistiquement, celles qui ne satisferont pas aux conditions précédentes. La troisième étape de la méthode consiste dans les interviews proprement dits. Ces interviews sont l'œuvre d'enquêteurs sélectionnés à raison d'un certain nombre de qualités intellectuelles (intelligence, curiosité scientifique), caractérielles (bon contact social), morales (honnêteté et objectivité) et aussi, de leur localisation géographique et sociale rendant immédiatement possible l'intégration de leurs sujets dans le plan d'enquête. Les questions sont toujours posées oralement, en privé, et pendant une période de temps très limitée. Les réponses recueillies par les enquêteurs et consignées sur des bulletins d'enquête doivent ensuite être dépouillées, et c'est la quatrième phase du travail. Elles sont codifiées, c'est à dire rangées dans un certain nombre de classes de réponses, ensuite reportées sur cartes perforées, enfin triées par trieuses électro-mécaniques. Des tris partiels par catégories sociologiques, des tris

combinés par catégories de réponses aux autres questions, rendent en outre possible une analyse plus poussée des attitudes du public.

Il reste enfin l'étape finale, où ces résultats bruts seront analysés et interprétés. Le rapprochement des éléments numériques obtenus permet de tirer des conclusions sur les différences des tendances dans les diverses catégories de sujets, et les tableaux combinés font ressortir des associations d'opinions souvent inattendues et toujours révélatrices de la véritable portée des réponses. C'est ici naturellement que l'auxiliaire de la statistique redevient impérieusement nécessaire. Les épreuves de signification peuvent seules permettre de discriminer entre les différences et les associations qui sont dues au hasard des fluctuations d'échantillonnage, et celles qui sont fondées avec une probabilité suffisante sur des causes objectives. Le problème statistique qui se pose à cette occasion est beaucoup moins un problème théorique qu'un problème de pratique et de rendement.

Au début du XVIII^e siècle, Bernoulli a posé les bases de toute théorie et de toute pratique de la méthode des sondages. La loi des grands nombres qui, dans sa précision actuelle, dérive directement de son *Ars Conjectandi*, établit en effet le double fait suivant : si on examine, au point de vue de deux caractères s'excluant mutuellement, un certain nombre d'échantillons dans une population, on peut atteindre une probabilité aussi grande qu'on veut pour que l'écart entre la proportion trouvée pour la présence de l'un des caractères et la vraie proportion, soit au plus égal à une valeur fixée à l'avance; il suffira de prendre un nombre d'échantillons convenable, donné par un calcul préalable. En second lieu, pour une probabilité donnée, l'écart est proportionnel à la racine carrée du nombre des échantillons. Il est sans doute inutile de rappeler les démonstrations et les formules, qui sont classiques. Citons seulement quelques données numériques pour fixer les idées. Par exemple, si la proportion vraie des deux caractères dichotomiques est de 50 %, si l'on a pris deux mille échantillons, on a moins de une chance sur cent pour que l'écart positif ou négatif soit plus grand que trois pour cent; si l'on a pris trois mille échantillons, le même écart sera atteint ou dépassé environ une fois toutes les mille expériences. Si les proportions vraies étaient de 95 % et 5 %, dans les mêmes conditions les écarts seraient respectivement d'environ un et quart pour cent et un et demi pour cent.

Malheureusement, cette théorie n'est aussi simple qu'à raison des hypothèses mathématiques restrictives qui ont été posées au départ. Certaines de ces hypothèses ne présentent pas de véritables difficultés en pratique. Sans doute faut-il supposer en toute rigueur que la population d'où l'on tire les échantillons est infinie; mais les corrections que l'on pourrait être tenté d'apporter de ce fait aux données ci-dessus sont si faibles dès que la population est un peu grande (ce qui est presque toujours le cas), qu'elles sont pratiquement négligeables. Il n'en va pas de même d'une autre hypothèse, dont l'expérience peut s'écarter dans des proportions considérables. C'est en effet une condition essentielle de la théorie ci-dessus, que l'univers doit être homogène au point de vue du caractère considéré : si l'on examine un élément quelconque de l'univers, il doit présenter exactement la même probabilité que n'importe quel autre, de posséder le caractère considéré. Ce sera le cas d'un dé qu'on lance d'une manière répétée, en observant par exemple le fait qu'il donne, soit un as, soit une autre valeur. Il n'en va pas de même dans les sondages sur une population d'hommes, et notamment en matière d'opinions. L'expérience montre précisément, tout au contraire, que les opinions politiques varient avec la catégorie sociologique (sexe, âge, situation économique sociale, habitat, etc...). En d'autres termes, l'univers dans lequel on est amené à échantillonner, est un univers stratifié. Un siècle après Bernoulli, Poisson a investigué cette situation et s'est trouvé ainsi au point de départ de recherches théoriques d'un grand intérêt pour la théorie des sondages. Une mise au point récente, qui apporte elle-même quelques nouveaux progrès à la question, est l'article de M. H. Hansen et W. H. Hurwitz, paru dans le numéro de décembre 1943 des *Annals of mathematical statistics*, sous le titre « Sur la théorie de l'échantillonnage dans les populations finies » (*On the theory of sampling from finite populations*).

Il apparaît actuellement que, dans le cas d'une population hétérogène, il y a intérêt, plutôt que de tirer au hasard les échantillons dans l'ensemble de la population, à subdiviser l'univers en un certain nombre de groupes, et à prendre les échantillons dans ces groupes proportionnellement à l'importance numérique relative des groupes les uns par rapport aux autres. Dans ces conditions, les erreurs auxquelles on doit s'attendre avec une probabilité donnée, sont inférieures à celles qu'on aurait encourues par le procédé précédent. C'est précisément la méthode qui a été adoptée régulièrement par les différents organismes d'étude des opinions. L'édification du plan d'enquête a justement pour effet de suivre, à différents niveaux, les stratifications de la population, et de prendre dans ces groupes des échantillons en nombre proportionnel à leur importance numérique connue par ailleurs. Là est sans doute une des explications des succès répétés obtenus, à partir d'échantillonnages bien faits, par les différents instituts d'opinion dans leurs prédictions des résultats d'élections.

Une expérience simple illustre ce qui se passe effectivement. Une enquête française entreprise une semaine avant les élections du 21 octobre 1945, a trouvé qu'il y aurait 67 % de « oui » en réponse à la deuxième question du referendum. En fait, il y eut environ 66,5 % de « oui » pour la France métropolitaine. Un dépouillement centaine par centaine des réponses recueillies au cours de l'enquête donne les résultats suivants :

Nombres et proportions de « oui » recueillis, centaine par centaine, et cumulativement, en réponse à la deuxième question du referendum du 21 octobre 1945, au cours d'un sondage effectué une semaine avant le vote.

CEN TAINES succes sives	NOMBRE % de « oui »	ÉCART pour la probabilité de 0,20	ÉCART pour la probabilité de 0,50	ÉCART effectif	POUR CENTAGE cumulatif	ÉCART pour la probabilité de 0,20	ÉCART pour la probabilité de 0,50	ÉCART effectif
1	68	6 0	3 2	+ 1,5	68,0	6 0	3 2	+ 1,5
2	55	6 0	3 2	— 11,5	61,5	4 2	2 2	— 5,0
3	62	6 0	3 2	— 4,5	61 6	3 5	1 7	— 4,9
4	66	6 0	3 2	— 0,5	62 8	3 0	1 6	— 3,7
5	64	6 0	3 2	— 2,5	63 0	2 7	1 4	— 3,5
6	61	6 0	3 2	— 5,5	62 7	2 5	1 3	— 3,8
7	74	6 0	3 2	+ 7,5	64 3	2 3	1 2	— 2,2
8	75	6 0	3 2	+ 8,5	65 6	2 2	1 1	— 0,9
9	69	6 0	3 2	+ 2,5	66 0	2 0	1 0	— 0,5
10	69	6 0	3 2	+ 2,5	66,3	1 9	1 0	— 0,2
11	77	6 0	3 2	+ 10,5	67 2	1 8	1 0	+ 0,7
12	68	6 0	3 2	+ 1,5	67 3	1 7	0 9	+ 0,8
13	69	6 0	3 2	+ 2,5	67,4	1 7	0 9	+ 0,9
14	67	6 0	3 2	+ 0,5	67 4	1 6	0 9	+ 0,9
15	66	6 0	3 2	— 0,5	67 3	1 6	0 8	+ 0,8
16	68	6 0	3 2	+ 1,5	67 3	1 5	0 8	+ 0,8

On constate que, si la vraie proportion est bien 66,5 %, les proportions obtenues dans les différents échantillonnages d'une centaine ont atteint ou dépassé quatre fois l'écart correspondant à la probabilité de 0,20 — qu'ils devaient donc atteindre ou dépasser entre trois

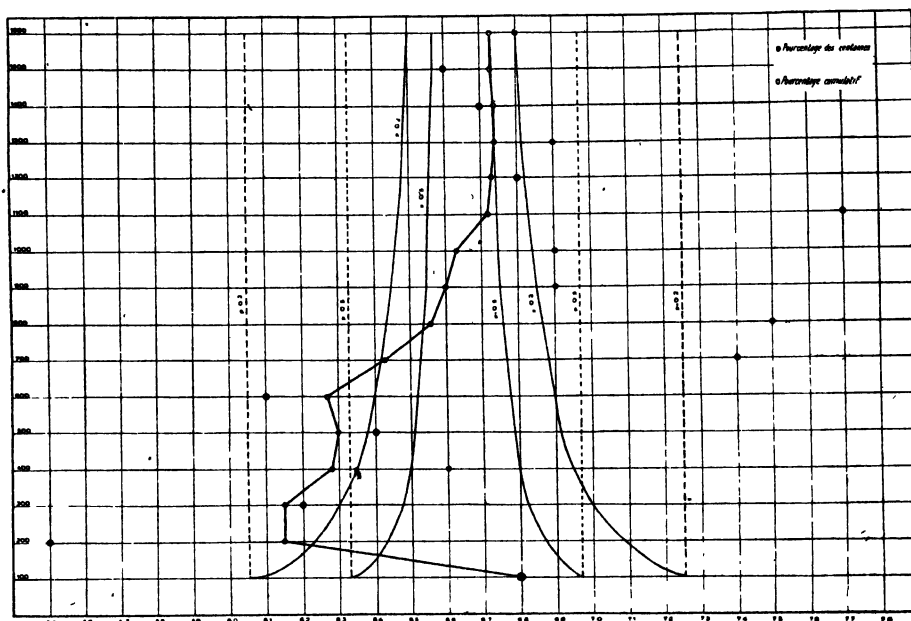


Fig. 1 — Dépouillement fractionné d'une enquête sur les intentions du public relativement à la deuxième question du referendum du 21 octobre 1945 (réponses « Oui »).

1.600 réponses « Oui » et « Non » recueillies au cours de l'enquête, ont été placées dans un ordre aléatoire et réparties en seize groupes de 100. Dans chaque centaine, on a déterminé le nombre de réponses « Oui ». Ces nombres sont figurés sur le graphique par des points noirs au niveau correspondant au rang de la centaine à laquelle ils appartiennent. Les marges en pointillés marquées $p = 0,5$ déterminent les zones extérieures dans lesquelles on doit s'attendre à trouver 50 % des résultats. On observera que 6 points seulement se trouvent dans ces zones. Les deuxièmes marges marquées $p = 0,2$ déterminent les zones dans lesquelles doivent se trouver 20 % des résultats. On y trouve 4 points.

Les points reliés par une ligne brisée continue sont les pourcentages cumulatifs du nombre des « Oui ». La vraie proportion du nombre des « Oui » au Referendum avait été d'environ 66,5 %.

et quatre fois — et ils ont atteint six fois l'écart correspondant à la probabilité 0,50 — qu'ils devaient donc atteindre ou dépasser huit fois (fig. 1). Il semble qu'on puisse conclure que indépendamment du fait global et capital de l'accord fréquent entre les expectations et la

réalité dans les enquêtes par sondage, sur un cas précis qui a pu être soumis à l'analyse, bien qu'elle ne s'applique pas à la rigueur, une théorie de l'échantillonnage fondée sur le schéma de Bernoulli ne rend pas tellement mal compte des faits.

L'exemple précédent est particulièrement favorable : on dispose de la vraie proportion. Il en est exceptionnellement ainsi. Les sondages ne sont pas un jeu de l'esprit. Il peut être intéressant, occasionnellement, de se placer dans des conditions telles qu'on puisse ultérieurement vérifier la validité des résultats. La fonction normale des études d'opinion est cependant de suppléer précisément à l'absence de toute indication sur les tendances et les attitudes du public. Lorsqu'il s'agit d'apprécier la valeur des résultats obtenus, les statisticiens se trouvent dans une situation délicate. En effet, l'ensemble des raisonnements probabilistes construits sur la méthode de l'échantillonnage donne par hypothèse la vraie proportion. La loi de Laplace-Gauss permet notamment, sachant le nombre des échantillons utilisés et la valeur de la véritable proportion, de déterminer, avec la probabilité qu'on voudra choisir, la proportion empirique limite qui pourra être atteinte. En pratique, le problème qui se pose est le réciproque du précédent : ayant obtenu une proportion empiriquement, on désire savoir, avec une probabilité donnée, la limite de l'écart qui peut séparer ce résultat de la véritable proportion. Pour tourner cette difficulté, S. S. Wilks a proposé l'artifice suivant : pour les différents niveaux de probabilité, il construit un abaque portant en ordonnées les différentes valeurs, de 0 % à 100 % de la vraie pro-

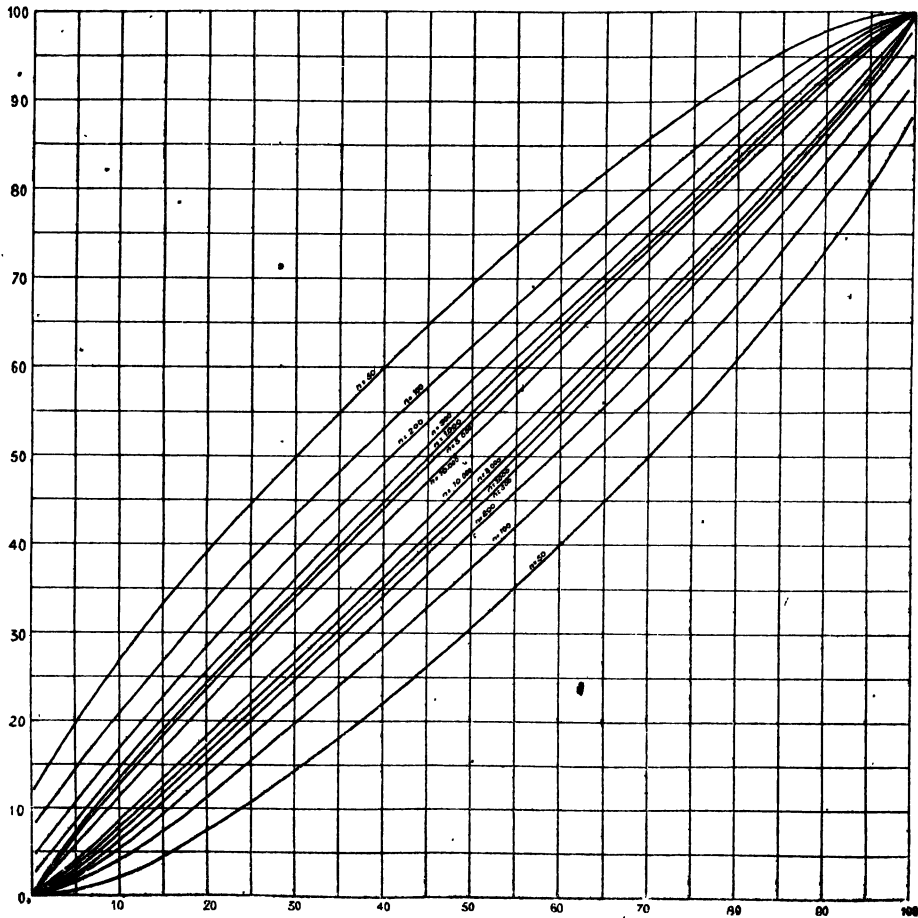


FIG. 2. — Abaque de S. S. Wilks.

Cet abaque construit par S. S. Wilks pour le niveau de probabilité 0,01 donne en abscisses les proportions trouvées dans un sondage, en ordonnée, les vraies proportions. Les nombres N d'échantillons utilisés se trouvent sur les courbes.

(Reproduit avec la permission de l'auteur.)

portion. A ce niveau de probabilité, pour chaque vraie valeur de la proportion, il correspond en général, pour un nombre donné d'échantillons, deux valeurs limites de la proportion empirique. Ces deux valeurs seront portées en abscisses. Chaque couple de coordonnées détermine un point dans le plan. En réunissant tous les points qui correspondent au même

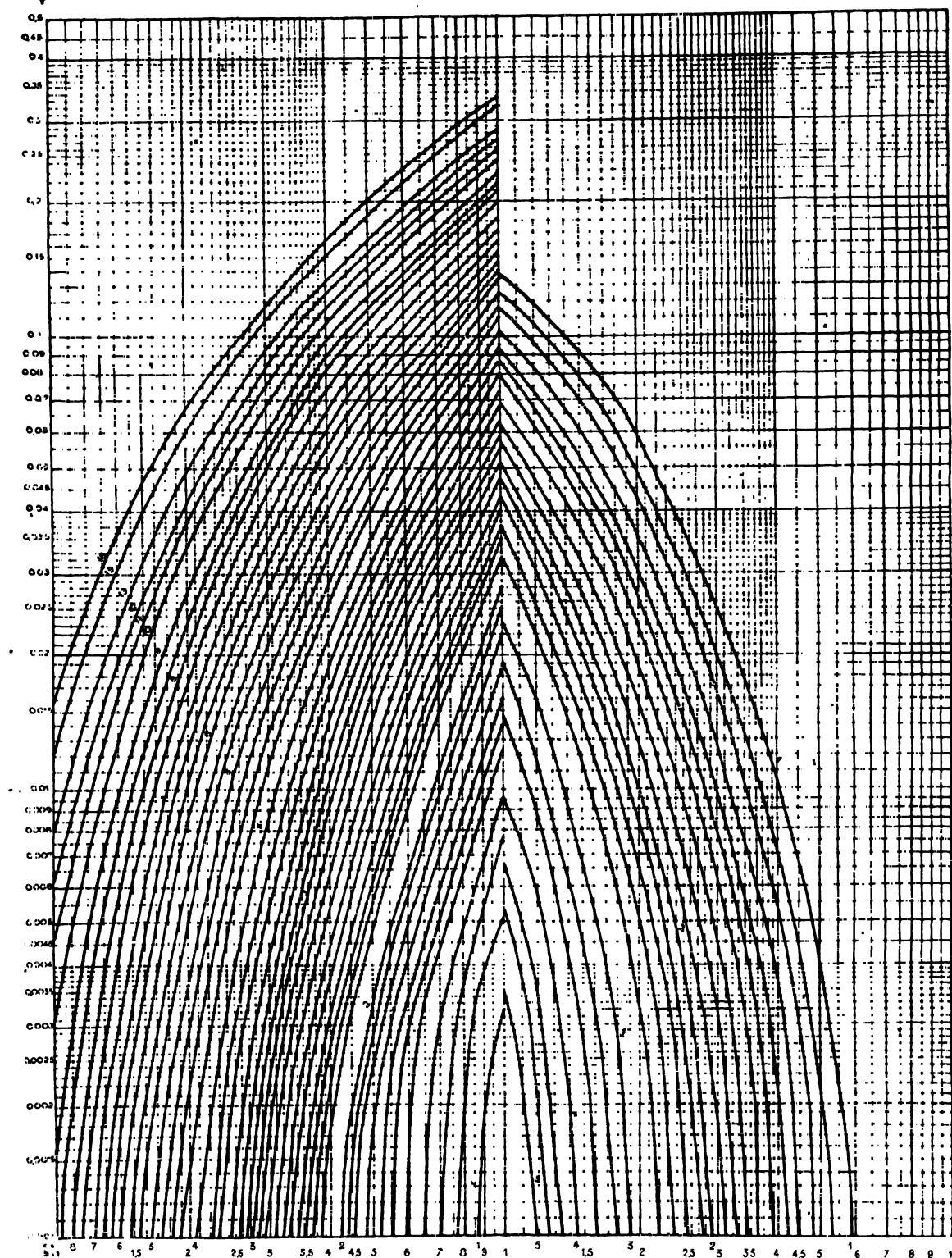
ϕ^2 $d=7$ 

Fig. 3. — Abaque de ϕ^2 pour la valeur $d = 7$.

Cet abaque donne en abscisses les valeurs de b jusqu'à 10, sur les courbes les valeurs de c jusqu'à 18, en ordonnées les valeurs de ϕ^2 ; b et c sont interchangeable. On a séparé, pour la commodité de lecture, les arcs croissants et décroissants de chaque courbe correspondant à la même valeur de c . Les différentes valeurs de b sont inscrites par suite deux fois sur l'abaque.

nombre d'échantillons, on obtient une courbe, et l'ensemble de la figure représente vaguement un ballon de rugby. Cet abaque « en ballon de rugby » de Wilks permet ainsi, en partant d'une proportion empirique quelconque donnée en abscisse, de lire, pour le niveau de probabilité choisi, les limites de la vraie valeur de la proportion avec le nombre d'échantillons utilisés (fig. 2). L'abaque de Wilks ne permet pas seulement de prévoir les limites entre lesquelles on doit s'attendre à trouver la vraie valeur de la proportion. Il doit encore permettre d'apprécier la probabilité qu'une différence entre deux proportions trouvées, par exemple, dans deux lieux différents, ou à deux moments différents, soit une vraie différence, et non pas une différence due à l'effet du hasard. Mais il ne saurait convenir dans d'autres cas rencontrés fréquemment dans la pratique, à savoir dans les cas où certains comportements d'opinions semblent associés à certains facteurs, et notamment avec une certaine prise de position sur un autre problème d'opinion. Dans ce cas, il s'agit de comparer le tableau des valeurs empiriques obtenues, très souvent un tableau 2×2 , avec le tableau qui aurait été obtenu théoriquement en supposant que les deux caractères dichotomiques, qu'on a des raisons de croire associés, aient été indépendants. Il reste alors à décider si les différences entre les deux tableaux sont significatives. On reconnaît ici l'usage de l'épreuve de χ^2 . Théoriquement, cette épreuve ne présente aucune difficulté, puisque nous disposons de différentes tables très pratiques de χ^2 et notamment de celle de Fisher. Malheureusement, sans être d'une application extrêmement laborieuse, l'épreuve de χ^2 dans le cas le plus simple et d'ailleurs le plus fréquent d'une table 2×2 , nécessite encore quatorze opérations, savoir quatre additions, quatre soustractions, deux multiplications et quatre divisions. Dès que le nombre des épreuves devient un peu élevé, et c'est généralement le cas, comme le travail d'étude des opinions est toujours très pressé, lié qu'il est à l'actualité, la charge de ces calculs devient rapidement intolérable. C'est pourquoi on doit recommander l'usage de la fonction ϕ^2 , liée à χ^2 , par la relation :

$$\chi^2 = \phi^2 N,$$

où N est le nombre des échantillons dans la table. Comme on le voit, χ^2 a l'avantage d'être indépendant de N, et peut donc être calculé une fois pour toutes, au moins pour les tables 2×2 . J'ai donc construit des abaques de cette fonction, très suffisants pour la pratique, et le nombre des opérations à effectuer se trouve ainsi ramené à quatre, et même à trois (fig. 3).

III

Les démarches dont il vient d'être question n'ont guère d'autre résultat que d'établir quelles opinions sont en faveur ou en défaveur dans le groupe étudié. Un début d'analyse apparaît avec la recherche des différences, c'est-à-dire en réalité des facteurs psycho-sociaux des opinions, et avec celle des associations d'opinions. Tout ce travail cependant ne pénètre guère en profondeur sous la surface des apparences. Le scepticisme dont certains bons esprits ont commencé par faire montre en présence des enquêtes d'opinions vient, au fond, d'une déception : la mise en œuvre de moyens relativement considérables aboutit finalement à des résultats apparemment sommaires et superficiels. Dans bien des cas, il apparaît que les opinions sont infiniment plus complexes qu'une alternative exprimée par deux pourcentages. Ce que les sceptiques ne savent pas, c'est que les psycho-sociologues, en présence d'une matière infiniment fluide et difficile à saisir, essaient, dans le silence, des méthodes dont on ne peut encore imaginer tous les fruits qu'elles porteront.

Une de ces méthodes en cours d'élaboration est celle qui a été imaginée, sous le nom de *scale analysis* (analyse des degrés) par Louis Guttman. Guttman a l'idée que certaines questions d'opinion constituent des ensembles tels que les réponses bi — ou multipartites qui leur sont faites possèdent un ordre hiérarchique qu'il s'agit de rechercher. Celui qui peut répondre exactement à une difficile question de mathématiques, remarque Guttman, pour des raisons culturelles évidentes, fera, *a fortiori*, des réponses exactes à des questions plus simples; la réciproque n'étant évidemment pas vraie, l'association des réponses, justes ou fausses, doit permettre à un observateur ignorant des mathématiques, par simple inspection intrinsèque des groupements de réponses, de décider de la hiérarchie en difficulté des différentes questions. De même, on peut rechercher entre les opinions relatives à un certain domaine d'opinion, un ordre hiérarchique qui constituera une « échelle des degrés de l'opinion » (*scale*) dans ce domaine. Un exemple aidera à fixer les idées :

Au mois d'août 1945, les quatre questions suivantes ont été posées à un échantillon représentatif du public français :

- (1) Êtes-vous favorable à la fusion en un seul parti des partis socialiste et communiste?
- (2) Si Franco ne s'en va pas de lui-même, les Alliés et notamment la France devront-ils intervenir en Espagne?
- (3) Si la Constituante, qui sera élue en automne, est une assemblée unique, cette assemblée devra-t-elle être souveraine, c'est-à-dire le gouvernement devra-t-il être responsable devant elle?
- (4) Êtes-vous satisfait du résultat des élections britanniques?

Les réponses de 78 % des sujets analysés se répartissent de la manière ci-dessous :

- 48 % ont répondu « oui » aux quatre questions;
- 13 % ont répondu « oui » à (2), (3), (4) et « non » ou « indécis » à (1);
- 10 % ont répondu « oui » à (3) et (4) et « non » ou « indécis » à (1) et (2);
- 9 % ont répondu « oui » à (4) et « non » ou « indécis » à (1), (2) et (3);
- 20 % ont répondu « non » ou « indécis » aux quatre questions.

Cette répartition peut être schématisée par le graphique suivant :

Fusion des partis socialistes et communistes.	OUI 48 %	NON et Indécis 52 %				
Intervention alliée en Espagne.	OUI 61 %	NON et Indécis 39 %				
Souveraineté de l'Assemblée.	OUI 71 %	NON et Indécis 29 %				
Satisfaction des élections britanniques.	OUI 80 %	NON et Indécis 20 %				
Distribution	48 %	13 %	10 %	9 %	20 %	

On constate que, dans le groupe des 78 % de sujets représenté ci-dessus, il existe un ordre hiérarchique entre les quatre questions : tout sujet qui a répondu « oui » à la première question, répondra « oui » aux trois autres; tout sujet qui a donné une réponse négative ou d'indécision à la question (4) répondra de même à (1), (2) et (3); et, d'une manière générale, les réponses affirmatives à une question d'un certain rang entraînent des réponses de même nature aux questions d'un rang supérieur, et inversement pour les réponses négatives ou d'indécision. Dans la mesure donc où les questions concernent un domaine politique commun, au sujet duquel les attitudes s'ordonnent dans une seule dimension, il est possible d'identifier cinq degrés de ces attitudes, dont les extrêmes sont exprimés par le groupe des 48 % qui ont répondu « oui » aux quatre questions et par le groupe des 20 % qui y ont répondu « non » ou « indécis », et comportant trois intermédiaires. A supposer (ce qui est du reste fort douteux, mais ce qu'on postulera pour les simples besoins de la cause) que les notions de « droite » et de « gauche » politiques puissent ici trouver leur utilisation, l'échelle ainsi établie permettra de raffiner cette alternative un peu grossière, et d'y substituer une gradation qualitative à cinq étages. Et, dans la mesure au moins où il s'agira des sujets appartenant au groupe des 78 %, il sera possible de localiser chacun d'eux à un étage déterminé.

L'établissement de telles échelles n'est évidemment pas toujours possible. L'existence de graduations parfaites n'est d'ailleurs jamais que théorique, et il ne peut s'agir dans la pratique que d'un ordre hiérarchique relatif. En ce qui concerne l'exemple précédent, il apparaît que cet ordre hiérarchique est particulièrement imparfait, puisqu'il ne peut être établi qu'à peine pour plus des trois quarts des sujets. La raison en est, très vraisemblablement, que le domaine d'opinions atteint par notre échantillonnage des quatre questions n'est pas unidimensionnel, qu'il comporte plus d'un facteur. La confusion qui paraît caractériser la situation politique actuelle de notre pays quand on essaie de la penser non pas en simples termes de partis, mais de programmes et de doctrines, vient précisément de ce qu'on aborde le problème avec le préjugé d'y retrouver la vieille dichotomie de la « droite » et de la « gauche ». La recherche des différents facteurs parmi les données expérimentales relativement nombreuses dont on dispose déjà, ne peut évidemment être entreprise que si l'on s'est muni des ressources de l'analyse statistique. Nous espérons être en mesure à brève échéance de rendre compte du résultat de certains efforts que nous avons tentés dans ce sens. Ces quelques indications suffiront sans doute pour montrer la diversité des perspectives qui sont ouvertes par la statistique à l'étude des opinions.

Jean STOETZEL.

DISCUSSION

M. BOURDON. — On pourrait essayer de préciser les domaines auxquels s'applique la très intéressante méthode qui vient d'être exposée. En Angleterre et aux États-Unis existe, en dehors des partis politiques, une masse d'électeurs flottants, dont le déplacement assure succès ou défaites : on cherchera donc à prévoir ce déplacement. En France, la première élection où se soit dessinée la répartition géographique des partis fut celle de 1849 et depuis lors jusqu'en 1939, cette répartition a changé progressivement, sans perdre ses lignes essentielles ni marquer des fluctuations en sens contradictoire.

Dans notre pays, la méthode des échantillons conduirait bien moins que la considération du passé à prévoir la répartition des suffrages. Je ne dis pas la répartition des sièges législatifs, qui ne dépend pas seulement du nombre des suffrages mais aussi du mode de scrutin et de la façon dont se font ou ne se font pas les coalitions entre les partis (ces derniers facteurs expliquent que, sans grand changement dans le nombre des suffrages, celui

des mandats obtenus par chaque parti ait été très différent en 1922 de ce qu'il avait été en 1919).

M. STÖETZEL. — M. Bourdon estime que d'autres méthodes que celles des sondages permettent de faire des prédictions sur les résultats des élections, et que cette méthode n'est peut-être pas la meilleure. Je suis, sur ce point, assez d'accord avec lui.

Ce que je tiens à dire, c'est que la véritable fonction des sondages n'est pas la prédiction des élections. Les instituts de recherches sur l'opinion acceptent volontiers de se soumettre à l'épreuve de l'expérience avant les élections : c'est pour eux, d'une part, un moyen de **montrer que les résultats qu'ils obtiennent ne s'écartent pas dans des proportions considérables, en général, d'un recensement exhaustif des attitudes politiques sur un point précis**; d'autre part, c'est aussi l'occasion de s'instruire sur la structure de la population qu'ils échantillonnent.

Mais il est indispensable de comprendre que la véritable portée des enquêtes d'opinion est autre chose qu'un jeu de société, consistant à anticiper de quelques jours sur un vote. Elle est, dans l'intervalle des élections ou en leur absence, tenue de faire connaître soit l'évolution des attitudes politiques, soit l'état des opinions sur des questions qui ne seront jamais soumises à une consultation électorale.
