

CHARLES LENAY

Galton, eugénisme et statistique

Publications de l'Institut de recherche mathématiques de Rennes, 1987-1988, fascicule 2
« Science, histoire et société », , p. 170-175

http://www.numdam.org/item?id=PSMIR_1987-1988__2_170_0

© Département de mathématiques et informatique, université de Rennes,
1987-1988, tous droits réservés.

L'accès aux archives de la série « Publications mathématiques et informatiques de Rennes » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

GALTON, EUGENISME ET STATISTIQUE

Charles LENAY

Il s'agit ici de montrer comment les enjeux de l'eugénisme de Francis Galton (1822-1911) le conduisirent à développer une analyse probabiliste des relations héréditaires, analyse qui allait déboucher sur la découverte des méthodes de la régression et de la corrélation.

Très frappé en 1859 par la lecture de l'Origine des espèces de Darwin, Galton pensa avoir trouvé sa mission: fonder une nouvelle religion basée sur le savoir scientifique. De là découlait tout un programme de recherche: mettre en évidence un certain déterminisme évolutif, déterminer la loi de l'hérédité des facultés intellectuelles, et trouver l'explication de leurs variations. C'est à l'aide de la courbe binomiale, dont il venait de découvrir l'importance grâce aux travaux de Quételet, que Galton pensa pouvoir maîtriser le hasard des variations dans la sélection naturelle. Mais en introduisant pour la première fois les probabilités dans les explications biologiques, Galton allait rencontrer une série de problèmes difficiles. Et tout d'abord, comment expliquer les tirages au hasard "naturels" qui donnent une répartition binomiale d'un caractère comme la taille ?

La théorie de la pangénèse que Darwin proposa en 1868 parut à Galton parfaitement adaptée à cet usage. L'hérédité correspondrait à la transmission d'une multitude de gemmules dont certains restent à l'état latent tandis que d'autres constituent les caractères de l'enfant. Les gemmules latents transmis à la descendance expliqueraient la réapparition de caractères ancestraux que l'on observe souvent dans la suite des générations, ce que les biologistes appellent la "réversion". Galton raisonnait

toujours de façon très concrète. Pour se représenter la formation d'une courbe en cloche dans la nature, il cherchait donc comment, mécaniquement, pouvait s'effectuer l'équivalent d'une prise d'échantillon dans une urne. L'ensemble des gemmules transmis à chaque génération constituerait une "stirpe", c'est-à-dire l'équivalent d'une urne dans laquelle seraient tirés au hasard les gemmules qui constituent l'individu.

Après bien des efforts pour rassembler des données expérimentales suffisantes, Galton se trouva enfin, en 1877, en possession des résultats d'une importante expérience qu'il avait organisée sur l'auto-fécondation du pois de senteur. Il pu ainsi décrire une tendance des enfants, de parents dont les tailles sont éloignées de la moyenne de la population, à se rapprocher plus de cette moyenne, phénomène qu'il dénomma "reversion" puisqu'il lui semblait que ce retour représentait la réapparition des caractéristiques spécifiques ancestrales. Bien qu'il rencontrât quelques problèmes, les instruments théoriques qu'il avait développés dans ses discussions proprement biologiques lui permirent de donner une interprétation originale de la relation entre les tailles des graines mères et celles de leurs graines filles. En effet, à l'aide de sa théorie de la stirpe il pouvait faire le raisonnement suivant: la répartition des graines mères suivant une courbe binomiale devrait résulter de deux phénomènes aléatoires cumulés. D'une part, elles sont issues de stirpes différentes dont la répartition devrait être binomiale. D'autre part, chaque graine est le résultat d'un tirage au hasard dans une de ces stirpes. Par exemple, la taille des graines résulterait de la sommation des effets de gemmules de deux types, ceux qui tendent à donner une grande taille, et ceux qui tendent plutôt à donner une petite taille. La grande taille d'une graine parentale peut être soit le résultat d'un tirage moyen dans une

stirpe contenant beaucoup de gemmules "grande taille", soit le résultat d'un tirage exceptionnel dans une stirpe moyenne. Dès lors, la déviation par rapport à la moyenne de la taille d'une graine fille peut se comprendre, soit comme le résultat d'un tirage probable dans la stirpe exceptionnelle de sa mère, soit comme le résultat d'un tirage exceptionnel dans la stirpe moyenne de sa mère, soit encore comme un mélange de ces deux causes. On voit donc comment se combinent deux niveaux de fluctuations autour de la moyenne. Seules les déviations dans la composition des stirpes auront un effet (seulement statistique) sur les variations de la descendance.

En 1889, dans Natural Inheritance, Galton exprimait encore dans les termes de sa théorie de la stirpe, cette partition entre une variation expliquée par la ressemblance familiale et une variabilité propre à la descendance:

" Two causes affect family resemblance; the one is Heredity, the other is Circumstance. That which is transmitted is only a sample taken partly through the operation of "accidents", out a store of otherwise unused material, and circumstance must always play a large part in the selection of the sample."

C'est très lentement, au fur et à mesure de l'arrivée de nouvelles données, en particulier sur l'homme, que Galton en vint à affiner ses analyses. Il ne possédait que de faibles connaissances mathématiques, mais il comprit que l'ensemble du diagramme de relation entre parents et enfants devait pouvoir être retrouvé à partir de trois données seulement: la réversion, la variabilité

1) Galton (F.), Natural Inheritance, 1889, infra., p. 195.

totale de la population, et la variabilité dans la descendance d'une classe parentale. C'est ce que confirma le mathématicien J.D. Hamilton. Mais pour Galton, toutes ces considérations restaient spécifiquement biologiques. Ce qui était mis en évidence par la réversion c'était seulement une loi de l'hérédité.

Pourtant, les explications biologiques attachées à la réversion lui paraissaient de plus en plus difficile à maintenir pour rendre compte de tous les phénomènes statiques observés. C'est certainement pourquoi il changea, en 1885, le terme de "reversion" par celui de "régression" pour exprimer le retour de la descendance vers la moyenne populationnelle. Et c'est seulement après 1888, alors qu'il était chargé de présenter en Angleterre les travaux d'Alphonse Bertillon, que Galton réalisa que la méthode qu'il avait mise au point pour décrire les relations de parenté pouvaient tout aussi bien s'appliquer à la description des corrélations entre les parties d'un même organisme. Ce n'est qu'à cette occasion qu'il découvrit soudain la généralité de ses concepts de régression et de corrélation.

Néanmoins, le fondement de ces techniques statistiques consistait en une partition de la variation, en une part "expliquée" et une part restant aléatoire. Cette idée avait été comprise à partir du cas concret de l'hérédité: la part expliquée correspond à la régression, c'est-à-dire ce qui, dans la déviation des enfants, était dû à celle de leurs parents. L'autre part de la déviation serait due à une variabilité indistincte.

Une telle découverte n'était pas le résultat d'une quelconque habileté mathématique de la part de Galton. Dès qu'il rencontrait une difficulté dans ce domaine, il s'adressait à des

amis mathématiciens.² D'ailleurs, le problème qu'il avait posé à Dickson n'était pas très difficile du point de vue mathématique. Ce que découvrit Galton, ce fut donc surtout une certaine façon d'apprécier le rapport entre deux variables statistiques en calculant le taux de variation expliqué. Cette idée nous paraît maintenant évidente, mais beaucoup de ses premiers lecteurs doutèrent de sa validité ou de son utilité. Ainsi, dans ses mémoires, Galton s'étonnait encore de l'incompréhension qu'il avait rencontrée chez certains mathématiciens:

"....certain applications of the Gaussian Law, which, for some reasons that I could never clearly perceive, seemed for a long time to be comprehended with difficulty by mathematicians. [.....] They were unnecessarily alarmed lest the well-known rules of Inverse Probability should be unconsciously violated which they never were. I could give a striking case of this, but abstain because it would seem deprecatory of a man whose mathematical powers and hability were far in excee of my own. Still, he was quite wrong. The primary objects of the Gaussian Law of error were exactly opposed, in one sense, to those to which I applied them. They were to get rid of, or to provide a just allowance for errors. But those errors or deviations were the very things I wanted to preserve to know about."³

2) Au long de sa vie, Galton collabora ainsi avec la plupart des mathématiciens ou philosophes anglais qui s'intéressaient aux statistiques. Citons: John Venn, Herbert Spencer, William Farr, Francis Edgeworth, Henry W. Watson, S.H. Burbury, Donald MacAlister, J.D. Hamilton Dickson, W.F. Sheppard, et finalement Karl Pearson.

3) Galton (F.), Memories of My Life, London, Methuen, 1908.

La différence est fondamentale entre la conception de la "loi des erreurs" et celle de la "loi des déviations par rapport à la moyenne" de Galton. Dans le premier cas, le hasard sous-jacent est un artéfact qu'il faudrait éliminer. Dans le second cas, les variations sont l'objet de l'étude, elles sont donc considérées comme réelles, et à tort ou à raison, on fait comme si le hasard était inclu dans le monde des phénomènes à étudier.

On comprend donc que la méthode des corrélations ne fut vraiment développée que dans ce cadre biologique, et non à partir d'une science comme l'astronomie où pourtant, Auguste Bravais, avait donné dès 1846, une formule mathématique semblable pour démêler les observations variables dépendant de différents observatoires aux positions imprécises.

Bibliographie

- 1855: The Art of Travel; or, Shifts and Contrivances Available in Wild Countries, London, J. Murray, 1855, réimprimé in David and Charles Reprint, 1971.
- 1865: "Hereditary Talent and Character", publié en deux parties, MacMillan's Magazine, XII, juin puis août 1865, pp. 157-166, 318-327
- 1869: Hereditary Genius: an Inquiry into its Laws and Consequences, London, MacMillan, 1869; seconde ed. 1892; troisième ed., Watts, 1950. La seconde édition a été réimprimée en 1962 avec une introduction de C. D. DARLINGTON, London, Collins, et New York and Cleveland, World Publishing Company, 1962.
- 1871: "Gregariousness in cattle and in men" Macmillan's Magazine, 23, 1871, pp. 353-57.
- 1872: "On blood-relationship", Proceedings of the Royal Society of London, 20, 1872, pp. 394-402, repris in Nature, 6, 1872, pp. 173-176.
- 1875: "A theory of heredity", Contemporary Review, 27, 1875, pp. 80-95. Traduction française: "Théorie de l'hérédité", La Revue Scientifique, t. X, 1876, pp. 198-205.
- 1877: "Typical Laws of Heredity", Nature, 15, 1877, pp. 492-533; ou bien, Proceedings of the Royal Institution of G.B., 8, 1877, pp. 282-301.
- 1885: "Regression towards mediocrity in hereditary stature", Journal of the Antropological Institute, 15, 1885, pp. 246-263.
- 1888: "Co-relations and their measurement, chiefly from anthropological data", Proceedings of the R.S. of London, 40, 1888, pp. 135-145.
- 1889: Natural Inheritance, London, MacMillan, 1889.