JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

ROBERT GIBRAT
MAURICE GIRAULT
DANIEL DUGUÉ
GEORGES MORLAT

La statistique et l'art de l'ingénieur. Ce que la statistique apporte aux ingénieurs. Nécessité d'une formation statistique pour les ingénieurs

Journal de la société statistique de Paris, tome 105 (1964), p. 68-95
<http://www.numdam.org/item?id=JSFS_1964__105__68_0>

III

LA STATISTIQUE ET L'ART DE L'INGÉNIEUR

CE QUE LA STATISTIQUE APPORTE AUX INGÉNIEURS NÉCESSITÉ D'UNE FORMATION STATISTIQUE POUR LES INGÉNIEURS (1)

INTRODUCTION

PAR M. ROBERT GIBRAT

Vous allez entendre successivement quatre membres de la Société de Statistique de Paris traiter un ensemble de problèmes statistiques touchant l'art de l'ingénieur. Mon rôle sera en choisissant comme exemple les phénomènes atmosphériques, de montrer à ceux d'entre vous qui ne sont pas statisticiens ce qu'il y a derrière les problèmes que nous traitons. Pour mon compte, je n'essaierai à aucun moment de faire ici un exposé systématique.

(1) Communication présentée au cours de la séance organisée en commun par la Société des Ingénieurs civils de France et la Société de Statistique de Paris.

M. GIRAULT traitera ensuite d'un des problèmes fondamentaux de la statistique mathématique, celui de l'échantillonnage dont le but est de diminuer le coût des recherches statistiques.

M. DUGUÉ traitera des carrés latins. Comme La Fontaine descendant dans la rue un jour de découverte intellectuelle demandait à tous : avez-vous lu Baruch? j'ai, depuis un certain jour où j'avais entendu M. DUGUÉ, toujours conseillé à tous d'étudier les carrés latins. Un peu de patience et vous saurez ce que c'est.

Enfin, M. MORLAT terminera par l'examen de ce que peut apporter la statistique à l'art du commandement, car il vous parlera de la théorie de la décision.

PHÉNOMÈNES ATMOSPHÉRIQUES DÉBITS DES RIVIÈRES ET SUJETS CONNEXES

Depuis longtemps ingénieurs et statisticiens se sont rencontrés sur ce sujet : l'orateur a eu personnellement, dès le début de sa carrière d'ingénieur, à étudier à propos de la Truyère la fixation du débit de crue à retenir pour la détermination des ouvrages d'évacuation dans les installations hydroélectriques et s'est ainsi trouvé tout de suite plongé dans les problèmes de statistique.

Depuis lors les problèmes posés par l'optimisation de la production d'énergie par l'utilisation de réservoirs ont pris une très grande importance et l'école française, avec M. Pierre Massé en particulier, a obtenu là-dessus de grands succès. Mais la communication n'aura pas pour but, étant donné le temps accordé, de faire la revue des résultats obtenus dans ces divers domaines, mais de montrer à propos de quelques exemples les diverses méthodes employées, les résultats des succès ou des échecs, en un mot le conférencier essaiera de faire la « philosophie » des efforts; l'auditoire pourra ainsi, par lui-même, juger de ce que dans ce domaine la statistique a apporté et apportera aux ingénieurs.

*
* *

Pour vous montrer ce qu'il y a derrière la statistique je vais vous raconter une petite histoire et nous la commenterons ensemble. « Je ne veux pas boire d'alcool — me dit ma femme récemment — parce que notre deuxième fille attend un enfant dans deux mois, et si je ne bois pas d'alcool, ce sera une fille, ce que je désire. » En discutant cette affirmation et en l'approfondissant, nous allons retrouver les réflexions successives qui s'imposent à un statisticien qui veut s'occuper de phénomènes atmosphériques. Avec un commencement de culture statistique il pensera d'abord : ça n'est pas une raison si c'est une fille dans deux mois pour que le raisonnement fait par cette femme soit valable, car un résultat à une chance sur deux n'est pas suffisant en matière statistique et le voilà amené à essayer de rassembler des statistiques sur les résultats obtenus par un nombre suffisant de grand-mères acceptant de ne plus boire d'alcool deux mois avant une naissance. Un peu d'échantillonnage l'aidera. (Vous remarquerez cependant que la plus grande partie des statistiques médicales sont faites de cette façon, d'après les mauvaises langues (1).)

Ensuite, réfléchissant plus intensément, il se dira que le raisonnement est peut-être valable au fond malgré les apparences contraires, si l'ensemble des femmes qui ont assez de

(1) En fait, cela a été un garçon (note ajoutée par l'heureux grand-père à la correction des épreuves).

volonté pour refuser de boire de l'alcool pendant deux mois représente une classe particulière de l'ensemble des femmes qui ont de la volonté de manière exceptionnelle ce qui coïncide peut-être ou, tout au moins est en corrélation avec l'ensemble des femmes dont les filles ont plus souvent des filles que des garçons. Il faut donc que la statistique à construire utilise une enquête parmi toutes les grand-mères à volonté prouvée. (La théorie des carrés latins vous aidera à faire le plan d'expérience.)

Il pourra poursuivre les réflexions longtemps et compliquer en les raffinant de plus en plus les plans d'enquête jusqu'au moment où il notera que deux mois avant la naissance le sexe n'est plus aléatoire, car nous avons ici la chance de connaître relativement bien le phénomène. Le fait de boire ou non de l'alcool ne peut évidemment avoir de l'influence mais cependant les raisonnements précédents gardent toute leur valeur et ce serait une erreur grave d'attribuer à la phrase de ma femme une valeur nulle.

Je reviens donc aux phénomènes atmosphériques : il convient de se demander dès le début s'ils sont aléatoires ou mieux si, en le supposant, on peut tirer des résultats justes et intéressants mais, avant cela, il nous faut faire l'historique de l'étude statistique de ces problèmes et j'utiliserai un texte de M. MORLAT (1). Il a la gentillesse de rappeler que le livre que j'écrivais il y a une trentaine d'années sous le titre « Les inégalités économiques » avait montré « à la fois les avantages qu'on trouve à regarder les phénomènes économiques les plus variés sous l'angle statistique et l'ampleur du domaine d'application d'une loi simple et commode : la loi logarithmico-normale ». Vers la même époque ajoutait-il : « M. GIBRAT s'attachait l'un des premiers à l'étude des lois de probabilité des débits de rivière et à la recherche des règles de gestion des réservoirs ».

Dix ans plus tard, expose M. MORLAT, M. Pierre MASSÉ, l'actuel Commissaire au Plan, s'intéressait aux liens de l'hydrologie statistique et de l'économie électrique; son ouvrage sur « Les Réserves et la régulation de l'avenir dans la vie économique » devait jouer ensuite un rôle fondamental et reste aujourd'hui un des classiques de la recherche opérationnelle.

Il est très intéressant après 30 ans de refaire le point sur un problème un peu négligé; aussi ai-je lu ou relu pour vous beaucoup d'articles sur le traitement statistique des phénomènes atmosphériques grâce à une bibliographie établie par mon ami Stéphane FERRY et j'ai pu constater l'énorme travail fait depuis 20 ans par ce qu'on peut appeler l'École de l'Électricité de France dominant aujourd'hui tout le travail international en la matière.

Dans cette École, il faut signaler un personnage extraordinaire qui a eu une influence décisive. Malheureusement, il est mort beaucoup trop jeune et il est encore difficile d'apprécier l'exacte importance de ses contributions. Il s'agit d'Étienne HALPHEN dont l'action a été importante de 1940 à 1954. Il travaillait dans des conditions très originales puisque, par exemple, il a mis en évidence un certain nombre de fonctions mathématiques dites « factorielles » qui lui étaient nécessaires mais n'a su que dix ans après qu'elles étaient déjà connues, en particulier par HERMITE; il désirait se forger lui-même ses outils et ne souhaitait pas connaître ce qui avait pu être fait avant lui, heureux luxe à la portée de peu d'élus. Il a cherché inlassablement les fonctions qui permettent de bien représenter un ensemble d'observations de débits, c'est-à-dire la fonction donnant le nombre d'observations de débit dépassant une certaine valeur.

Avant d'aller plus loin, revenons à nos réflexions sur la nature aléatoire de ces phénomènes. Il paraît incontestable que si l'on considère la suite annuelle des quantités de pluie

(1) Texte communiqué à titre personnel daté du 19 juillet 1963.

tombées pendant un an, on aura un phénomène aléatoire souvent représenté par une courbe de Gauss classique.

Si, par contre, on prend la suite des débits journaliers, on verra vite que le débit d'une rivière est lié à celui de la veille, celui de la veille à celui de l'avant-veille et ainsi de suite, on est donc amené à essayer de représenter ces liaisons. J'ai été amusé de voir que ce problème n'avait presque pas avancé en 30 ans, car j'ai trouvé deux opinions récentes à peu près contemporaines, l'une du Président d'alors de la Société de Statistique de Paris, l'autre d'un futur Président, l'un indiquant que les débits journaliers successifs formaient une chaîne simple, l'autre affirmant l'inverse.

L'influence des débits passés se traduit-elle sur le débit d'aujourd'hui uniquement par le débit d'hier? Rien n'est aujourd'hui démontré, semble-t-il!

La suite des débits mensuels est plus facile à étudier, on peut négliger les liaisons, la courbe qui les représente n'est plus celle de GAUSS, mais une courbe dissymétrique à plusieurs paramètres. Pour les débits journaliers, il n'y a pas non plus de difficultés à considérer les différents débits d'une année, par exemple, et d'essayer de représenter la fonction donnant le nombre de ceux dépassant un certain chiffre, mais on utilise ainsi une très faible quantité des informations que pourraient donner la suite des débits et il faut donc que les questions posées soient de nature très particulière pour simplifier autant.

Mon idée sur ce sujet, il y a plus de 30 ans, fut d'utiliser la loi de GAUSS en l'appliquant au logarithme du débit et les résultats furent très frappants. J'essayais d'ailleurs de lui donner une base théorique. Mais cette loi a l'inconvénient de mal se prêter à la détermination des débits très petits ou des très grands. HALPHEN a recherché, un peu plus de dix ans après, des lois nouvelles sans utiliser les types de fonctions de l'arsenal statistique à sa disposition à son époque; il a créé, dès 1941, des lois d'un type spécial qu'il perfectionnait peu à peu en cherchant à avoir une décroissance exponentielle pour les grands débits. En 1948 il a établi les comportements symptotiques de ces lois en les classant en deux types très distincts et, en 1952, il s'est aperçu, comme nous l'avons déjà dit, que ces fonctions obtenues par une marche lente étaient déjà connues d'HERMITE.

Au passage, ses travaux l'ont amené à réfléchir très profondément à la notion de vraisemblance, introduisant des problèmes qui ont été repris depuis par des statisticiens et il est un peu dommage que l'attention des chercheurs n'ait pas été suffisamment attirée par les travaux de HALPHEN.

Ces recherches sont particulièrement importantes pour la détermination des crues à prendre en compte, pour la détermination d'un ouvrage, et les travaux ont été très nombreux depuis 30 ans. Certains se demandent encore si le phénomène des crues peut être traité par la statistique et s'il s'agit bien de phénomènes purement aléatoires car nous avons tous été frappés par le fait qu'après une crue centenaire il en arrive fréquemment une autre une année après, ce qui nous paraît insensé si la chance n'en était vraiment qu'une sur cent.

Le plus grand spécialiste mondial dans la cueillette des données statistiques sur les crues, le Français M. PARDE, est tenté de suspecter le caractère aléatoire mais on peut avancer des influences psychologiques pour expliquer ces anomalies. Quand une population vient de subir une catastrophe elle est indiscutablement beaucoup plus sensible à un phénomène du même genre même s'il est beaucoup plus faible. Chat échaudé craint l'eau froide... Elle aura donc tendance à les grouper par deux ou par trois en leur donnant la même amplitude.

Le danger que présente l'ouvrage menacé par la crue est très variable et la probabilité à prendre en compte sera ainsi très différente d'un ouvrage à l'autre. SCHNACKENBERG a donné pour la probabilité acceptable de la donnée entre deux crues pouvant détruire l'ou-

vrage, des chiffres variant de 20 ans à 1 000 ans. Vingt ans sera par exemple le cas où la crue correspondante apporterait seulement un dommage financier par ailleurs faible. Mille ans est envisagé pour les grandes catastrophes, j'ai d'ailleurs essayé moi-même, il y a quelques années, à propos de la grande inondation hollandaise à la suite d'une marée-tempête, d'approfondir cette idée de catastrophe, en distinguant aussi nettement que possible celles où la stabilité de l'État lui-même est mise en question, soit par l'ampleur des pertes financières ou humaines, au point de créer une tentation, soit par un affaiblissement durable. Dans ce cas, les raisonnements de matière financière que l'on est amené à faire d'une manière générale doivent être profondément modifiés et on est conduit à rejeter tous les raisonnements d'optimum.

La première méthode que j'ai utilisée pour essayer de déterminer le débit de crue à prendre en compte utilisait l'ensemble de tous les débits journaliers et essayait de déduire de la courbe de leurs fréquences, les probabilités des débits les plus élevés. Puis une approche nouvelle s'est introduite : elle consiste à s'intéresser uniquement à la plus grande valeur annuelle, c'est-à-dire à estimer que le maximum maximorum sur une certaine période s'obtient en étudiant les maxima annuels; en d'autres termes, pour savoir quelle sera la moyenne la plus forte du Major d'entrée dans une grande École sur une durée de 100 ans, il y a deux méthodes, soit prendre les moyennes de tous ceux qui ont été reçus à cette École et chercher la valeur extrême sur 100 ans, soit ne considérer que les moyennes des cent majors. La théorie du Major des majors, c'est-à-dire la distribution des plus grandes valeurs a été faite par M. Fréchet dès 1928 dans un Mémoire difficile à retrouver (1); il a démontré qu'il y avait trois types de lois possibles pour la distribution des plus grandes valeurs et ses travaux très puissants ont permis de mettre en ordre les travaux divers des statisticiens sur ce sujet. On n'a certes pas encore épuisé tout l'intérêt de ses recherches.

Je vais maintenant, pour rester dans le temps qui m'est imparti, vous parler d'un seul problème, à vrai dire, très célèbre à l'heure actuelle, celui de la pluie artificielle, car nous avons sur ce sujet la chance d'avoir un très gros rapport assez récent résumant les études très importantes faites aux U. S. A. pour savoir si réellement les essais avaient influencé les précipitations atmosphériques (E. D. F. a aussi beaucoup étudié le problème).

Plutarque avait remarqué qu'il y avait toujours des pluies le lendemain des grandes batailles; que ce soit vrai ou non cela relève du genre de problème traité à l'occasion de la déclaration de ma femme. On a essayé d'expliquer ce phénomène par le fait que les généraux de l'époque de Plutarque semblaient choisir, par un accord tacite, les jours de beau temps pour les batailles, qu'ils le faisaient, ai-je lu, là où il pleut beaucoup et qu'il n'est donc pas étonnant qu'il pleuve après puisqu'il faisait beau le jour de la bataille. Je reste sceptique, en tout cas on ne saurait généraliser cette méthode pour plaire aux agriculteurs ou à l'E. D. F.

Les Américains avaient essayé des explosifs en 1891 pour obtenir de la pluie et le Congrès a fourni les fonds, mais il n'y a pas eu de résultats.

Ils ont eu ensuite l'idée que des grands feux de forêts favorisaient l'arrivée de la pluie sans doute en créant dans l'atmosphère des courants de convection et l'idée a été reprise de nos jours sans que nous connaissions bien le mécanisme physique correspondant.

Mais c'est par un acte de « *serenpidité* », « art de profiter d'une circonstance inattendue », que l'on a pu faire avancer la question. [Je noterai au passage qu'un des plus beaux actes de

(1) J'en ai la photocopie sans indication de revue d'origine (elle est sans doute polonaise, l'article est daté du 12 septembre 1927 et la pagination va de 93 à 116). L'article s'intitule : « Sur la loi de probabilité de l'écart maximum ».

« serenpidité » fut celui de BECQUEREL, trouvant dans un tiroir une plaque photographique voilée et tirant de cette observation la radio-activité]. En Amérique, SCHAEFER, compagnon de LANGMUIR dans des recherches sur les gaz de combat au cours de la première guerre mondiale, a constaté dans son « freezer » domestique une fusion accidentelle de glace sous l'influence du gaz carbonique. A partir de cette observation il a construit une théorie de formation artificielle de cristaux qui peuvent être ensuite des amorces de condensation. Il est curieux de constater que 8 ans avant, un Suédois appelé BERGERON TOR avait fait la théorie du même phénomène et l'avait publiée sans que l'attention soit éveillée. Tout ceci a conduit à ensemencer les nuages par de l'iodure de sodium, soit par des fumées créées au sol, soit par avion.

Après que d'innombrables sociétés fondées pour la cause aux U. S. A. aient dépensé environ 6 millions de dollars en essais, les organisations officielles américaines ont voulu savoir si réellement il y avait eu, après ces ensemencements de nuages, plus de pluie ou non.

Seules les méthodes statistiques peuvent intervenir et ce n'est pas un problème commode même en utilisant la méthode de la cible-témoin comparant deux régions de climat analogue et assez voisines l'une où on ensemence, l'autre qui reste vierge.

La conclusion est assez précise dans les régions montagneuses ou séparées par des chaînes de montagnes, il y aurait moins d'une chance sur 1 000 pour que l'iodure de sodium soit sans influence, l'augmentation des précipitations ainsi provoquée étant de l'ordre de 8 %. Dans les pays non montagneux, au contraire, l'étude statistique des observations ne permet pas de conclure.

L'E. D. F. était, à cette époque, arrivée à des conclusions assez voisines avec des chiffres un peu plus faibles de l'ordre de 7 à 8 % dans la région de la Truyère, mais il semble qu'elle soit aujourd'hui très prudente dans ses conclusions et que les études les plus récentes ainsi que les discussions aujourd'hui très vives en cours aux U. S. A. sur la valeur des études statistiques faites, rendent les chercheurs français très sceptiques. La question serait donc encore très ouverte.

J'ai voulu, dans ce court exposé, renonçant *a priori* à toute présentation systématique, vous faire sentir toute la difficulté de bien utiliser la statistique dans les questions nouvelles, difficulté d'autant plus grande qu'un appareil mathématique très vaste est tout à fait au point et qu'il est bien tentant de s'en servir vite sans que la réflexion ait été suffisamment complète.

Robert GIBRAT
Directeur général d'INDATOM

UN PROBLÈME FONDAMENTAL DE LA STATISTIQUE MATHÉMATIQUE : L'ÉCHANTILLONNAGE

— *La notion d'échantillon issu d'une population.*

— *Observation de plusieurs échantillons parents : différences et analogies.*

Pour préciser les analogies et les exprimer clairement, il est indispensable d'avoir recours à un modèle mathématique; celui-ci est fourni par le calcul des probabilités.

La méthode ne conduit pas à une induction scientifique sous sa forme classique; mais elle propose des règles d'action. *Illustration sur un exemple.*

— *Aperçu sur les problèmes que pose l'échantillonnage.*

***

Dans la plupart des expériences faites actuellement, tant en recherche théorique qu'en recherche appliquée, les modèles déterministes ne conviennent pas : si l'on renouvelle plusieurs fois une expérience dans des conditions jugées analogues, les résultats ne sont pas identiques.

Exemple : Fabrication de tôles en alliage léger :

Une usine fabrique des tôles : des lingots d'alliage léger sont transformés en tôles par une série de laminages : les mêmes suites d'opérations effectuées à partir des mêmes matières premières ne produisent pas des produits identiques. Des mesures d'épaisseurs au centre de la bande obtenue fournissent par exemple les résultats suivants (en microns) :

2 483 2 509 2 473 2 457 2 513...

On reconnaîtra là l'exemple-type de toute fabrication industrielle. Ici on peut parler d'expérience au sens classique du terme : on en connaît les conditions (tout au moins les principales). On *sait* renouveler une expérience et l'on *peut* le faire à tout moment.

Il existe par ailleurs tout un ensemble de phénomènes qui intéressent l'ingénieur ou l'administrateur et qui se présentent d'une manière différente : le phénomène étudié se produit parfois; on sait reconnaître s'il se produit mais on n'est pas libre de le provoquer :

Exemple : pluies recueillies au mois d'août à l'observatoire du Parc Saint-Maur. Hauteurs en mm :

en 1880 : 60	1925 : 93
1885 : 66	1930 : 81
1890 : 43	1935 : 76
1895 : 42	1940 : 6
1900 : 60	1945 : 89

A cet exemple se rattachent les phénomènes météorologiques, climatologiques, etc., de nombreux phénomènes économiques, commerciaux et ceux que l'on étudie dans les « Sciences Humaines ».

Tous ces phénomènes, qu'ils soient provoqués ou simplement observés, entrent dans un schéma général qu'on désigne sous le nom *d'épreuve.*

Une épreuve est caractérisée par un ensemble de conditions qui la décrivent :

— fabrication par tel procédé, dans tel atelier, à partir de matières premières spécifiées;

— mois d'août au Parc Saint-Maur.

Ces « conditions » sont appelées des *hypothèses.*

Chaque épreuve produit un *résultat* auquel on s'intéresse :

— épaisseur de la tôle obtenue;

— hauteur de la pluie recueillie.

Dans les phénomènes étudiés en physique classique, les mêmes conditions d'une épreuve provoquent les mêmes résultats; on dit que les mêmes « causes » produisent les mêmes « effets » ou que les hypothèses de l'épreuve *déterminent* le résultat.

Il n'en va pas de même pour la catégorie de phénomènes indiqués plus haut : les mêmes hypothèses n'entraînent pas les mêmes résultats.

Naturellement, ce qu'on désigne sous le nom d'*épreuve scientifique* est une abstraction (aussi bien dans les modèles déterministes que dans les autres). Jamais les conditions d'une épreuve ne sont complètement observées et, par suite, ne sont complètement décrites.

Si l'on voulait être parfaitement rigoureux et lucide, il faudrait bien reconnaître qu'il est impossible de reproduire deux fois la même épreuve réelle. Cette attitude serait aussi stérile qu'elle voudrait être rigoureuse et l'efficacité de la méthode scientifique est si flagrante qu'elle n'a pas besoin d'être défendue : méthode qui admet délibérément un certain flou dans les clichés qu'elle prend et qu'elle retient. Toutefois, le passage d'un ensemble d'épreuves réelles à une épreuve abstraite soulève des difficultés : une épreuve doit être assez clairement définie pour que plusieurs observateurs soient également capables de reconnaître si elle se produit. Certaines conditions, considérées comme essentielles, seront complètement spécifiées (facteurs contrôlés); pour le reste, il est tout de même nécessaire de donner quelques indications.

Une épreuve abstraite étant définie, on désigne par « *population* » l'ensemble des résultats qu'on peut obtenir en effectuant une telle épreuve : ceux qu'on a obtenus, mais aussi ceux qu'on obtiendra. Il s'agit donc d'un ensemble abstrait qui n'est jamais complètement connu, mais dont on peut connaître des *échantillons.*

On appelle *échantillon de taille n* l'ensemble de *n* résultats d'une même épreuve.

Ainsi nous avons donné ci-dessus un échantillon de taille 10 de la population des hauteurs de pluies observables au mois d'août au Parc Saint-Maur.

Plusieurs échantillons issus d'une même population ne sont pas identiques; toutefois on peut observer qu'ils ne sont pas arbitraires :

Exemple : arrivées de bateaux dans un port : on note chaque jour le nombre de bateaux d'un certain type venant au port. Nous allons comparer deux types d'échantillons, notés A et B. Il s'agit d'un même port mais de deux types de bateaux. On admet que tous les échantillons (A) sont parents et qu'il en est de même des échantillons (B).

Les histogrammes (page 76) mettent en évidence une certaine *ressemblance* entre les échantillons parents : (échantillons A entre eux et échantillons B entre eux). Corrélativement ils accusent les différences qui distinguent échantillons A d'échantillons B.

Ces analogies et ces différences sont d'autant plus nettes que leseffectifs des échantillons sont grands.

Le rôle de la *statistique descriptive* est de *révéler* d'une manière aussi nette que possible ces caractères de parenté : soit en présentant des échantillons d'une manière globale (histogramme, fonctions de répartition...), soit en donnant des indices.

Ainsi, dans l'exemple ci-dessus les moyennes des 8 échantillons sont :

$$\text{effectif } N = 20 \begin{cases} a_1 \text{ (moyenne de } A_1) = 1{,}35 \quad b_1 = 3{,}25 \\ a_2 = 2{,}25 \quad b_2 = 3{,}5 \end{cases}$$

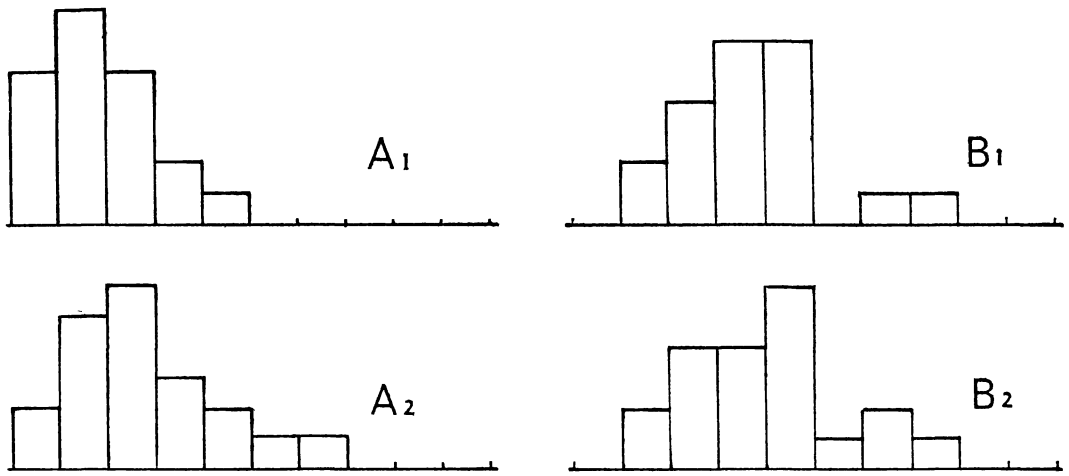
$$\text{effectif } N = 100 \begin{cases} a_3 = 1{,}91 \quad b_3 = 4{,}0 \\ a_4 = 1{,}87 \quad b_4 = 4{,}11 \end{cases}$$

Toutefois, quels que soient les effectifs des échantillons observés, les considérations précédentes, de la statistique descriptive, ne permettent pas d'exprimer clairement ni ces analogies ni ces différences, car les caractères qualitatifs retenus sont trop vagues. Dans cette voie, il n'est pas possible de donner des règles précises qui permettraient de distinguer d'une manière objective des échantillons non parents.

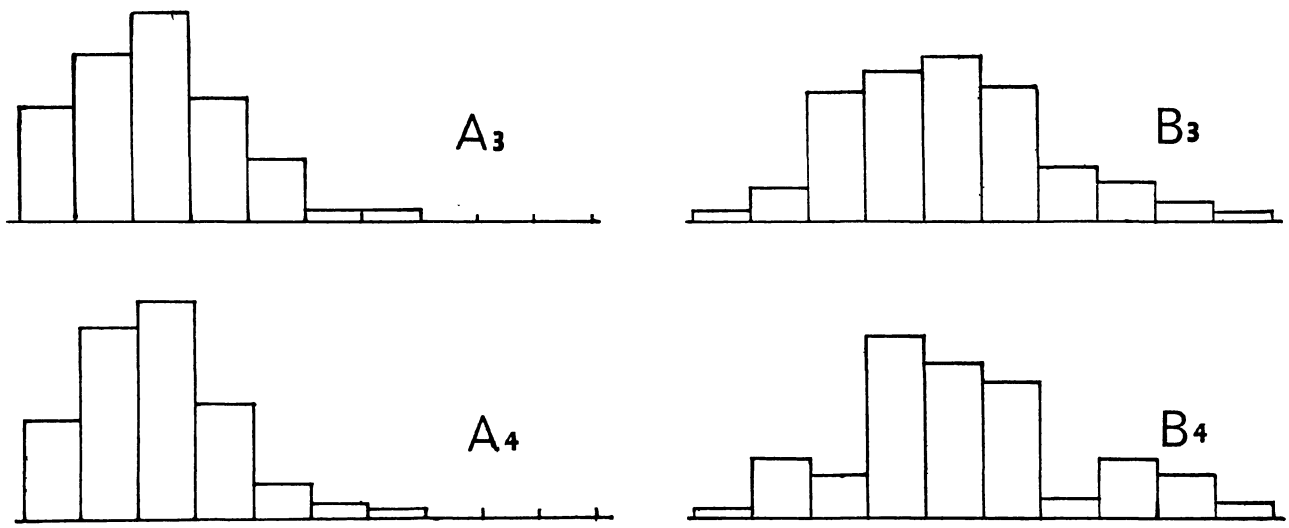
La théorie de l'échantillonnage se propose précisément d'étudier et de décrire des échantillons, donc de caractériser des échantillons parents. Ici comme ailleurs, pour élaborer une théorie, il faut construire un modèle abstrait sur lequel le raisonnement mathématique peut s'exercer rigoureusement : c'est le *modèle aléatoire* ou *calcul des probabilités.*

Dans cette théorie, chaque événement pouvant résulter de l'épreuve (1) est affecté d'une mesure : le degré de croyance qu'on lui attribue et qu'on appelle la *probabilité* de cet événement. Les probabilités des événements possibles sont liées les unes aux autres par certaines relations qui permettent un calcul et qui caractérisent la structure de cette théorie.

ÉCHANTILLONS D'EFFECTIFS N = 20



ÉCHANTILLONS D'EFFECTIFS N = 100



Pendant longtemps (XVIII^e et XIX^e siècles) on a cherché à utiliser le Calcul des Probabilités pour effectuer une véritable induction scientifique au sens classique du terme; c'est-à-dire de « remonter » des « effets » aux « causes ». Naturellement, la connaissance des « causes » ne peut s'exprimer ici qu'en langage probabiliste. La théorie est simple et claire; elle est basée sur l'application du théorème dit de BAYES, théorème très banal en lui-même; toutefois

(1) On appelle événement toute partie de l'ensemble des possibles.

l'application de cette théorie se heurte le plus souvent (pas toujours) à une difficulté grave qui en limite l'intérêt : la méconnaissance *a priori* des lois de probabilité.

Le point de vue actuel sur cette question est très différent et le problème dit de jugement sur échantillon se trouve posé d'une toute autre manière : si l'on cherche à connaître le comportement d'une famille d'échantillons parents, c'est en définitive pour agir, pour prendre des décisions. C'est à ce problème plus complet que s'attache actuellement la statistique mathématique. On ne recherche plus une connaissance pour elle-même, ni des « causes » (le terme a bien vieilli !) mais une *règle d'action.* Nous allons illustrer le procédé sur un exemple emprunté au contrôle de fabrication.

Un problème de décision statistique

Un atelier fabrique en série des *pièces.* Celles-ci doivent satisfaire certaines conditions techniques pour être déclarées bonnes. L'étude de cette fabrication a conduit à admettre qu'en régime « normal » chaque pièce a la probabilité $p = 0{,}02$ d'être mauvaise (et donc la probabilité $p = 0{,}98$ d'être bonne) et que ces probabilités sont mutuellement indépendantes pour toutes les pièces fabriquées.

Pour s'assurer que les conditions de fabrication restent satisfaisantes (qu'il n'y a pas de déréglages d'appareils par exemple) on prélève 100 pièces. Celles-ci sont contrôlées et soit K le nombre de pièces mauvaises trouvées dans l'échantillon. A la connaissance de K, on veut fixer une règle d'action : le choix entre les deux termes de l'alternative suivante :

1) admettre que la production est normale et donc continuer à produire dans les mêmes conditions, ou bien :

2) admettre que la production est perturbée et donc arrêter la fabrication pour vérifier les appareils.

Les difficultés du choix

Si certains appareils sont mal réglés, la qualité de la production peut encore être représentée par le schéma précédent mais avec une valeur du paramètre p supérieure à 0,02. Or quelle que soit la valeur réelle de p, le nombre K de pièces défectueuses peut prendre toutes les valeurs entières de 0 à 100; de sorte qu'il est impossible de déduire de K la valeur de p avec certitude. Quelle que soit la décision prise, on risque de se tromper :

Soit en choisissant (2) quand les conditions sont normales (on commet alors une erreur dite de première espèce).

Soit en décidant (1) alors que certains appareils sont déréglés (on commet une erreur dite de seconde espèce).

La statistique mathématique apprend à réduire considérablement ces risques.

Si la fabrication est « normale », compte tenu des hypothèses admises, le calcul des probabilités nous apprend que le nombre K peut varier de 0 à 100 par valeurs entières en obéissant très sensiblement à la *loi de Poisson* de paramètre $m = 2$. Si tous les entiers de 0 à 100 sont des valeurs possibles de K, celles-ci sont très inégalement probables :

En désignant par P_n la probabilité d'avoir $K = n$, on a :

$$P_0 = 0{,}1353 \qquad P_1 = 0{,}2707 \qquad P_2 = 0{,}2707$$
$$P_3 = 0{,}1804 \qquad P_4 = 0{,}0902 \qquad P_5 = 0{,}0361$$
$$P_6 = 0{,}0120 \qquad P_7 = 0{,}0034 \qquad P_8 = 0{,}0009$$
$$\text{Prob (d'avoir } K > 8) = 0{,}00023$$
$$\text{Prob (d'avoir } K > 10) = 0{,}000008$$

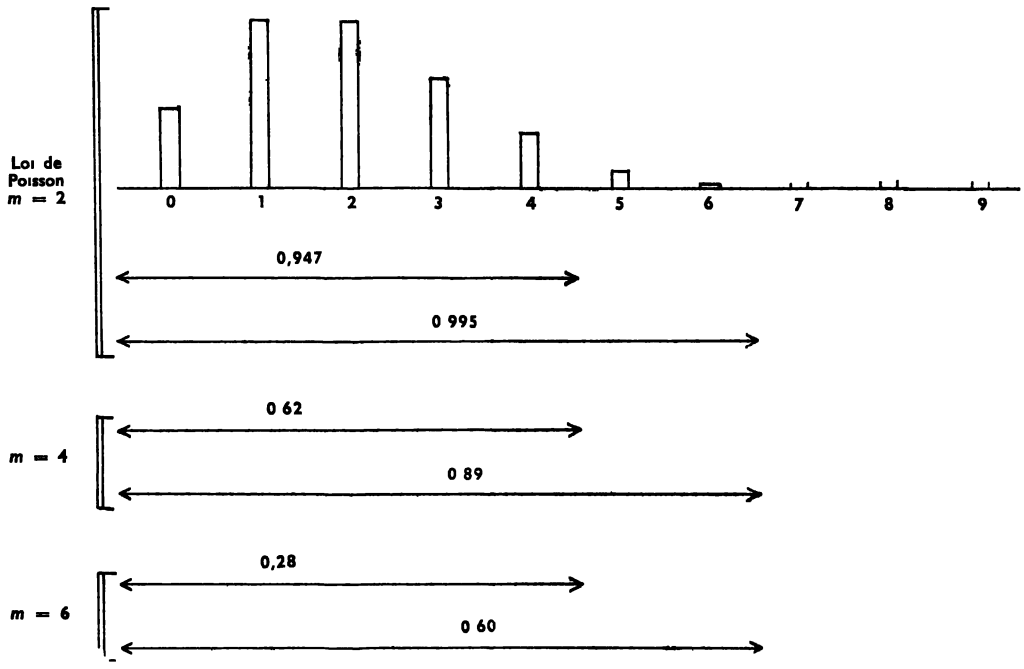
Donc, en fait, la distribution de probabilité de K est très fortement concentrée sur les petites valeurs (0, 1, 2, 3...) et il existe un très petit intervalle ayant une très forte probabilité de contenir K.

Choisissons un seuil de probabilité (0,95 par exemple) et cherchons le plus petit intervalle ayant cette probabilité de contenir K; c'est l'intervalle [0; 4] dit *intervalle d'acceptation.*

La règle est alors la suivante :

Si l'on obtient $K \leqslant 4$ on décide (1) (pas d'intervention).

Si l'on obtient $K > 4$ on décide (2) (intervention).



Test $m = 2$ contre $m > 2$
K obeit à la *loi de Poisson* de parametre m
Tableau donnant les probabilites qu'ont les intervalles de contenir K si $m = 2$; $m = 4$ ou $m = 6$.

Pour trancher entre ces risques, il faut tenir compte de leurs probabilités et surtout de leurs coûts. Ces techniques assez complexes relèvent de la *science des décisions;* elles seront abordées par M. MORLAT dans son exposé.

Quoi qu'il en soit des problèmes théoriques de décision, il faut remarquer que la marge de choix de l'intervalle d'acceptation est faible, du moins lorsque l'effectif de l'échantillon observé est grand. La probabilité accuse une décroissance très brutale à partir de certaines valeurs (ici $K = 5$). Si un déréglage se produit, il sera en fait très vite repéré par la méthode précédente.

Terminons ce rapide survol de la statistique en mentionnant quelques problèmes importants qui constituent autant de chapitres de la théorie :

La probabilité de commettre une erreur de première espèce (intervention inutile) est 0,05 (très exactement ici 0,053) c'est la probabilité d'obtenir $K > 4$ lorsque $p = 0{,}02$. On pourrait être tenté de réduire ce risque en choisissant un intervalle d'acceptation plus grand : [0; 6] par exemple, qui a la probabilité 0,995 de contenir K. Ce faisant, on augmen-

terait l'autre risque; car cet intervalle plus grand aurait plus de chances de contenir le nombre K, même si ce dernier obéit à une *loi de Poisson* de paramètre > 2. Supposons par exemple que, par suite de mauvais réglages, p prenne la valeur 0,06. Dans ces conditions K obéirait à la *loi de Poisson* de paramètre 6. La probabilité de ne pas intervenir (à tort) qui est seulement 0,28 pour l'intervalle [O ; 4] deviendrait égale à 0,60 pour l'intervalle d'acceptation [O ; 6]. Il y a donc conflit entre les deux risques.

Lois d'échantillonnage : études des lois exactes ou de lois approchées de certaines « statistiques ».

Tests dits « non paramétriques ». Tests basés sur les rangs des valeurs constituant l'échantillon. Ils ont l'avantage de ne pas faire intervenir les lois de probabilité des quantités étudiées.

Les considérations *d'économie* conduisent aux plans séquentiels d'échantillonnage (ou échantillonnage progressif) et aux notions de plans d'expérience. Cette dernière théorie, extrêmement séduisante, n'a toutefois été développée jusqu'ici que dans le cadre très étroit d'analyse de variance laplacienne. Son intérêt pratique n'en demeure pas moins incontestable.

Maurice GIRAULT

Professeur à l'Institut de Statistique de l'Université de Paris

CARRÉS LATINS (1)

L'étude de l'influence de divers facteurs, chacun d'entre eux pouvant se manifester à plusieurs niveaux, sur un ensemble de résultats, constitue ce que l'on appelle, selon Sir R. A. Fisher, le plan d'expérience (Design of experiments). L'un d'entre eux est le carré latin, qui est la figure formée par n *lettres latines, chacune répétée* n *fois et placée dans les* n^2 *cases d'un tableau carré, chaque lettre figurant une fois et une seule dans chaque ligne et dans chaque colonne.*

Dans le cas où un plus grand nombre de facteurs intervient, on peut utiliser les carrés latins orthogonaux (c'est-à-dire tel qu'en les associant deux à deux chaque couple de lettres se retrouve une fois et une seule).

La construction de ces carrés latins orthogonaux utilise l'algèbre pure et en particulier le corps de Galois, qui constitue ainsi un trait d'union curieux entre la statistique et l'algèbre.

***

Cette histoire commence comme un conte de fées : il était une fois en Écosse un vieil homme original du nom de Mac MAHON; il y avait deux traits marquants dans sa personnalité : des dispositions pour l'abstraction et une solide aversion pour le genre humain. Ayant des loisirs, il avait choisi de se consacrer à des recherches mathématiques et des recherches telles qu'elles ne puissent jamais être appliquées à quoi que ce soit. C'était ce que nous appellerions maintenant un mathématicien « pur ». Son sujet? les carrés latins. Voici le problème : considérons un carré de 5 cases de côté (soit 25 cases dans le carré) et 5 lettres latines A, B, C, D, E. Écrivons dans ces 25 cases, 5 fois la lettre A de manière qu'elle ne figure qu'une fois et une

(1) Le texte ci-dessus est extrait d'un exposé fait le 23 janvier 1960, par M. Daniel Dugué, au Palais de la Découverte, sous la titre : « Un melange d'algèbre, et de statistique : le plan d'expérience. » Il constitue la synthèse de la conférence de M. Daniel Dugué à notre Société, du 24 octobre dernier, et nous le reproduisons avec l'aimable autorisation de l'auteur et du Palais de la Découverte.

seule dans chaque ligne et dans chaque colonne, de même 5 fois B, 5 fois C, 5 fois D et 5 fois E. Nous avons réalisé ce que Mac MAHON appelait un carré latin.

Le problème est possible puisque le tableau suivant répond aux conditions posées.

A	B	C	D	E
D	E	A	B	C
B	C	D	E	A
E	A	B	C	D
C	D	E	A	B

On voit très aisément que si 5 est remplacé par un nombre quelconque le problème est encore possible. Il admet même un grand nombre de solutions. Mac MAHON s'est attaché à trouver le nombre de ces solutions et les moyens de passer d'un carré à un autre. Il a laissé une œuvre importante sur ce sujet. Il espérait, je pense, s'il était logique avec lui-même — et un mathématicien doit l'être — que cette production tomberait dans l'oubli sitôt lui-même disparu.

Le malheureux ne se doutait pas de deux choses : tout d'abord qu'un mathématicien (hélas ! un mathématicien appliqué selon la terminologie moderne) suisse EULER s'était déjà intéressé à un aspect de ce problème, et qu'un statisticien dont le nom marque toute l'école statistique de notre époque, Sir Ronald FISHER, allait utiliser ce modèle dans ce qu'on appelle d'après lui le plan d'expériences.

EULER avait posé vers 1750 le problème qu'il appelait problème des 36 officiers. Il s'agissait ici d'un carré de 6 cases de côté, donc de 36 cases en tout, sur lesquelles devaient être placés 36 officiers : 6 colonels, 6 lieutenants-colonels, 6 commandants, 6 capitaines, 6 lieutenants et 6 sous-lieutenants. Ces 36 officiers étaient de 6 armes différentes : 6 artilleurs, 6 fantassins, 6 cuirassiers, 6 dragons, 6 hussards, 6 sapeurs. De plus, dans chaque grade, chaque armée était représentée une fois et une seule (un colonel d'artillerie, un colonel d'infanterie, un cuirassier, un dragon, un hussard et un colonel de génie, etc.). Il fallait placer ces 36 officiers sur les 36 cases du carré de manière que dans chaque ligne et dans chaque colonne figure une fois et une seule chaque arme et chaque grade. *Le problème est impossible.* On peut bien placer 6 lettres A, B, C, D, E, F chacune répétée 6 fois de manière à constituer un carré latin puisque nous avons dit qu'il existe des carrés latins d'un nombre quelconque de cases. De même, les 6 lettres grecques $\alpha, \beta, \gamma, \delta, \varepsilon, \zeta$ étant répétées chacune 6 fois on peut aussi réaliser avec elles un carré latin. Mais quand on superpose ces deux carrés on n'aura pas une fois et une seule chacun des 36 couples formés en choisissant une lettre latine et une lettre grecque. Certains couples manqueront, d'autres seront répétés plusieurs fois. Il n'y aura pas, comme on dit, orthogonalité des deux carrés. Les lettres latines symbolisant les grades, les lettres grecques les armes, cela signifie que l'on pourra bien ranger les 36 officiers de manière que, dans chaque ligne et chaque colonne, il y ait un officier et un seul de chaque grade et un officier et un seul de chaque arme, mais que l'on serait obligé par exemple de ne pas avoir de colonel fantassin et que l'on aurait deux colonels artilleurs.

Considérons les deux carrés latins suivants :

A	B	C	D	E	F
C	D	E	F	A	B
E	F	A	B	C	D
B	C	D	E	F	A
D	E	F	A	B	C
F	A	B	C	D	E

α	β	γ	δ	ε	ζ
β	γ	δ	ε	ζ	α
γ	δ	ε	ζ	α	β
δ	ε	ζ	α	β	γ
ε	ζ	α	β	γ	δ
ζ	α	β	γ	δ	ε

Si l'on superpose les deux carrés on voit que A est associé une fois avec toutes les lettres

sauf α et δ, qu'il est associé 2 fois avec α et 0 fois avec δ. De même pour toutes les autres lettres latines un couple manque et un autre est répété 2 fois.

Cette impossibilité du problème des 36 officiers, EULER l'avait prévue dès le milieu du XVIII^e siècle. C'est seulement en 1900 que TARRY a réussi à l'établir. J'ajoute que sa démonstration ne présente aucune explication du phénomène. Elle se contente d'énumérer, en les résumant en plusieurs catégories, les cas possibles de carrés latins de 36 cases et de constater sur ces différentes catégories l'impossibilité de l'orthogonalité. Il semble d'ailleurs que le mystère de la raison profonde de cette impossibilité s'épaississe encore depuis des résultats acquis en avril 1959 par BOSE et que l'on soit là en face d'une propriété particulière du nombre 6. L'intuition géniale qu'a eue EULER à ce sujet n'en est que plus admirable. Il s'agit là d'une véritable prémonition de la vérité dont il y a plusieurs exemples en mathématiques (je pense à Henri POINCARÉ ayant la révélation que les groupes fuchsiens étaient des groupes de déplacement d'une géométrie non-euclidienne, à Paul LÉVY énonçant trois ans avant sa démonstration le théorème de LÉVY-CRAMER, à Denjoy donnant en 1906 la limite supérieure du nombre des valeurs asymptotiques d'une fonction entière, théorème qui devait attendre la démonstration d'AHLFORS jusqu'en 1928).

Quoi qu'il en soit et pour des raisons que je vous exposerai tout à l'heure, Sir Ronald FISHER s'est demandé si, en dehors du cas de 6 pour lequel la cause est entendue, il existe des carrés latins de n cases de côté pour lesquels il existe deux couples orthogonaux et même plus généralement N couples orthogonaux deux à deux. Cette question est restée en suspens jusqu'en 1938. Sir Ronald FISHER l'avait posée environ dix ans auparavant.

Cette année-là, à quelques jours d'intervalle, W. L. STEVENS en Angleterre et R. C. BOSE en Inde eurent l'idée qu'une des algèbres (pour employer un mot à la mode) les plus simples et les plus anciennement connues, les corps finis ou corps de Galois, fournissait une partie de la solution. Tout d'abord si l'on considère des carrés latins de n cases de côté on démontre très aisément qu'on ne peut en avoir plus de n-1 orthogonaux deux à deux (par exemple, il est impossible d'avoir plus de 4 carrés latins orthogonaux deux à deux, chacun ayant 5 cases de côté). Si l'on peut avoir un ensemble de n-1 carrés latins de n cases de côté, on dira que l'on a construit un ensemble *complet* orthogonal de carrés latins. La méthode de STEVENS-BOSE permet de construire des ensembles complets orthogonaux pour tous les nombres n (nombre de cases d'un côté) égaux à une puissance d'un nombre premier soit p^k, c'est-à-dire 2, 3, 4, 5, 7, 8, 9, 11, 13, 16, 17, 19, 23, etc.

Prenons par exemple $n = 5$ et considérons la table d'addition modulo 5.

0	1	2	3	4
1	2	3	4	0
2	3	4	0	1
3	4	0	1	2
4	0	1	2	3

Au croisement de la ligne commençant par 3 et de la colonne commençant par trois nous avons écrit 1 qui est le reste de la divison 3 + 3 par 5. Autrement dit, dans cette algèbre qui est un corps de GALOIS tous les nombres ordinaires sont remplacés par le reste de leur division par 5. Nous pouvons de même écrire une table de multiplication dans le même système.

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

(3 × 2 = 6 dont le reste de la division par 5 est 1)

Et maintenant écrivons 4 tables d'addition dans chacune desquelles la première colonne sera écrite dans un ordre différent déterminé par la règle suivante. Pour le premier carré, cette première colonne sera la suite des nombres 0, 1, 2, 3, 4, chacun étant multiplié (au sens de la multiplication que nous avons définie) par 1. Cela donnera naturellement 0, 1, 2, 3, 4.

Pour le second carré nous prendrons toujours 0, 1, 2, 3, 4, multipliés chacun par 2, ce qui, d'après la table de multiplication que nous avons écrite donnera : 0, 2, 4, 1, 3.

Pour le troisième carré nous aurons 0, 3, 1, 4, 2 et pour le quatrième 0, 4, 3, 2, 1. Nous en déduisons les tables d'addition suivantes :

0	1	2	3	4		0	1	2	3	4	
1	2	3	4	0		2	3	4	0	1	
2	3	4	0	1	pour 1	4	0	1	2	3	pour 2
3	4	0	1	2		1	2	3	4	0	
4	0	1	2	3		3	4	0	1	2	
0	1	2	3	4		0	1	2	3	4	
3	4	0	1	2		4	0	1	2	3	
1	2	3	4	0	pour 3	3	4	0	1	2	pour 4
4	0	1	2	3		2	3	4	0	1	
2	3	4	0	1		1	2	3	4	0	

Naturellement chacune des tables forme un carré latin avec les 5 éléments 0, 1, 2, 3, 4 et il est facile de voir que ces 4 carrés latins sont orthogonaux deux à deux. Je vous ai dit qu'on ne pouvait faire mieux. Pour tout nombre premier p, on peut construire un corps de GALOIS de p éléments qui sont les restes de la division d'un nombre par p. Ainsi les restes de la division par 11 (soit 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10) forment un corps de Galois dont on peut construire, de la même façon que pour 5, une table d'addition et une table de multiplication. Et les 10 tables d'addition obtenues comme ci-dessus constituent un ensemble complet orthogonal de carrés latins de 11 cases de côté. Pour les nombres p^k qui sont des puissances de nombres premiers (4, 8, 9...) il existe encore des corps de GALOIS ayant p^k éléments mais ils sont un peu plus compliqués à construire. Ces corps conduisent également à des ensembles complets orthogonaux. Par exemple pour 4 éléments A, B, C, D on a la table d'addition :

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

Ce qui signifie que dans l'algèbre ainsi définie $B + C = D$ et $C + C = A$

et la table de multiplication :

B	C	D
C	D	B
D	B	C

Ce qui signifie que dans cette algèbre $C \times D = B$

Selon les mêmes règles que pour 5 nous aurons les 3 tables d'addition suivantes :

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

A	B	C	D
C	D	A	B
D	C	B	A
B	A	D	C

A	B	C	D
D	C	B	A
B	A	D	C
C	D	A	B

qui forment 3 carrés latins orthogonaux deux à deux. Par conséquent tout corps de GALOIS conduit à un ensemble complet orthogonal de carrés latins. On démontre que les corps de

GALOIS ont obligatoirement p^k éléments. On voit donc que l'on a une forme de la solution pour les carrés ayant p^k cases de côté.

Ici, deux questions se posent :

1° pour les nombres de p^k cases de côté, a-t-on toutes les solutions au moyen des corps de GALOIS?

2° pour les nombres qui ne sont pas de la forme p^k, existe-t-il ou n'existe-t-il pas d'ensembles complets orthogonaux?

Autrement dit : y a-t-il réciprocité dans la propriété énoncée tout à l'heure? L'existence d'un ensemble complet orthogonal implique-t-il un corps de GALOIS?

La question pour le moment n'est pas tranchée. On a là un exemple de problème extrêmement simple d'algèbre finie et dont la solution dépasse de beaucoup les possibilités actuelles.

On peut se montrer moins exigeant. Sans rechercher un ensemble complet orthogonal on peut se demander si, pour des carrés latins de n cases de côté, il existe au moins 2 carrés orthogonaux. (Nous avons vu que pour 6 cette exigence réduite ne pouvait être satisfaite.) On a pu établir, en utilisant toujours le corps de GALOIS, qui joue évidemment dans cette théorie un rôle fondamental mais que l'on n'a pu encore tirer au clair, que pour les nombres n qui ne sont pas de la forme $2\ (2k + 1)$ il existe au moins 2 carrés latins orthogonaux. Par conséquent on sait qu'il existe, et on peut les construire, 2 carrés latins orthogonaux de 12 cases de côté, 2 carrés latins orthogonaux de 15 cases de côté; on sait même qu'il existe 3 carrés latins orthogonaux deux à deux de 20 cases de côté.

Il ne subsistait, il y a encore peu de temps, de doute que pour les nombres de la progression arithmétique $2\ (2k + 1)$, soit 6, 10, 14, 18... Nous avons vu que pour 6 la vieille hypothèse d'EULER de l'impossibilité de l'orthogonalité s'était trouvée vérifiée en 1900. On pensait qu'il en était de même pour les autres nombres de la progression et que, par conséquent, on ne pouvait pas non plus avoir 2 carrés latins orthogonaux de 10, 14, 18 cases de côté. J'ai même tenté d'utiliser les machines électroniques actuelles pour résoudre le problème au moins pour 10. Hélas! il aurait fallu les faire tourner pendant un temps dépassant largement les possibilités humaines (on m'a parlé d'un milliard d'années). J'ai d'ailleurs été heureux de constater que, malgré tout, dans certains domaines, l'esprit est supérieur à la machine puisqu'en avril 1959 un de mes élèves m'informait que la presse américaine publiait un compte rendu d'une réunion de la Société Mathématique Américaine où R. C. BOSE avait établi l'existence d'un couple de carrés latins orthogonaux de 10 cases de côté, ainsi que d'un couple de 22 cases de côté. Il pourrait donc se faire que 6 soit dans la progression $2\ (2k + 1)$ le seul nombre pour lequel l'impossibilité d'un couple orthogonal existe et que ce fait soit une propriété qui, pour le moment est mystérieuse, de 6. Nous avons jusqu'à présent beaucoup parlé d'algèbre. Parlons maintenant de statistique. Prenons par exemple 2 carrés latins orthogonaux de 4 cases de côté.

$A\alpha$	$B\beta$	$C\gamma$	$D\ \delta$
$B\gamma$	$A\delta$	$D\alpha$	$C\beta$
$C\delta$	$D\gamma$	$A\beta$	$B\alpha$
$D\beta$	$C\alpha$	$B\delta$	$A\gamma$

Imaginons que le grand carré représente un champ disons de 400 mètres de côté, chaque

petit carré représentant une parcelle de 100 mètres de côté. Ce champ va servir de terrain d'expérience pour étudier l'influence de 2 sortes d'engrais (un engrais potassique et un engrais azoté) sur la culture du blé. Les lettres latines vont représenter un niveau d'engrais potassique dans l'ordre suivant :

A niveau inférieur, puis B, puis C, enfin D niveau supérieur. Les lettres grecques représenteront de la même façon les différents niveaux d'engrais azoté : depuis α niveau inférieur jusqu'à δ niveau supérieur. Le champ sera ensuite ensemencé, la même quantité de blé étant semée dans chacune des 16 parcelles, chacune d'elles ayant reçu au préalable un engrais correspondant au couple de lettres qui y figure. Dans la parcelle Aγ (par exemple celle du coin inférieur droit) on aura mis la plus faible quantité d'engrais potassique et une quantité d'engrais azoté qui sera celle immédiatement inférieure à celle du niveau supérieur. Ensuite la récolte fournira 16 chiffres : le poids de blé recueilli dans chacune des parcelles. Le problème est de comparer l'influence des deux sortes d'engrais sur la récolte. Appelons x_{ij} la quantité de blé recueillie dans la parcelle située sur la i^e ligne et la j^e colonne : cette parcelle aura reçu un engrais déterminé par le « plan d'expériences » établi. Par exemple la parcelle donnant x_{32} aura reçu l'engrais potassique au niveau D et l'engrais azoté au niveau γ.

x_{ij} sera la somme des 5 quantités :

1° une quantité due à la variation de la fertilité du sol suivant une composante horizontale (parallèle à la base du grand carré);

2° une quantité due à la variation de la fertilité du sol suivant une composante verticale (parallèle à la hauteur du grand carré);

3° une composante due à l'influence éventuelle de l'engrais potassique;

4° une composante due à l'influence éventuelle de l'engrais azoté;

5° une dernière composante due aux influences n'entrant pas dans les 4 causes précédentes et que nous appellerons (la chose peut se discuter) composante aléatoire.

A cause de la disposition orthogonale des carrés latins utilisés, le plan d'expériences ainsi construit permet d'éliminer les 2 premières influences ainsi que la dernière pour ne conserver, en les séparant, que la troisième et la quatrième. La fertilité du sol constitue ici le facteur certain à éliminer. L'influence des deux engrais constitue les facteurs certains à contrôler.

Voici le procédé que l'on utilisera :

Nous avons défini x_{ij}. Introduisons maintenant les quantités suivantes :

$x_{i.}$ avec i prenant les valeurs 1, 2, 3, 4. Ce sont les moyennes des quantités x_{ij} situées dans la i^e ligne.

$x_{.j}$ de même avec j variant de 1 à 4 ce sont les moyennes des quantités x_{ij} situées dans la j^e colonne.

x_t avec t prenant les différentes valeurs A, B, C, D vont être les moyennes des x_{ij} dans les 4 cases contenant le même traitement potassique A, B, C. ou D. x_A par exemple sera la somme divisée par 4 des résultats de la récolte dans les cases situées sur la première diagonale.

et x_τ avec τ prenant les différentes valeurs, α, β, γ, δ sont les moyennes des x_{ij} dans les 4 cases contenant le même traitement azoté soit α, β, γ, δ.

$x..$ sera la moyenne générale, soit la somme des résultats divisés par 16. Ces différentes notations vont nous permettre de définir :

$$S^2 = \frac{1}{3} \sum_{i,j} (x_{ij} - x_{j.} - x_{ij} - x_t - x_\tau + 3\, x..)^2$$

$$S_t^2 = \frac{1}{3} \sum_t 4\, (x_t - x..)^2 \qquad S_\tau^2 = \frac{1}{3} \sum_\tau 4\, (x_\tau - x...)^2$$

Dans S^2 qui est la somme de 16 termes, l'un d'entre eux, celui relatif à x_{32} par exemple, sera :

$$(x_{32} - x_{3.} - x_{.2} - x - x\gamma + 3\,x..)^2$$

Les 2 valeurs de t et τ, soient D et γ, sont données par le plan d'expériences. Les résultats expérimentaux vont donc conduire à 3 valeurs numériques que l'on peut calculer. La théorie de l'analyse de la variance a permis sous certaines hypothèses (indépendance aux sens du calcul des probabilités, des quantités x_{ij}, normalité des x_{ij}, c'est-à-dire que leur loi de probabilité est de Laplace Gauss, identité de l'écart type de chacune d'elles) de calculer la loi de répartition des deux quantités $\frac{s_t^2}{S^2}$ et $\frac{s_\tau^2}{S^2}$ dans le cas où l'engrais potassique et l'engrais azoté sont sans influence. C'est ce que l'on appelle la loi de Snedecor (Snedecor ayant, en hommage à Sir Ronald Fisher, appelé F le quotient $\frac{s_t^2}{S^2}$ ou $\frac{s_\tau^2}{S^2}$). Dans ce cas $\frac{s_t^2}{S^2}$ comme $\frac{s_\tau^2}{S^2}$ ont une valeur probable voisine de l'unité et l'écart à l'unité a une probabilité donnée par la table de Snedecor. Ici intervient ce qu'en statistique on appelle le seuil de signification. Dans le cas en question où il y a 4 lignes, 4 colonnes, 4 niveaux du premier engrais et 4 niveaux du second engrais, il y a une probabilité de $\frac{1}{20}$ pour que le quotient F soit supérieur à 9,28 ou inférieur à $\frac{1}{9,28}$ et une probabilité de $\frac{1}{100}$ pour que F soit supérieur à 29,46 ou inférieur à $\frac{1}{29,46}$. $\frac{1}{20}$ et $\frac{1}{100}$ sont en général les seuils de signification adoptés. Supposons que l'on soit arrivé par le calcul à trouver $\frac{s_t^2}{S^2}$ égal à 12,45. Si l'on travaille avec le seuil de signification de $\frac{1}{20}$ cela entraîne que l'on accepte l'hypothèse de l'influence du traitement représenté par les lettres latines, c'est-à-dire ici de l'engrais potassique. Si ce traitement était sans influence il y aurait une probabilité égale à $1 - \frac{1}{20}$ pour que $\frac{9,28}{1} < \frac{S_t^2}{S^2} < 9,28$. L'événement en question aurait donc une probabilité inférieure à $\frac{1}{20}$ ce qui amène le rejet de l'hypothèse de non-influence étant donné nos conventions.

Si l'on était plus exigeant en adoptant le seuil de $\frac{1}{100}$ pour accepter l'hypothèse de l'influence du traitement potassique il faudrait que l'on ait :

$$\frac{s_t^2}{S^2} < \frac{1}{29,46} \text{ ou } \frac{s_t^2}{S^2} > 29,46$$

Dans le cas actuel, avec un quotient de 12,45 et un seuil de signification de $\frac{1}{100}$ on ne peut donc accepter l'hypothèse de l'influence de l'engrais potassique : le quotient $\frac{s_t^2}{S^2}$ s'écarte trop peu *significativement* de l'unité pour que l'on puisse accepter l'hypothèse.

L'idée fondamentale dans ces méthodes d'analyses de variance est donc de rejeter une hypothèse, si dans cette hypothèse l'événement observé a une probabilité trop faible de se produire (trop faible voulant dire inférieur au seuil de signification adopté à l'avance).

La méthode décrite est naturellement susceptible d'extensions à des ensembles de carrés latins orthogonaux deux à deux. On pourrait donc, encore dans le cas des carrés de

4 cases de côté, ajouter un troisième carré latin orthogonal aux 2 précédents, c'est-à-dire « tester », pour employer ce néologisme, l'influence d'un troisième engrais, par exemple de phosphates.

Je vous ai décrit ici un modèle de plan d'expérience agricole (c'est dans ce domaine qu'à la station expérimentale de Rothamstead, Sir Ronald FISHER l'a tout d'abord expérimenté). Mais il va de soi que cette décomposition de la variance par l'emploi de carrés latins orthogonaux rend les mêmes services dans d'autres domaines : la psychologie, la sidérurgie par exemple, où on les a souvent utilisés.

Je crois donc pouvoir dire que le lien entre l'algèbre et la statistique est établi par le plan d'expériences.

Daniel DUGUÉ

Directeur des études de l'Institut de Statistique de l'Université de Paris

STATISTIQUE ET THÉORIE DE LA DÉCISION

Entre la définition donnée par Littré au siècle dernier : « statistique : science qui a pour but de faire connaître l'étendue, la population, les ressources agricoles et industrielles d'un état » et une définition prise au hasard à la première page d'un des nombreux traités modernes de statistique : « statistics is a body of methods to take rational decisions in the face of uncertainty », se situe la statistique mathématique. Tous les hommes et tous les groupes humains étant amenés à prendre des décisions en présence d'incertitudes, la statistique les concerne tous. Elle a maintenant conquis droit de cité dans un certain nombre de domaines scientifiques, industriels ou économiques, où les incertitudes sont clairement formalisables en termes de calcul des probabilités : erreurs de mesures, biométrie, contrôle des fabrications, conjoncture, etc. Il reste une foule de problèmes où l'on ne soupçonne même pas toujours que le statisticien puisse jouer au moins un rôle d'utile conseiller. Quelques notions simples sur la théorie de la décision pourront aider à prendre conscience de ces possibilités.

***

Même si les statisticiens, qui sont gens raisonnables, s'en plaignent rarement, la statistique est une discipline bien mal nommée, et de ce fait, parfois mal renommée. Combien de nos contemporains pensent encore, avec un personnage de LABICHE, que le métier du statisticien consiste à compter le nombre de sexagénaires qui traversent le Pont-Neuf un dimanche après-midi.

Il est bien vrai que le rôle de la statistique, il y a un siècle ou deux, c'était de compter les choses.

« Science, dit LITTRÉ, qui a pour but de faire connaître l'étendue, la population, les ressources agricoles et industrielles d'un État... Les économistes ont créé un mot pour désigner la science de cette partie de l'économie politique — les dénombrements — et l'appellent statistique. »

Cette vocation de compter les choses, dont on pouvait se gausser au siècle dernier, a revêtu quelque noblesse depuis qu'on a reconnu combien elle est essentielle à la bonne marche des systèmes économiques, et donc des sociétés humaines. D'ailleurs, si les Incas prenaient

très au sérieux, avec raison, ceux de leurs fonctionnaires qui étaient chargés de faire des nœuds sur des ficelles de couleurs diverses, tenant ainsi à jour les statistiques vitales de leur État — comment nos contemporains n'auraient-ils pas de la considération pour des hommes qui disposent maintenant de machines à calculer électroniques, puissantes et coûteuses, qui savent avec une bonne précision de quoi sont faites la population et la production de leur pays et qui peuvent fournir, sur demande, au gouvernement, des chiffres pour éclairer ses décisions?

Nous devons donc constater que la statistique, au sens de LITTRÉ, a connu au cours des dernières décennies une promotion considérable, et cela suffirait à en faire une discipline pour l'honnête homme. Mais en même temps le terme de statistique a progressivement désigné un domaine d'activité intellectuelle infiniment plus vaste, et c'est de cette évolution-là que je veux parler.

La première phrase d'un traité élémentaire de statistique, publié il y a quelques années aux États-Unis, est celle-ci : « La statistique est un ensemble de méthodes pour prendre des décisions raisonnables en présence d'incertitudes ». Comme toutes les décisions humaines sont prises en face d'incertitudes, cette modeste phrase indique donc que la statistique, c'est la méthode pour prendre des décisions. Nous voici assez loin de la science des dénombrements. Par quel chemin la statistique — sans changer son étiquette, ce en quoi elle a peut-être eu tort, puisque c'est source de malentendus, mais il est trop tard de toute façon pour y revenir — par quel chemin a-t-elle pu passer des dénombrements à la théorie des Décisions? C'est ce que nous nous proposons ici de mettre en lumière, en retraçant sommairement les étapes par lesquelles la méthode statistique a progressé; nous n'aurons guère le souci de respecter l'ordre historique de cette évolution, mais seulement demontrer comment s'enchaînent, d'une manière toute naturelle, les diverses conceptions de la statistique qui vont de la « comptabilité des choses » à la logique des décisions.

*
* *

Les premières étapes sont les mieux connues. Pour rendre commodes les dénombrements, et pour présenter leurs résultats sous forme succincte et maniable, on al'habitude de calculer des moyennes et des écarts moyens, d'établir des polygones de fréquence, ou des histogrammes, des fibrogrammes, des courbes d'observations classées, et quelques autres représentations commodes. La manière d'établir des tableaux à double entrée — ou davantage — les graphiques représentant les variations d'un phénomène selon les valeurs d'un autre, le concept de régression, et le calcul des courbes de régression par la méthode des moindres carrés, la définition et le mode de calcul des divers indices de liaison entre des séries d'observations, tels le coefficient de corrélation, le carré moyen de contingence, et beaucoup d'autres — tout cela fait partie de la *statistique descriptive;* c'est l'art et la méthode pour manier, décrire et résumer des observations nombreuses. Mais lorsqu'il s'agit d'interpréter — lorsqu'on se demande, par exemple, si deux séries d'observations peuventêtre regardées comme homogènes — ou si tel phénomène dépend de tel autre — les recettes de la statistique descriptive ne nous permettent guère de fonder des raisonnements solides. Il faut, pour de telles interprétations, faire appel à des modèles, et les modèles convenables sont fournis par le *calcul des probabilités.*

Est-ce suffisant? Le calcul des probabilités, au sens strict, nous apprend par exemple comment on peut tirer, d'un modèle probabiliste donné, la loi de probabilité de la moyenne d'une série d'observations. Cela éclaire grandement la situation, mais le statisticien voudrait, connaissant des observations, et ayant quelque idée sur les modèles convenables, préciser

ces modèles ou les mettre à l'épreuve : il veut induire et non pas déduire. Voilà pourquoi le calcul des probabilités n'est pas suffisant, et pourquoi il a fallu élaborer une discipline nouvelle, étroitement fondée sur le calcul des probabilités, mais parfaitement originale cependant, par les problèmes qu'elle aborde, et par son objet précis, qui est de formaliser l'induction, et de fournir le modèle de ce que l'observation nous apprend des modèles. Cette nouvelle discipline a continué à s'appeler statistique — ou, pour être plus précis, on la dénomme *statistique mathématique.*

On convient généralement de considérer que la première pierre de la statistique mathématique fut posée par Karl PEARSON, avec son célèbre mémoire de 1900 dans le Philosophical Magazine : « On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. » En réalité, le problème de l'induction statistique avait été bien vu depuis fort longtemps, et le Révérend Thomas BAYES avait écrit un mémoire, publié après sa mort, en 1764, dans les Philosophical Transactions : « An essay towards solving a problem in the doctrine of chances » où il proposait une solution générale à ce problème. Mais la « formule de BAYES » — dont CONDORCET, LAPLACE et beaucoup d'autres usèrent et abusèrent parfois — pouvait paraître bien arbitraire. Au début de ce siècle il semblait souhaitable d'établir la statistique mathématique comme une discipline indépendante d'éléments subjectifs ou arbitraires, comme les probabilités *a priori* de BAYES — et il est assez plaisant de constater que c'est en s'assignant un tel but, que nous savons aujourd'hui littéralement insensé, que de grands statisticiens tels que Karl PEARSON, Sir Ronald A. FISHER, Jerzy NEYMAN (pour ne citer que lès noms les plus justement célèbres) ont pu, en effet, échafauder les principaux chapitres de la statistique mathématique, et donner à cette science une impulsion telle qu'elle a connu, au cours des dernières décennies, d'immenses succès théoriques et pratiques, et qu'elle est devenue indispensable dans la plupart des domaines d'activités scientifiques.

* * *

Mais les différents problèmes de la statistique mathématique — estimation des paramètres, tests d'hypothèses, analyse de la variance, plans d'expériences, statistiques d'ordre, etc. — pouvaient apparaître comme une collection de problèmes ayant certes en commun ce trait fondamental d'impliquer une induction, des observations au modèle, mais pour le reste assez différents entre eux, et presque hétéroclites, en ce qui concerne leurs formalisations respectives. Il a fallu attendre l'œuvre d'Abraham WALD pour voir se réaliser l'unité de la statistique mathématique, sous la forme de la *Théorie des Fonctions de Décision statistique,* dont les divers chapitres que nous avons énumérés à l'instant apparaissent comme des cas particuliers. C'est une étape un peu moins généralement connue, et nous nous y arrêterons un instant, pour décrire d'une façon sommaire la théorie proposée par WALD. L'intérêt est double à nos yeux : montrer l'unité des problèmes de la statistique mathématique et mettre en lumière la nécessité de fondements logiques que l'on ne peut trouver que dans une autre théorie, dont nous dirons quelques mots *in fine.*

Un problème statistique comporte les éléments formels suivants :

L'ensemble des observations, X, est l'ensemble de tous les « points » x qui constituent tous les résultats d'observations possibles *a priori,* pour le phénomène étudié. Naturellement, x peut être un nombre, un vecteur, une fonction du temps, ou tout autre élément que l'on voudra.

Les diverses lois de probabilité prises en considération, F, forment un ensemble Ω.

Ce sont des distributions de probabilité sur X, et elles peuvent être en nombre fini, ou bien former une famille paramétrique, ou être définies de quelque autre façon.

Le statisticien doit, au vu des observations, prendre une décision. Les décisions possibles d forment un ensemble D. Dans certains cas, on peut être amené à prendre en considération le tirage au sort entre plusieurs décisions, avec des probabilités choisies au mieux : cela signifie qu'on prend en considération l'ensemble D^* des distributions de probabilité sur D. Du point de vue qui nous occupe ici, cela constitue un détail technique, dont nous ne nous encombrerons guère par la suite.

Le problème posé est de choisir une décision d chaque fois qu'on disposera d'observations x, c'est-à-dire de choisir une manière de faire correspondre un d à tout x; en d'autres termes, nous devons envisager les applications de X dans D. Une telle application δ, constitue ce que WALD appelle une fonction de décision, et nous noterons Δ l'ensemble des fonctions de décision δ.

Maintenant, en vertu de quelles considérations peut-on être amené à choisir une fonction de décision plutôt qu'une autre? Si l'on prend une décision d, alors que la loi qui représente le mieux le phénomène étudié — ce que nous appellerons conventionnellement la « vraie » loi — est F, alors on encourt certains avantages et certains désagréments. Admettons, avec WALD, que ces conséquences peuvent être représentées valablement par un nombre, que nous appellerons la « perte » : ce nombre est donc une fonction de F et de d, que nous représenterons par

$$W\ (F,\ d) \qquad \text{(fonction de perte)}$$

Dans certains problèmes de contrôle des fabrications, ce concept peut être rattaché à une perte au sens ordinaire, chiffrée en argent (ce sera quelque chose comme l'espérance des divers coûts actualisés résultant d'une décision de rejet ou d'acceptation d'un lot de fabrication). Mais dans la plupart des problèmes statistiques, il faut admettre que la fonction de perte représente, forfaitairement, des inconvénients de toute nature; on verra plus loin qu'il suffit de lui attribuer des formes extrêmement simples pour retrouver et, dans une certaine mesure, justifier les techniques classiques de la statistique.

On doit choisir, avons-nous dit, non pas une décision mais une fonction de décision : la fonction de perte n'est qu'un intermédiaire, et l'on attachera à toute fonction de décision δ l'espérance de la perte correspondante, ce que nous écrivons :

$$r\ (F,\ \delta) = \int W\ (F,\ d) \qquad dF\ (x)$$

et que nous appellerons, en reprenant toujours la terminologie de WALD, fonction de risque. Le sens de ce nouveau concept peut être rendu plus explicite; à une distribution de probabilité F sur X, la fonction de décision δ, qui est une application de X dans D, associe une distribution de probabilité sur D, et, en conséquence, également une distribution de probabilité pour W : d'où la possibilité d'en calculer l'espérance, qu'exprime bien la formule ci-dessus.

La considération de la fonction de risque permet-elle de choisir une règle de décision? Ce sera vrai dans des cas exceptionnels où il existera une fonction de décision δ rendant minimum le risque quel que soit F. Mais en général, on peut s'attendre à ce que la règle de décision, qui minimise le risque, dépende de F. Avant de signaler comment on peut surmonter cette difficulté, illustrons les notions que nous venons d'introduire, en examinant quelques problèmes statistiques classiques.

***

On connaît le problème de l'estimation d'un paramètre : la famille Ω des lois de probabilité est par exemple une famille à un paramètre, F (x, θ) et on veut, d'après des observations, assigner une valeur approximative à θ. C'est bien un cas particulier des problèmes que nous venons de décrire. Une décision est constituée par la valeur t qu'on attribuera au paramètre, une règle de décision n'est pas autre chose qu'une fonction $t(x)$ — c'est ce qu'on appelle en statistique mathématique un estimateur.

Supposons que la fonction de perte admise soit proportionnelle au carré de l'écart entre la valeur vraie du paramètre et sa valeur estimée :

$$W = k\ (\theta - t)^2$$

Alors la fonction de risque sera, dans le cas d'un estimateur sans biais :

$$r \quad \theta, t(x) \quad = \int k\ (\theta - t)^2\ dF\ (x) = k \text{ var } t$$

Minimiser la fonction de risque, c'est donc ici rechercher les estimateurs de variance minimum : on voit comment on pourra justifier la méthode du maximum de vraisemblance et quelques autres méthodes classiques d'estimation.

Considérons maintenant un problème de test — et pour simplifier l'exposé, nous admettrons qu'on s'intéresse exclusivement à des modèles entièrement déterminés : on veut tester une hypothèse simple F, contre une alternative simple, G. L'ensemble X étant au reste quelconque. Ω est formé des deux lois F et G. L'ensemble des décisions possibles comporte aussi deux éléments, à savoir :

d_1 : conserver l'hypothèse F
d_2 : rejeter l'hypothèse F

Dès lors, une fonction de décision est simplement une partition de X en deux sous-ensembles — ou encore, c'est la donnée d'une région critique, ou région de rejet.

Prenons pour fonction de perte la fonction représentée par le tableau ci-dessous :

	F	G
d_1	0	1
d_2	1	0

Autrement dit, la perte est nulle quand on prend une décision correcte, égale à l'unité quand on commet une erreur. On calcule sans peine la fonction de risque, qui vaut α pour F et β pour G, α désignant le seuil de signification du test, et β la probabilité d'erreur de seconde espèce. Là encore, nous retrouvons tout naturellement des notions classiques.

Dans l'exemple de l'estimation — et du moins pour la plupart des formes de lois de probabilité d'usage courant — l'estimateur de variance minimum, ou l'estimateur de FISHER, ne dépend pas de la valeur du paramètre que l'on cherche à estimer : nous sommes dans ce cas privilégié où il existe une fonction de décision uniformément meilleure que toutes les autres — cette proposition étant vraie sans restriction lorsqu'il existe un estimateur exhaustif, et n'étant vraie qu'asymptotiquement dans d'autres cas.

Nous constatons qu'il en va tout différemment dans l'exemple des tests; même dans les cas les plus simples, il faut choisir un compromis entre les deux risques d'erreurs α et β : on sait bien qu'il n'existe pas de test qui rende à la fois α et β aussi petits que possible.

C'est la difficulté déjà signalée plus haut : la fonction de risque dépend des deux arguments F et δ. Il faut décider d'un principe de choix.

WALD a suggéré qu'on pourrait trancher cette ultime difficulté en adoptant le principe du minimax : la fonction de décision retenue serait celle qui réalise la condition :

$$\min_{\delta} \max_{F} r(F, \delta)$$

C'est une règle de prudence extrême, et l'analogie formelle entre les problèmes de la statistique et ceux de la Théorie des Jeux à deux joueurs a pu rendre cette règle fort séduisante. Mais il convient de noter que la justification du minimax dans la Théorie des Jeux réside essentiellement en ceci, que le minimax conduit à un point d'équilibre : si l'un des joueurs s'écarte de la stratégie optimale, il sera pénalisé. Dans les problèmes statistiques, l'un des joueurs est le statisticien, l'autre est la « nature », qui est censée choisir la loi de probabilité F inconnue. On ne peut guère considérer sérieusement que les intérêts de la nature sont opposés à ceux du statisticien — et le minimax perd, de ce fait, toute espèce de justification.

WALD a étudié par ailleurs une catégorie de fonctions de décision dont il a montré l'importance : ce sont les fonctions de décision de BAYES, c'est-à-dire celles qui réalisent une condition du type

$$\min_{\delta} E_{\xi} r(F, \delta)$$

ξ étant une distribution de probabilité sur l'ensemble Ω des lois de probabilité F (ξ est appelée habituellement « distribution de probabilité *a priori* »). WALD a montré que dans des cas très généraux, les fonctions de décision de BAYES sont les seules fonctions de décision admissibles — c'est-à-dire telles qu'il n'existe aucune autre fonction de décision dont le risque serait plus faible pour tout F. Aux yeux de WALD, il s'agit là d'une propriété purement formelle et la distribution de probabilité *a priori* n'est qu'un élément formel intermédiaire, permettant de sélectionner une classe de fonctions de décision intéressante : c'est que la statistique était encore fortement tournée, il y a une dizaine d'années, vers la recherche de principes objectifs.

Aujourd'hui, un nombre croissant de statisticiens admettent que la recherche d'un principe de choix objectif est vaine. Les problèmes de décisions statistiques dont nous venons d'esquisser sommairement les contours, entrent dans la catégorie des problèmes de décision en face d'incertitudes, et la *Théorie Générale des Décisions* nous apprend que des choix cohérents — ou rationnels — la définition précise de ces adjectifs peut être traduite dans des postulats aussi convaincants qu'on peut le souhaiter — de tels choix sont nécessairement ceux que l'on commettrait en maximisant l'espérance d'une fonction d'utilité des conséquences — ou, ce qui revient au même, en minimisant l'espérance d'une fonction de perte convenablement choisie. L'espérance doit être prise par rapport à une distribution de probabilité traduisant tous les éléments d'incertitude sur lesquels n'influe pas la décision du statisticien — et en particulier des probabilités doivent être affectées aux éléments de l'ensemble Ω des modèles.

C'est dire que la seule façon de donner un fondement logique solide à la statistique, semble bien être de considérer que les distributions *a priori* ξ doivent représenter un certain degré de connaissance, ou de confiance, préalable, à l'égard des diverses loi F.

Cela, c'est en quelque sorte le droit le plus général. Dans la pratique faut-il recommander de traduire toujours les connaissances *a priori* par une distribution de probabilité, et considérer que tous les efforts de la statistique classique — celle qu'on appelait moderne il y a dix ans — doivent être passés par pertes et profits? Certains n'hésitent pas à l'affirmer.

Personnellement, nous admettons sans réserve que la Théorie Générale des Décisions, ou encore Théorie de la Probabilité-Utilité, constitue le seul modèle logique valable. Mais le choix effectif des probabilités *a priori* éveille dans beaucoup de situations pratiques des scrupules trop vifs pour qu'on puisse penser qu'ils sont dénués de fondement. Il convient de juger en connaissance de cause et, dans chaque cas particulier, quels inconvénients pèsent le plus lourd, de ceux d'un choix hasardeux des probabilités *a priori*, et de ceux d'une règle de choix conventionnelle, permettant de s'en passer. Heureusement, dans les problèmes les plus courants de la statistique, on peut reconnaître que ce dilemne est facilement tranché. Nous avons eu l'occasion de constater que le problème de l'estimation des paramètres donne lieu, sous certaines conditions, à une règle de décision optimale évidente. On peut montrer, d'autre part, que dans beaucoup de problèmes, de tests, des hypothèses très généralement raisonnables concernant à la fois des probabilités *a priori* des diverses hypothèses et les coûts des diverses erreurs, permettent de justifier les méthodes classiques — en montrant que la règle optimale, au sens de la Théorie des Décisions, s'écarte assez peu de la convention courante qui consiste à choisir des tests ayant un niveau de signification faible. Bien entendu, cette propriété n'est nullement générale, c'est une des raisons pour lesquelles il nous semble nécessaire que les Ingénieurs, de même que les autres utilisateurs possibles de la statistique, soient informés des rapports entre les techniques statistiques et la Théorie de la Décision; nous avons essayé, d'une façon beaucoup trop succincte sans doute, d'en donner seulement les grandes lignes. Peut-être eût-il été bon d'esquisser avec plus de précision le contenu de la Théorie Générale des Décisions; mais ceci est une autre histoire...

Note bibliographique

A. Wald — Statistical Decision Functions. Wiley, 1950. J.-L. Savage — The Foundations of Statistics. Wiley, 1954. Concernant la Théorie Générale des Décisions, on pourra consulter :

La Décision — Colloques Internationaux du C. N. R. S., 1961. R.-D. Luce et H. Raiffa — Games and Decisions. Wiley, 1957. G. Morlat — Des Poids et des Choix. Mathématiques et Sciences Humaines n° 3, ou Bulletin S. E. D. E. I. S., Futuribles, n° 26.

Georges Morlat

Conseiller scientifique
à la direction des Études économiques générales
de l'E. D. F.

DISCUSSION

M. le Président Delaporte (Président de la Société de Statistique de Paris). — Je voudrais tout d'abord vous remercier, Monsieur le Président, de l'accueil si aimable que vous faites aujourd'hui à la Société de Statistique de Paris dans cette salle et je suis particulièrement heureux de ce contact entre la Société des Ingénieurs civils de France et la Société de Statistique de Paris. Je suis convaincu que notre collaboration d'aujourd'hui sera particulièrement fructueuse.

Avant de remercier les orateurs qui nous ont donné des indications aussi précieuses sur quelques-uns des grands problèmes de la statistique, je voudrais demander tout de suite à ceux d'entre vous qui ont quelques questions à poser de bien vouloir nous les exprimer.

M. GIRSCHIG. — Après les très intéressants exposés que nous venons d'entendre, je voudrais faire une remarque plus terre à terre mais qui se situe bien dans l'état d'esprit d'un ingénieur. Enseignant la Statistique à l'École Centrale, je pense qu'il est, en effet, indispensable d'attirer l'attention des élèves ingénieurs sur les hypothèses qui sont à la base des modèles mathématiques correspondant aux divers plans d'expérience étudiés. L'ingénieur peut être un peu effrayé par la complexité des méthodes qui lui sont proposées, d'autant que le choix est parfois difficile et qu'il existe des cas où les plans d'expérience classiques sont physiquement inapplicables. Ainsi, dans un carré latin, l'élimination des différences éventuelles de fertilité selon les lignes et selon les colonnes n'est possible qu'à la condition que ces différences puissent s'exprimer par des termes additifs dont l'un correspond à la ligne et l'autre, à la colonne se croisant sur l'une des cases du tableau. En d'autres termes, il faut qu'il n'existe aucune interaction entre les effets des lignes et des colonnes ou entre les traitements et les différences de fertilité du terrain. Dans les applications industrielles ces hypothèses ne sont pas obligatoirement vérifiées et je pense qu'il est indispensable qu'elles soient toujours présentes à l'esprit des ingénieurs qui se proposent d'utiliser ce type de plan d'expérience.

En ce qui concerne la conférence de M. GIBRAT, je me permets de faire remarquer que, si le sexe d'un enfant est bien déterminé deux mois avant sa naissance, ce sexe demeure cependant, pour tout observateur humain, un événement aléatoire. Dès lors, s'il a pu être prouvé par des observations antérieures qu'il existe une corrélation significative entre le sexe des petits-enfants et le comportement de leur grand-mère antérieurement à leur naissance, il me semble que le refus de M^me^ GIBRAT de boire de l'alcool peut conduire à pronostiquer, avec plus de 50 chances sur 100 de succès, qu'elle aura la joie d'être grand-mère du petit-fils qu'elle désire. Je serais heureux de savoir si mon interprétation est conforme à la pensée de M. GIBRAT.

M. GIBRAT. — J'avais voulu montrer par l'exemple de mon petit-fils toute l'importance de la réflexion préliminaire. Mais je vais cependant faire part à ma femme de votre réflexion et je ne sais si cela va la faire renoncer à ses idées; la connaissant bien, je crains que non.

M. LE PRÉSIDENT AUBERT (Président de la Société des Ingénieurs civils de France). — Je n'avais pas l'intention de demander la parole, mais à la suite des exposés si intéressants que nous venons d'entendre, je serais désireux de poser quelques questions.

M. GIBRAT a signalé la concordance profonde des problèmes d'hydrologie avec le calcul des probabilités et la statistique. Je ne puis qu'être d'accord avec lui sur ce point, et pendant 30 ans, j'ai effectivement parlé, dans mon cours aux futurs ingénieurs des Ponts et Chaussées, de la loi de GIBRAT et de son application à la détermination des dimensions des évacuateurs de crue accolés aux grands barrages.

Il a parlé de la nécessité de ne pas avoir plus d'une chance sur 1 000 de voir le barrage emporté. Dans ce domaine des chiffres, je pense qu'il faudrait préciser que cette destruction éventuelle doit être rendue invraisemblable pendant toute la durée d'un ouvrage destiné à être utilisé pendant 100 ans et peut-être même pendant 1 000 ans. Si l'on veut ne pas avoir plus d'une chance sur 1 000 de voir le barrage emporté pendant toute la durée de son utilisation, il faut évidemment que l'évacuateur de crue soit capable de laisser passer sans catastrophe la crue susceptible de se produire, en moyenne, un fois par million d'années.

Incidemment, je me permettrai de signaler qu'avant l'hydraulique, une autre technique a été, elle aussi, l'occasion de travaux intéressants de statistiques et de calculs de probabilité. Il s'agit de la technique du tir des canons, à propos de laquelle les grands artilleurs

français, et je citerai parmi eux le colonel HENRY, ont, avant la guerre de 1914, bâti de belles théories sur « les grandeurs éventuelles ». En entendant dire tout à l'heure qu'EULER s'était trompé, j'en venais à me demander si Joseph BERTRAND, qui a enchanté ma jeunesse, ne s'était pas, lui aussi, trompé dans sa théorie de la décision. Je m'explique. Ayant admiré la grande généralité et la beauté de l'exposé de M. MORLAT, j'ai été étonné de constater que la probabilité *a priori* ne jouait aucun rôle, cette expression ayant bien été prononcée mais une seule fois. Je rappellerai donc que les grands anciens, POINCARÉ et Joseph BERTRAND, en étudiant la théorie de la décision, ont estimé que la probabilité *a posteriori* était indéterminée tant que l'on ne faisait pas une hypothèse sur la probabilité *a priori*. Si l'on pouvait éliminer la considération de celle-ci, on aurait montré que Joseph BERTRAND avait tort.

Ce problème de décision correspond à notre vie de tous les jours et habituellement, sans le savoir, nous sommes amenés à en résoudre un grand nombre. Ils se présentent habituellement sous une forme très simple, analogue à l'exemple sur lequel je voudrais raisonner.

Vous prenez un certain nombre de cobayes auxquels vous inoculez une certaine maladie, et, à la moitié d'entre eux, vous appliquez également un remède. Vous comparez ensuite la mortalité des deux groupes. Tout au moins si les deux pourcentages de mortalité ne sont pas fondamentalement différents, que conclure sur l'efficacité du remède? Si celui-ci est constitué par de l'eau distillée (probabilité *a priori* faible de constituer un bon remède), la probabilité *a posteriori* de son efficacité sera, elle aussi, faible. Si, au contraire, le remède est un vaccin préparé suivant les règles de l'art, la probabilité *a posteriori* de son efficacité sera beaucoup plus forte.

Je pense que ceux des probabilistes d'aujourd'hui qui ne connaissent pas les très belles pages que Joseph BERTRAND a consacrées à ce problème prendraient beaucoup d'intérêt à cette lecture.

M. MORLAT voudra sans doute bien me dire si Joseph BERTRAND s'est effectivement trompé.

M. GIBRAT. — Ce n'est pas moi qui suis en cause, mais M. DUGUÉ. Pourriez-vous exposer le théorème exact pour répondre à M. AUBERT?

M. DUGUÉ. — Prenons un carré latin de 4 cases de côté, ce qui fait 16 résultats. La variance totale se répartit entre 4 variances différentes : une variance résiduelle de 6 degrés de liberté, une variance par ligne, une par colonne et une par traitement, chacune de ces trois dernières ayant 3 degrés de liberté. Dans le cas de ce que l'on appelle l'hypothèse nulle, toutes ces variances indépendantes et le quotient de la variance par traitement par la variance résiduelle permet par comparaison avec les valeurs fournies par les tables de SNEDECOR de de tester cette hypothèse nulle.

M. GIRAULT. — Je voudrais ajouter une remarque pour répondre à une ou deux questions qui viennent d'être posées. Chaque fois qu'on fait une expérience (qu'elle soit de type déterministe ou de type statistique) on met en doute une hypothèse qui n'est qu'une partie des conditions admises. Dans le cas de la statistique, il y a un ensemble d'hypothèses qui ne sont pas mises en cause par le test, et une partie supplémentaire des hypothèses qui, elle, est mise en doute par le test. Cette remarque ne s'applique pas seulement à la statistique, mais à toute expérience. Dans toutes les disciplines on a l'habitude d'admettre un certain nombre de résultats établis antérieurement, car chaque génération ne peut pas refaire la science depuis le début. On fait confiance sur un certain nombre de points qui sont admis et on étudie un petit quelque chose supplémentaire. Naturellement, il arrive qu'un résultat inexact soit

admis pendant de longues périodes. Pour éviter ce risque on ne peut pas, à propos de chaque expérience faite, mettre en cause tous les fondements de la science.

M. MORLAT. — Il est vrai que sur le sujet des probabilités *a priori* mon exposé était un peu bref, et risquait d'être mal interprété, aussi bien l'intervention de M. le Président AUBERT était-elle justifiée; il aurait fallu ajouter quelques mots pour montrer comment se placent les discussions du temps de Joseph BERTRAND. Il s'agissait le plus souvent de répondre à cette fameuse question : « Comment faut-il choisir les probabilités *a priori* de la formule de BAYES lorsqu'on ne sait rien? » Comme souvent, dans des cas semblables, les problèmes qui ont fait couler beaucoup d'encre étaient sans solution et ils ont été résolus parce qu'on a dépassé le problème; les résultats de la théorie générale de la décision auxquels j'ai fait allusion répondent à une question un peu différente, ils répondent à une question de structure des décisions, mais ne donnent pas des règles précises pour déterminer les probabilités *a priori* dans tel cas particulier. Cela entre dans le cadre d'une remarque qu'a faite M. GIRAULT, selon laquelle les résultats fondamentaux de la théorie des décisions comme du reste de toute théorie, sont des résultats de structure. Dans le cas présent la théorie nous enseigne que les seules règles cohérentes sont des règles de BAYES : tout se passe comme si, lorsque l'on est cohérent, on avait une probabilité *a priori;* cela ne montre pas comment choisir cette probabilité *a priori* dans telle ou telle situation concrète.

LAPLACE et Joseph BERTRAND recommandaient l'emploi de la formule de BAYES, et cette recommandation était fort arbitraire. C'est pourquoi on a cherché d'autres solutions. On a montré que, sous des hypothèses fort générales, il n'y en a point d'autres. Aussi bien, lorsque le théoricien moderne parle de la formule de BAYES, sa position ne doit pas être tout à fait confondue avec celle des anciens; car il sait dans quelle mesure la formule de BAYES est une nécessité logique. Du même coup, la question de savoir comment doivent être choisies les probabilités *a priori* et si on peut se passer de les évaluer dans certains cas, se présente sous un jour assez différent.

M. DELAPORTE. — Je voudrais, pour terminer, remercier vivement nos orateurs, MM. GIBRAT, GIRAULT, DUGUÉ et MORLAT de leurs très intéressants et très brillants exposés. Ils nous ont introduits progressivement dans des problèmes de plus en plus compliqués concernant la méthode de statistique et la statistique mathématique, je les en remercie vivement ainsi que ceux qui ont bien voulu participer à cette discussion.