

ANTOLOGIA

MATHÉMATIQUES ET BIOLOGIE *

par G. Teissier

L'organisation actuelle des études secondaires et supérieures en France est telle qu'un futur médecin ou un futur naturaliste fait son dernier problème le jour de son baccalauréat et que c'est à cette même date que le futur ingénieur, le futur physicien, le futur chimiste, prennent définitivement congé des sciences biologiques.

Je ne suis pas sûr que cette situation soit absolument sans inconvénient pour ceux qui cultivent les sciences exactes ou les sciences appliquées. Je suis certain, en revanche, qu'elle interdit à de jeunes biologistes l'accès de disciplines modernes et vivantes, qu'ils sont condamnés à ignorer leur vie durant, et qu'elle est responsable, pour une large part, de la décadence des recherches médicales dont s'alarment actuellement les meilleurs de nos collègues des facultés de médecine.

Le titre donné à cette conférence couvre un si large sujet qu'il ne me sera pas possible de faire mieux que d'en donner une esquisse et de faire entrevoir à ceux qui voudront bien me suivre, l'existence de domaines de la biologie dont l'entrée n'est permise qu'à ceux que ne rebute pas la rigueur des mathématiques. Puisse-t-elle éveiller, chez quelques jeunes biologistes, la nostalgie de ces terres interdites et apprendre à quelques jeunes mathématiciens qu'ils peuvent trouver dans la Biologie mathématique un champ de recherches comparable à celui qu'ont trouvé leurs devanciers dans la Physique mathématique.

La mesure en Biologie

Il n'est guère de chapitre de la Biologie où n'interviennent aujourd'hui *mesures* ou *dénombrements*. S'il entrerait dans un laboratoire moderne, un biologiste du siècle dernier serait, sans doute surpris d'y trouver un outillage qui, de son temps, était réservé au physicien ou au chimiste. Nous avons en effet emprunté à nos collègues leurs procédés de mesure pour les appliquer à l'étude des êtres vivants: le physiologiste définit l'état du système nerveux par la vitesse de réaction à l'excitation électrique, le médecin vérifie l'état de la thyroïde en mesurant la consommation d'oxygène, le

zoologiste définit races, variétés, parfois même espèces, par les dimensions de tel ou tel de leurs organes; le psychologue cherche à définir les aptitudes intellectuelles par des tests quantitatifs variés.

La vie courante même nous montre cette utilité de la mesure. Chacun de nous s'est alarmé ou rassuré suivant l'indication lue sur un thermomètre médical. Nous avons tous eu, dans notre premier âge, une feuille de pesée où une ligne sinueuse a fait, tantôt la fierté, tantôt le désespoir de notre mère. En quoi l'allure de cette ligne peut-elle expliquer, sinon justifier, cette joie ou cette alarme? Pourquoi l'écart plus ou moins marqué, dans un sens ou dans l'autre, de la ligne joignant les points figuratifs des pesées journalières et d'une courbe ascendante merveilleusement régulière, que d'avisés commerçants impriment sur les feuilles de pesées, intéresse-t-il parents et médecins? Que signifie au juste cette courbe que l'on prétend définir la croissance normale d'un enfant; que signifient de même ces tailles et ces poids correspondant aux divers âges qu'indiquent les bascules des gares?

Questions auxquelles bien peu de gens, à commencer par les médecins, seraient capables de répondre; questions d'ailleurs qu'ils ne se sont jamais posées, tant l'esprit scientifique fait défaut à l'immense majorité de nos concitoyens.

Quoi qu'il en soit de ce dernier point, sur lequel il y aurait, hélas! beaucoup à dire, nous sommes précisément ici pour nous poser des questions de ce genre. Le fait qui importe est que, dans un nombre toujours plus grand de recherches, le biologiste est appelé par le jeu même de son travail, à recueillir des séries de chiffres qui lui permettent, soit de suivre pas à pas un phénomène, soit de caractériser une population. Il se trouve fréquemment qu'une fois ses mesures faites et bien faites, le biologiste se trouve embarrassé pour interpréter correctement et qu'il doive recourir à l'aide d'un spécialiste.

Essayons de comprendre comment il se peut faire qu'il ait été nécessaire d'élaborer tout un corps de techniques mathématiques, la *Biométrie*, pour interpréter les plus simples des mesures biologiques.

* Conferência da série «Enseignement et Culture» promovida pela «Union Française Universitaire» e realizada em Paris em 13 de Abril de 1945. O Prof. Georgis Teissier, da Faculdade de Ciências da Universidade de Paris, seu Autor, procura elevar em França o ensino da Biologia e sobretudo o da Genética ao nível que lhe compete. Lamentamos também, com o seu Autor, a necessidade de realizar uma conferência deste tipo na época actual e transcrevemo-la na nossa revista por julgarmos infelizmente que a situação entre nós é ainda mais grave.

L'analyse de la variabilité et la loi normale

Le Biologiste se heurte à chaque instant dans ses recherches à une difficulté qu'ignorent presque complètement ceux qui ne s'attachent pas à l'étude des êtres vivants. Lorsqu'un chimiste, isolant un corps nouveau, en détermine la densité, la température d'ébullition ou l'indice de réfraction, il ne doute pas que les chiffres obtenus soient des *constantes* caractéristiques du corps étudié. Si celui-ci a été obtenu à un état de pureté convenable, si le chimiste est suffisamment habile et s'il dispose d'instruments suffisamment précis, les résultats de ces mesures pourront être retrouvés, en tout temps et en tout lieu, par tout autre chimiste qui voudra s'en donner la peine. La température de fusion de la glace et la température d'ébullition de l'eau sous pression donnée sont partout les mêmes et peuvent être déterminées à partir d'un échantillon d'eau pure tout à fait quelconque. Il en est de même dans tous les chapitres de la physique ou de la chimie. Les conditions d'une expérience peuvent être reproduites assez exactement pour qu'on puisse être assuré d'observer chaque fois qu'on le voudra le même phénomène. La précision des résultats n'est limitée que par le degré de perfection des appareils dont on fait usage et il est toujours possible d'étudier deux objets pratiquement identiques.

Rien de tel lorsque l'objet d'étude est un être vivant. Il n'est rien de plus difficile que de reproduire exactement un phénomène biologique que l'on a observé une fois ; il est impossible de trouver deux organismes qui soient vraiment identiques.

A cette *variabilité*, aucune espèce animale ou végétale n'échappe. Dans une collection d'animaux ou de végétaux de même espèce, de même race, de même sexe, de même âge, choisis aussi semblables que possible, il n'en est pas deux qui puissent être considérés comme réellement identiques, dès que l'on cesse de se contenter d'une approximation grossière. Les moutons d'un troupeau indiscernables l'un de l'autre à l'oeil d'un citadin ont chacun une physiologie propre pour le berger qui vit avec eux. Quelle que soit l'espèce étudiée, quel que soit l'organe ou la fonction sur lesquels on fait porter les mesures, il existe des différences individuelles non négligeables. Une étude comparative de deux individus montrera toujours des différences entre les dimensions de chacune des parties de leur corps, des différences dans la composition chimique ou dans les caractéristiques physiques de chacun de leurs tissus.

Ainsi, au seuil même de la Biométrie, la variabilité des êtres vivants apparaît comme un grave obstacle, qu'aucun artifice expérimental ne pourra jamais tourner, qu'aucun progrès technique ultérieur ne permettra

de franchir puisqu'il appartient à la nature même de l'objet d'étude. En faut-il conclure qu'il est vain d'espérer faire de la Biologie une science quantitative et qu'elle est condamnée éternellement à l'imprécision et au vague ? Beaucoup de biologistes l'ont cru, beaucoup le croient encore. Mais d'autres n'acceptent pas de s'arrêter à cette solution paresseuse. Ils pensèrent puisqu'il n'était pas possible d'ignorer la variabilité, ni légitime d'en faire abstraction, qu'il fallait étudier le phénomène en lui-même. En abordant de front la difficulté, ils trouvèrent dans l'obstacle prétendu insurmontable, l'occasion des plus précieuses découvertes.

De cette étude de la variabilité, Darwin avait d'ailleurs montré la nécessité. Lorsque son «*Origine des Espèces*» eut fait du problème de l'Evolution la question centrale de la Biologie, il devint indispensable d'analyser de plus près ce qu'il n'avait pu le faire Darwin lui-même, les phénomènes qu'invoquaient ses théories. Au premier rang de ceux-ci est la variabilité qui différencie les individus d'une même espèce et leur donne des chances inégales de survivre à la Sélection naturelle. Aussi, dans l'étude de la variabilité, partisans et adversaires de l'Evolution s'affrontèrent-ils bientôt. Mais tous rivalisèrent dans la découverte de nouveaux faits.

Dans cet ensemble de travaux, deux courants se dessinent bientôt. Certains chercheurs, s'intéressant surtout à l'Homme, ne pouvaient pas disposer des ressources qu'offre la méthode expérimentale aux autres biologistes. Ils étudièrent la variabilité en elle-même, sans prétendre en analyser la signification biologique. Ils durent, pour interpréter les innombrables mesures qu'ils avaient faites, élaborer des techniques mathématiques nouvelles et créèrent ainsi les méthodes de la Statistique moderne, dont la Biométrie constitue la première codification.

D'autres chercheurs firent porter leurs études sur des animaux ou des végétaux élevés au laboratoire ou dans des champs d'expérience. Moins portés aux études mathématiques et plus «*naturalistes*» que leurs émules, ils s'attachèrent surtout à élucider le déterminisme de la variation. Après une longue période de travail confus et passablement désordonné, ils créèrent un des chapitres les plus éclatants de la Biologie moderne, la Génétique.

Nous verrons plus tard que la Génétique, science expérimentale, ne peut pas plus se passer de l'aide des mathématiques que la Biométrie, science d'observation, mais dès maintenant nous devons utiliser un de ses résultats.

L'étude expérimentale de l'hérédité a montré que les caractères individuels distinguant les divers représentants d'une même espèce, n'ont pas tous la même signification. Il en est qui sont inhérents à la nature même de chacun des êtres comparés, tandis que d'autres,

beaucoup plus contingents, sont imputables aux circonstances dans lesquelles chaque individu s'est trouvé placé au cours de son existence. Les premiers, qui sont héréditaires, peuvent porter sur les traits d'organisation les plus dissemblables, affecter la couleur aussi bien que la forme et porter sur d'infimes détails aussi bien que sur les dimensions globales de l'organisme. Les seconds, strictement individuels et non transmissibles des parents aux enfants, portent le plus souvent sur les dimensions du corps tout entier ou sur telle ou telle de ses parties.

Un exemple bien connu de tous permettra d'illustrer cette notion capitale. Une collection de chiens rassemblés au hasard, telle que celle que l'on peut observer dans une fourrière, présente dans la taille, la forme ou la couleur, beaucoup plus de diversité que n'en montreraient tous les fox-terriers du monde si l'on pouvait les réunir. Il s'agit là évidemment d'un cas extrême, mais l'on retrouverait partout, à un moindre degré, le phénomène qu'il met en évidence; la variabilité est moins forte dans une race que dans un mélange de races; elle est d'autant plus réduite que la pureté de la race est plus grande ou, pour employer le langage des généticiens, que le patrimoine héréditaire est plus homogène. A la limite, lorsque le patrimoine héréditaire est exactement le même pour deux individus, la ressemblance peut parfois presque atteindre l'identité. Cela ne se produit que dans un cas et un seul, celui des jumeaux vrais, qui, à la vérité, sont beaucoup moins deux êtres distincts qu'un seul être en deux personnes.

Hors ce cas très exceptionnel, il subsiste entre les divers représentants d'une même race des différences très appréciables. Elles s'expliquent pour une part, par de légères différences dans les patrimoines héréditaires des différents individus, mais surtout par la diversité des circonstances qu'ont traversées, au cours de leur existence, les individus que l'on compare. Ceux qui se sont trouvés dans les circonstances les plus favorables, ceux qui ont bénéficié de la nourriture la plus abondante et la meilleure, deviendront plus grands, plus forts et plus beaux que les autres. Quelque soin que prenne un expérimentateur pour rendre comparables les conditions d'existence des animaux qu'il élève, il ne réussira pas à faire disparaître ces différences individuelles, et on en retrouvera d'au moins aussi grandes dans les représentants de toutes les espèces sauvages. Ainsi, dans le cas le plus simple où puisse être tentée une étude biométrique, il subsiste toujours entre les divers individus utilisés dans une même série de mesures des différences morphologiques ou physiologiques très notables. Mais il se trouve, et c'est là le premier et l'un des plus importants résultats de la Biométrie, que cette variabilité des races

pures obéit à des règles précises et simples, règles qui permettent de comprendre et d'interpréter les faits que l'on observe lorsqu'on fait porter les mesures sur des populations renfermant un mélange de plusieurs races.

La découverte fondamentale en la matière date de près d'un siècle et est due au savant belge Quetelet, qui s'occupa avec un égal succès d'astronomie, de météorologie et de statistique. Etudiant sur une population très étendue les variations de la taille de l'homme adulte, il précisa par des mesures une constatation banale: les variations individuelles autour de la taille moyenne, qui est en même temps la plus fréquente, sont d'autant moins nombreuses, qu'elles sont plus amples. Les individus très grands ou très petits sont d'autant plus exceptionnels qu'ils sont plus grands ou plus petits. Familiarisé avec le calcul des probabilités et entraîné par ses recherches antérieures à l'élimination des erreurs expérimentales, il n'eut pas de peine à constater que la distribution des tailles suivait une loi exactement de même forme que celle qui régit la distribution des erreurs commises dans la mesure d'une grandeur. Il vérifia qu'il ne s'agissait pas là d'une simple analogie, mais que les mêmes formules s'appliquaient aussi correctement à l'un qu'à l'autre phénomène. Ainsi se trouvait introduite en Biométrie la loi qui devait jouer dans cette science un rôle si fondamental, qu'elle fut bientôt considérée comme «loi normale» de la variabilité.

Tout se passe, en somme, comme si tous les individus obéissaient en principe à la même loi de croissance qui, si elle agissait seule, aurait pour effet de donner à tous la même taille finale, et comme si, en même temps, les circonstances venaient troubler inégalement, dans un sens ou dans l'autre, et plus ou moins selon les individus, la marche régulière du phénomène. La taille finale serait la résultante de la superposition de plusieurs phénomènes: un principal, qui nous intéresse particulièrement, et d'autres très nombreux et très divers, que nous ne pouvons pas prétendre connaître tous. Les calculs statistiques faits sur l'ensemble des mesures ont pour but de dégager le phénomène principal des phénomènes parasites et d'évaluer en même temps l'importance de ces derniers.

Cette interprétation de la variabilité coïncide très exactement avec celle que donne la Génétique, dont nous avons parlé tout à l'heure. Elle permet en outre de comprendre le rôle de la loi normale en Biologie et donne ainsi une base solide à la Biométrie.

Grâce à la Biométrie, il est possible d'atteindre des grandeurs objectives et de les déterminer, avec une précision d'autant plus grande que le nombre des mesures a été plus grand. L'incertitude qui subsiste est toujours chiffrable. Mais ces grandeurs, qui ont une signification aussi claire que celles que l'on rencontre

en physique, ne concernent que des collectivités et, transportées à l'échelle individuelle, ne s'appliquent qu'à un être idéal, l'individu moyen, que nul n'a jamais rencontré. En ce qui concerne les individus réels, la Biométrie ne nous indique que des *valeurs probables*. Un biométricien ne peut pas deviner, plus que n'importe qui, que Monsieur X, qu'il n'a jamais vu, mesure 1 m. 74 et pèse 64 kg. Mais, si la fantaisie l'en prenait, il pourrait faire, à bon escient, certains paris. Il pouvait par exemple, très raisonnablement, parier en 1939, à trois contre un, que Pierre, petit Parisien de 12 ans, dont il ne savait rien, sinon qu'il était élève d'une école primaire du XV^e arrondissement, ne dépassait pas la taille de 1 m. 44 et le poids de 37 kg. 500; et, à dix contre un, qu'il n'atteignait pas la taille de 1 m. 49 et le poids de 45 kg.

Tout ce qui précède permet de comprendre que la réponse de la Biométrie à la plus simple des questions soit nécessairement assez compliquée. Elle comporte, au moins, l'indication de la *moyenne*, celle de l'*écart-type*, qui est une mesure de la variabilité, celle des *erreurs probables* de l'une ou de l'autre de ces grandeurs, erreurs qui dépendent à la fois de l'écart-type et du nombre des individus mesurés. Pour interpréter ces chiffres, il pourra être nécessaire de consulter des tables numériques, notamment celles de la fonction de Gauss, qui permettent seules de faire avec chance de succès certaines prédictions.

Ces constantes biométriques une fois connues, il devient possible de pratiquer de fort instructives comparaisons. On peut, par exemple, décider si deux séries d'individus peuvent légitimement être considérées comme appartenant à une seule et même population ou si, au contraire, on doit admettre qu'ils proviennent de deux populations distinctes. Des questions de ce genre se posent constamment dans l'étude des problèmes biologiques les plus divers, distinction des races au sein d'une même espèce, analyse de l'action du milieu sur les représentants d'une même race, détermination de l'importance des caractères sexuels secondaires lorsque ceux-ci sont peu marqués, orientation scolaire ou professionnelle... Et cette liste n'a rien de limitatif.

La corrélation

D'autres problèmes sont plus difficiles. Ce sont ceux qui exigent l'étude simultanée de deux ou de plusieurs grandeurs biologiques mesurées sur le même être. Dans une population d'enfants du même âge, le poids et la taille sont deux grandeurs variables, et tous les individus de même taille ne pèseront pas le même poids. On peut faire correspondre à chaque taille un poids, à chaque poids une taille qui représenteront des valeurs moyennes, caractérisant dans une certaine mesure la population étudiée. Mais ces renseignements

sont insuffisants. Ils ne nous diront pas si les enfants de même taille ont une corpulence moyenne et des poids très voisins l'un de l'autre ou si, au contraire, à taille égale, le poids est très variable, les individus très maigres ou très gras étant très nombreux. Ils ne nous diront pas, en somme, jusqu'à quel point les deux grandeurs étudiées dépendent l'une de l'autre, jusqu'à quel point pour un individu donné la connaissance de l'une peut servir à prévoir l'autre.

Ces renseignements complémentaires sont apportés par le calcul, à partir des résultats de l'ensemble des mesures, d'un paramètre spécial, le *coefficient de corrélation*, dont la valeur exprime le degré plus ou moins étroit de dépendance des deux grandeurs que l'on compare. La notion de corrélation, dont l'idée première est due au biologiste anglais Galton, est beaucoup plus facile à expliquer en langage mathématique qu'en langage usuel. Essayons cependant de donner une idée de son contenu.

La taille et le poids d'un enfant ne sont évidemment pas indépendants. Parmi les phénomènes qui déterminent la taille, il en est, notamment les dimensions du squelette, qui interviennent aussi dans la fixation du poids. Si ces phénomènes existaient seuls, le poids serait connu lorsqu'on connaîtrait la taille, exactement comme le poids d'une boule de fer est déterminé lorsqu'on connaît son diamètre. La variabilité d'une des grandeurs serait déterminée par les mêmes causes que la variabilité de l'autre. Mais nous savons bien que quantité de circonstances qui n'affectent pas une des grandeurs affectent l'autre, et qu'un enfant peut grandir sans augmenter de poids, maigrir ou engraisser sans que sa taille varie. Ces circonstances, qui troubles en quelque sorte le jeu des phénomènes communs à la croissance en taille et à la croissance en poids, nous empêchent de calculer avec certitude l'une des grandeurs lorsque l'on connaît l'autre. Elles ajoutent à la variabilité commune à la taille et au poids, une variabilité propre à chacune des grandeurs. Elles transforment la relation rigoureuse qui existerait sans elle en une relation moins stricte, une relation de corrélation. Le coefficient de corrélation donne une mesure de la part de la variabilité commune aux deux grandeurs dans la variabilité totale de l'une et de l'autre.

Nous pourrions redire, à propos de relations de corrélation, ce que nous avons dit en étudiant la mesure d'une grandeur isolée. Elles peuvent fournir des descriptions précises, exprimables par des formules aussi strictes que celles que l'on rencontre en physique, du comportement moyen d'une grande population. Mais elles ne donnent pour un individu particulier que des prévisions plus ou moins probables, dont le degré de précision dépend étroitement de la grandeur du coefficient de corrélation.

Lorsque ce coefficient est nul, les deux grandeurs sont indépendantes, et la connaissance de l'une n'ap-prend rien sur l'autre. Lorsqu'il est égal à 1, les deux grandeurs varient proportionnellement l'une à l'autre. Toutes les valeurs comprises entre 0 et 1 peuvent se rencontrer; elles expriment tous les degrés de liaison, du plus lâche au plus étroit, entre les grandeurs que l'on compare. Et, de même qu'il est possible de comparer deux populations du point de vue des moyennes et des écarts-types de leurs caractères mesurables, il est possible de les comparer du point de vue de la valeur des coefficients de corrélation unissant ces caractères. Nous verrons tout à l'heure un exemple où cette comparaison a apporté des résultats extrêmement importants pour la Biologie générale.

Biometrie et statistique

Il ne nous est, malheureusement, pas possible d'aller plus loin dans l'exposé des procédés de travail en usage dans la Biométrie et nous ne pouvons pas davantage songer à examiner les problèmes biologiques que l'utilisation méthodique de ces procédés a permis ou permettra de résoudre. Ce que nous avons dit doit cependant suffire à montrer qu'un des problèmes les plus fondamentaux de la Biologie, problème qui touche à l'une des caractéristiques les plus essentielles des êtres vivants, celui de la variation, est tout entier tributaire de méthodes mathématiques. Nous avons appris, à cette occasion, que l'on ne peut jamais connaître d'une grandeur caractéristique d'un être vivant que des valeurs probables et que, dans aucun cas, un seul nombre n'est capable de résumer complètement une enquête biométrique, différence d'avec les grandeurs physiques qui suffit à donner aux mathématiques du biologiste une allure toute particulière.

Il arrive, cependant, et il arrivera de plus en plus fréquemment dans l'avenir, qu'un biologiste ait à faire usage de techniques expérimentales et de modes de raisonnement analogues à ceux du physicien ou du chimiste et qu'il cherche à établir des relations quantitatives entre deux grandeurs biologiques, ou entre une grandeur biologique et une grandeur physique. Les lois qu'il obtiendra ne seront pas, en général, d'un type très différent de celui des lois que manient ses collègues des sciences exactes; mais leur signification concrète sera bien différente. Alors que, pour ceux qui étudient la matière inanimée, les lois obtenues peuvent permettre la prévision rigoureuse des phénomènes, pour celui qui étudie la matière vivante, la loi la plus rigoureusement établie ne sera jamais qu'une loi *moyenne*, capable de décrire le comportement *moyen* d'une population infiniment nombreuse, mais incapable par essence de traduire avec rigueur aucune loi indi-

viduelle. Nous savons bien qu'aucun bébé n'a jamais été assez idéalement moyen pour se conformer à la loi de croissance exprimée par la courbe admirablement régulière tracée par avance sur les feuilles de pesée, courbe qui ne peut traduire que le comportement d'un être parfaitement abstrait, le Bébé Normal. Mais nous savons aussi que la connaissance de cette courbe moyenne n'est pas inutile et que, si notre enfant s'en écarte trop, il nous faut appeler le médecin. Aujourd'hui chacun de nous s'alarme plus ou moins vite suivant son degré propre de calme ou de nervosité, mais il est fort probable qu'un jour prochain, la courbe théorique, qui existe seule aujourd'hui, sera encadrée de deux autres courbes qui représenteront des sortes de marges de sécurité qu'un enfant bien portant ne doit pas franchir en principe. Ces marges de sécurité seront définies par la biométrie, à partir de cet écart-type du poids moyen dont il a été fait mention tout à l'heure. L'ensemble des trois courbes, la courbe moyenne et ses deux satellites, représentera l'équivalent de la courbe unique correspondant au phénomène physique. Différence qui s'explique par le fait que la matière vivante est toujours variable et, par là même, indéfinissable, tandis que la matière inerte est définissable et constante — au moins en principe.

Mais il peut évidemment arriver que certains phénomènes physiques complexes, ou certaines substances chimiques instables, présentent une moindre constance que d'autres phénomènes plus simples, que d'autres substances plus stables. Das des cas de ce genre, il faudra faire usage de la Biométrie, bien qu'il ne s'agisse plus d'êtres vivants, mais alors on ne parlera plus de Biométrie, mais bien de Statistique.

Au sens le plus large du terme, la Statistique peut être conçue comme l'ensemble des méthodes mathématiques qui peuvent être utilisées pour décrire ou pour interpréter les résultats quantitatifs des expériences ou des observations. Elle a son fondement dans le calcul des probabilités, branche des mathématiques pures, nées au XVIII^e siècle, de l'étude des jeux de hasard. Mais elle ne s'est réellement développée qu'au XIX^e siècle et, fait bien digne de remarque, c'est en essayant de résoudre des problèmes biologiques qu'elle s'est constituée.

Il y a cent ans juste, avec QUETELET, se fondait la Biométrie, origine des notions d'écart-type de l'erreur probable. Il y a trois quarts de siècle qu'avec GALTON, qui étudiait les lois de l'Hérédité, apparaissait la notion de corrélation, fondement de la Statistique moderne. Aujourd'hui encore, devenue discipline autonome, elle trouve dans des problèmes biologiques les plus précieux des stimulants; c'est ainsi que depuis vingt ans s'est édifié tout un corps de doctrine destiné à préparer, de façon rationnelle, des plans d'expéri-

mentation en agriculture et à interpréter au mieux les résultats des expériences.

De cette branche moderne de la Statistique, il ne saurait être question ici et nous ne parlerons pas davantage, faute de temps, d'une application plus ancienne, la démographie. Je voudrais utiliser le temps qui me reste à indiquer très rapidement que les mathématiques ont joué un grand rôle dans le développement d'une des disciplines biologiques les plus essentielles : la Science de l'Hérédité.

Hérédité et Statistique

L'histoire de cette science est curieuse. C'est en 1865 que MENDEL découvrit les lois qui ont immortalisé son nom, mais cette découverte, la plus importante qui ait été faite en Biologie au cours du XIX^e siècle, passa tout d'abord inaperçue et ne sortit de l'oubli que trente-cinq ans plus tard. De sorte que, lorsque, vers 1870, FRANCIS GALTON aborda l'étude de l'hérédité, tout paraissait encore à découvrir sur le sujet. GALTON, étudiant l'homme, sur lequel on n'expérimente pas, et s'intéressant à des caractères comme la taille, dont la transmission obéit, on l'a su plus tard, à des règles particulièrement complexes, ne pouvait redécouvrir les lois de Mendel. Ce qu'il cherchait, c'était un moyen d'exprimer par un nombre, le degré de ressemblance moyen existant entre parents et enfants, frères et soeurs, etc... Et c'est au cours de cette recherche qu'il élabora, au prix de longs efforts, la notion de corrélation, qui, nous le savons déjà, est devenue essentielle à la Biométrie tout entière. Appliquée aux pedigrees qu'il rassembla patiemment, la méthode lui permit d'établir un résultat fondamental, la *loi de corrélation ancestrale* qui dit que le coefficient de corrélation entre individus apparentés en ligne directe ou en ligne collatérale, diminue de moitié à chaque génération nouvelle qui les sépare ; la corrélation père-fils est égale à la corrélation frère-soeur et deux fois plus forte que la corrélation grand-père-petit-fils, etc... Les méthodes de calcul, assez rudimentaires à l'origine, se perfectionnèrent bientôt, de sorte que, vers la fin du siècle dernier, l'étude quantitative de l'hérédité semblait avoir reçu son orientation définitive. Mais en 1900, les lois de Mendel furent redécouvertes et prirent dans la Biologie une place prépondérante, d'ailleurs bien justifiée. Les biologistes, peu amis des mathématiques en général, furent trop heureux d'avoir ce prétexte pour rejeter dans l'ombre l'oeuvre monumentale de GALTON. Certains prétendirent même que les résultats de l'école biométrique étaient faux, parce que soûdais incompatibles avec les résultats mendéliens. Des travaux ultérieurs devaient montrer qu'il n'en était rien et que, bien au contraire, la seule justification de la loi

de l'hérédité ancestrale repose sur les lois de Mendel. Le hasard a voulu que les conséquences complexes de principes très simples aient été découvertes avant que ceux-ci aient été compris. Il en eût été évidemment autrement si les biologistes contemporains de Mendel n'avaient pas été dépourvus de compréhension mathématiques au point de méconnaître totalement la grandeur du don que le génie du moine tchèque venait de faire à la science.

Ce qu'eut, en son temps, de profondément original la pensée de Mendel, c'est l'idée que l'hérédité était régie par les lois du hasard et qu'elle ne pouvait être étudiée utilement que par des méthodes statistiques. Une telle hypothèse ne pouvait que heurter profondément la pensée biologique traditionnelle, mais toute la science moderne en montre l'admirable fécondité.

Une des espèces les plus propres à faire comprendre la découverte de Mendel est la Belle-de-Nuit que l'on cultive assez souvent dans nos jardins, et qui présente plusieurs variétés parfaitement stables. Deux d'entre elles ne diffèrent que par la couleur de la fleur, toujours blanche dans l'une, toujours rouge dans l'autre. Lorsque le pistil d'une fleur rouge reçoit du pollen provenant d'une fleur de la variété blanche, les graines qui résultent de cette fécondation produisent des plantes dont les fleurs sont roses. Le croisement inverse donne les mêmes résultats : pollinisation d'une fleur blanche par du pollen de la variété rouge donne des produits dont les fleurs sont roses. Tous les hybrides de première génération sont donc identiques.

En deuxième génération, tout change. Les plantes issues des fleurs roses fécondées par leur propre pollen sont de trois sortes : les unes ont des fleurs blanches, les autres des fleurs rouges, d'autres encore, les plus nombreuses, des fleurs roses. Les deux premières catégories de plantes ont tous les caractères des plantes parentes, la troisième tous les caractères des hybrides de première génération. Le croisement de ces hybrides à fleurs roses, et des plantes à fleurs blanches donne un mélange de plantes à fleurs roses et de plantes à fleurs blanches ; celui d'une plante à fleurs roses et d'une plante à fleurs rouges un mélange de plantes à fleurs roses et de plantes à fleurs rouges : de tels hybrides ressemblent donc à l'un et l'autre de leurs parents.

Cette diversité d'apparence des hybrides de deuxième génération était connue avant Mendel et avait fait le désespoir de théoriciens de l'hérédité. Mendel, le premier, eut l'idée de dénombrer les différentes catégories d'individus et constata que, dès que la descendance d'un hybride était assez nombreuse pour qu'une statistique ait un sens, il apparaissait toujours en deuxième génération un quart de plantes semblables à celle qui avait fourni le pollen et un quart de plantes

semblables à celle qui avait apporté les ovules ; et le reste, c'est-à-dire la moitié du nombre total, était composé de plantes identiques aux hybrides de première génération. Dans le croisement d'un hybride de première génération avec une des plantes de race pure, on trouvait toujours un nombre égal de plantes semblables à chacun des deux parents.

Comment expliquer cette proportion remarquablement simple que l'on retrouve, avec la même constance dans les croisements les plus divers, non seulement chez les végétaux, mais chez les animaux ?

Remarquons tout d'abord que l'élément mâle et l'élément femelle, le grain de pollen et l'ovule, jouent le même rôle dans les phénomènes d'hérédité puisque les croisements symétriques donnent le même résultat et que l'un comme l'autre peuvent porter les caractères de couleur qui nous intéressent.

La constance des races parentes signifie évidemment que tous les grains de pollen d'un côté, et tous les ovules de l'autre, ne sont semblables entre eux — qu'ils sont tous « blancs » dans un cas, tous « rouges » dans l'autre. Il est difficile de comprendre l'existence des différentes catégories de la deuxième génération d'hybrides autrement qu'en admettant que les hybrides de première génération produisent, non pas comme on pourrait le croire, des éléments sexuels d'un type particulier que nous pourrions appeler « roses » mais bien un mélange de pollen et d'ovules « blancs » et « rouges ». Il est facile de montrer que les proportions observées sont exactement celles que l'on doit attendre si les éléments « blancs » et « rouges » sont en nombre égal.

Une certaine graine résulte de la fécondation d'un certain ovule par un certain grain de pollen et c'est le hasard seul qui a décidé que ce grain de pollen atteindrait cet ovule plutôt qu'un autre, le hasard seul qui fait qu'un pollen « rouge » féconde un ovule « blanc » ou « rouge ». Dès lors on peut prévoir aisément les résultats d'un croisement.

Représentons le grain de pollen et l'ovule par deux pièces de monnaie, de 1 franc et 2 francs par exemple. Convenons que le côté face représentera le facteur « rouge », le côté pile le facteur « blanc » et lançons nos deux pièces. Quatre combinaisons sont possibles qui sont, en citant dans l'ordre la pièce de 1 franc et celle de 2 francs, face-face, face-pile, pile-face, pile-pile ; autrement dit dans l'ordre pollen-ovule : rouge-rouge, rouge-blanc, blanc-rouge, blanc-blanc. Mais nous savons que rouge-blanc et blanc-rouge sont identiques et donnent l'un et l'autre des plantes à fleurs roses. Il ne reste donc que trois cas possibles : deux combinaisons représentées chacune une fois et correspondant aux deux races pures, une combinaison représentée deux fois et correspondant à l'hybride. Le calcul des probabilités nous apprend que si nous lançons nos

deux pièces un assez grand nombre de fois, les proportions des différentes combinaisons seraient sensiblement égales aux probabilités que nous venons de déterminer, c'est-à-dire à $1/4$, $1/2$ et $1/4$, proportions observées par Mendel. Les proportions $1/2$ et $1/2$ observées dans le croisement d'un hybride avec l'une ou l'autre des souches parentes s'expliquent plus simplement encore.

La précision que l'on peut atteindre dans la détermination de ces proportions mendéliennes, puisque c'est ainsi que l'on nomme ces rapports, est d'autant plus grande que la population étudiée est plus nombreuse, et les règles du calcul des probabilités permettent d'estimer la grandeur des écarts que l'on peut attendre dans des dénombrements portant sur un nombre déterminé de plantes. Les vérifications faites se sont toujours montrées assez satisfaisantes pour que l'interprétation de Mendel puisse être considérée comme entièrement justifiée.

Tous les croisements faisant intervenir un seul couple de caractères ne sont pas aussi faciles à interpréter que celui que nous venons d'étudier. Très souvent, le plus souvent même, l'hybride n'a pas une apparence intermédiaire entre les parents mais paraît identique à l'un d'eux, qui est alors dit *dominant* par rapport à l'autre que l'on dit *récessif*. Dans ces conditions, on doit s'attendre à observer en deuxième génération un quart d'individus récessifs et trois quarts d'individus à forme dominante, les hybrides se confondant avec les dominants purs, deux fois plus nombreux qu'eux. Ce sont ces proportions $1/4$ et $3/4$ que Mendel lui-même observa dans ses croisements entre variétés de poids.

Ainsi, la loi fondamentale de l'hérédité est une loi statistique, la plus simple des lois statiques, celle-là qui est à la base du calcul des probabilités ; la loi de pile ou face. La connaissance de cette loi suffit à faire des prévisions dans des cas beaucoup plus compliqués que celui que nous venons d'envisager. Elle permet, entre autres choses, de prévoir l'évolution d'une population qui, à un moment donné, contiendrait en proportions quelconques les trois catégories d'individus que nous avons distinguées, lorsque l'on fait sur le mode de pollinisation, sur la fertilité et sur la vigueur des trois catégories de plantes, diverses hypothèses.

Si l'on suppose que toutes les plantes sont également fécondes et vigoureuses, quelle que soit la couleur de leur fleur, on peut montrer que s'il y a autofécondation, le nombre des plantes à fleurs roses ira constamment en décroissant. Si la pollinisation est croisée et se fait complètement au hasard (pollinisation par le vent, par exemple), on obtiendra, en une génération, une population telle que le nombre des plantes à fleurs roses soit le double de la moyenne géométrique du

nombre des plantes à fleurs blanches ou à fleurs rouges, ces proportions restent ensuite stables de génération en génération. Dans une hypothèse intermédiaire, une fraction définie du nombre total de plantes étant soumise à l'autofécondation et le reste livré à la fécondation croisée, la population tendra graduellement vers un état d'équilibre stable, comportant toujours moins de plantes à fleurs roses que dans le cas de la fécondation entièrement croisée.

Lorsque l'on suppose que les plantes à fleurs rouges, roses ou blanches, n'ont pas la même fécondité, ni la même vigueur, hypothèses plus conformes à la réalité, les calculs deviennent beaucoup plus compliqués, mais aussi beaucoup plus intéressants. Ils permettent en effet d'aborder un problème capital, d'où dépend toute la philosophie de la Biologie et toute la philosophie de la Science — le problème de l'évolution. Etudier en effet le problème de la persistance ou de l'extinction, au sein d'une population, d'un facteur conditionnant un caractère qui donne à son porteur une fécondité ou une vitalité, accrue ou diminuée, c'est en effet ébaucher la théorie de la Sélection naturelle, clef de voûte du Darwinisme, seule explication rationnelle de l'Evolution.

C'est sur cette indication — qu'il est naturellement impossible de développer ici — que je voudrais conclure. J'ai essayé de montrer l'intérêt que présentent

les mathématiques pour la Biologie, mais on me permettra de regretter qu'il soit périodiquement jugé utile de faire cette démonstration. Il ne viendrait à l'idée de personne de demander une conférence de même niveau intitulée «Mathématiques et Physique». Ce n'est peut-être pas un très bon symptôme pour la qualité de la culture scientifique donnée par notre enseignement secondaire, qu'aux approches du milieu du XX^e siècle, une conférence sur «Mathématiques et Biologie» ait encore sa raison d'être. Ce stade est dépassé depuis vingt ans, trente ou plus aux U.S.A., en U.R.S.S. et ailleurs. Il n'est que temps de rattraper un retard inexcusable, et toute réforme de l'enseignement serait incomplète si elle ne posait pas en principe que tous ceux que leur métier mettra en contact avec la matière ou qui auront à travailler, d'une façon ou d'une autre, sur les êtres vivants, c'est-à-dire les ingénieurs, les architectes... d'une part; les médecins, les agronomes et aussi, il faut le dire, les professeurs et les juristes d'autre part, devront recevoir un minimum de culture scientifique. Ce minimum, qui devra être impitoyablement exigé, comportera obligatoirement toutes les disciplines scientifiques, inégalement dosées selon les carrières préparées. La culture générale du biologiste, dont l'objet d'étude, quel qu'il soit, est toujours complexe, devra être particulièrement poussée et comportera, avec des éléments suffisants d'analyse, une large part de statistique.

P E D A G O G I A

RESULTADOS DE UM EXAME DE MATEMÁTICA - 1.º CICLO

por Maria Teodora Alves

No n.º 29 da «Gazeta de Matemática» foi prometido que seria feito o estudo dos pontos para as provas escritas de Matemática do exame do 1.º ciclo no Liceu de Passos Manuel no ano lectivo 1945-46.

Os elementos estatísticos, relativos a êsse exame, que determinámos, constam do quadro seguinte:

<i>Aritmética e Álgebra</i>		<i>Geometria</i>
Nota mínima	0	0
Nota máxima	18	19
média	9.9	11.2
σ	4.08	3.99
coeficiente de variação	0.41	0.35
Q_3	12.3	13.8
mediana	9.1	10.9
Q_1	6.7	8.0
Q	2.8	2.9
r		0.53
$N=236$		

Em $N=236$ estão incluídos 42 alunos internos do Liceu de Passos Manuel e 194 externos e destes, uns, do ensino particular em estabelecimento, outros do ensino particular individual e, ainda, maiores e emancipáveis.

Em Aritmética e Álgebra a distribuição das notas fez-se de 0 a 18 e em Geometria de 0 a 19. Isto é, os pontos que serviram às duas provas e os critérios de classificação adoptados pelos respectivos professores, distribuíram os alunos por quasi tóda a escala oficial de classificação (0 a 20).

A média 9.9 na prova de Aritmética e Álgebra difere sensivelmente da média 11.2 na de Geometria, o que mostra que o grupo de alunos que se apresentou a exame tinha melhor preparação em geometria do que em Aritmética e Álgebra, ou que o ponto de Aritmética e Álgebra foi sensivelmente mais difícil do que o de Geometria.

É ocasião de lembrar que o número de lições de